Differential geometry

Manuel Hohmann

April 18, 2024

Contents

Part I

Basic notions

Chapter 1

Manifolds and maps

1.1 Manifolds

The most important concept we will be dealing with in this lecture course is that of a *manifold*. We will follow the treatment in [Lan85, ch. II, § 1]. A manifold can be viewed as a set with an additional structure, called an *atlas*. In order to arrive at its definition, we take a few steps, which will turn out to be useful later. We start with the definition of a *chart*:

Definition 1.1.1 (Chart)**.** Let M be a set. A *chart* of dimension $n \in \mathbb{N}$ on M is a pair (U, ϕ), where $U \subset M$ is a subset of M and $\phi : U \to \mathbb{R}^n$ is an injective function, such that the image $\phi(U) \subset \mathbb{R}^n$ is open.

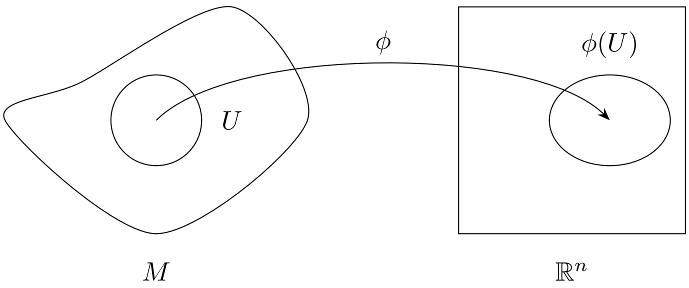


Figure 1.1: Illustration of a chart. The function ϕ establishes a bijection between a set $U \subset M$ and its image $\phi(U) \subset \mathbb{R}^n$.

The concept is illustrated in figure 1.1. Note in particular that ϕ must be *injective*. Further, recall that every function is surjective onto its image. Hence, ϕ establishes a bijection between U and the image $\phi(U)$. This is required for the following definition:

Definition 1.1.2 (Compatibility between charts)**.** Let M be a set and (U_1, ϕ_1) and (U_2, ϕ_2) charts of dimension $n \in \mathbb{N}$ on M. We call these charts *compatible of class* C^k if and only if the following conditions are satisfied:

- The images $\phi_1(U_1 \cap U_2) \subset \mathbb{R}^n$ and $\phi_2(U_1 \cap U_2) \subset \mathbb{R}^n$ are open sets.

- If $U_1 \cap U_2 \neq \emptyset$, then the functions

$$\phi_{12} : \phi_2(U_1 \cap U_2) \to \phi_1(U_1 \cap U_2)\,, \qquad \phi_{12} = \phi_1|_{U_1 \cap U_2} \circ \phi_2^{-1}|_{\phi_2(U_1 \cap U_2)}\,, \tag{1.1.1a}$$
$$\phi_{21} : \phi_1(U_1 \cap U_2) \to \phi_2(U_1 \cap U_2)\,, \qquad \phi_{21} = \phi_2|_{U_1 \cap U_2} \circ \phi_1^{-1}|_{\phi_1(U_1 \cap U_2)}\,, \tag{1.1.1b}$$

are of class C^k, i.e., they are k times continuously differentiable.

We also introduce the following notion:

Definition 1.1.3 (Transition function)**.** Let M be a set and (U_1, ϕ_1) and (U_2, ϕ_2) compatible charts on M such that $U_1 \cap U_2 \neq \emptyset$. The functions ϕ_{12} and ϕ_{21} defined in definition 1.1.2 are called the *transition functions* between these two charts.

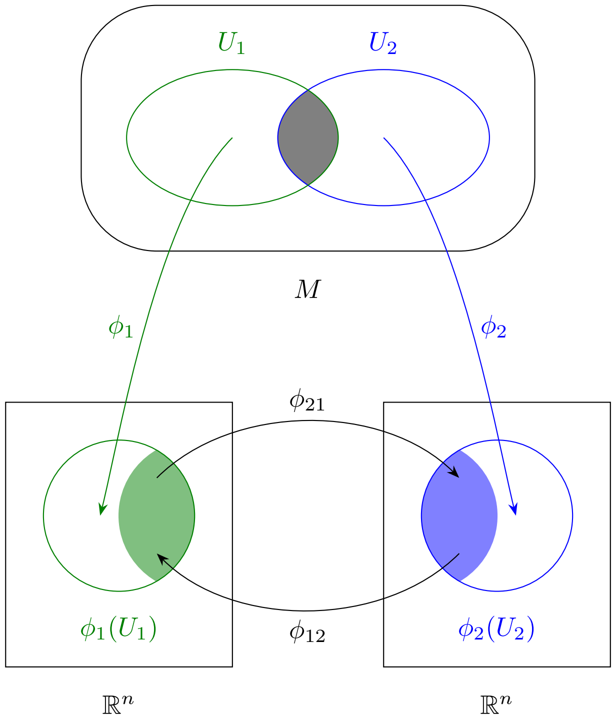


Figure 1.2: Illustration of a transition function. The functions ϕ_{12} and ϕ_{21} are bijections between $\phi_1(U_1 \cap U_2)$ and $\phi_2(U_1 \cap U_2)$.

It follows immediately from their definition that the transition functions ϕ_{12} and ϕ_{21} are bijective and inverses of each other. For our purposes it will be most useful to adopt the following definition of an atlas:

Definition 1.1.4 (Atlas)**.** Let M be a set. An *atlas* $\mathcal{A}$ of class C^k and dimension n on M is a collection of charts (U_i, ϕ_i) of M of dimension n, where $i \in \mathcal{I}$ and $\mathcal{I}$ is an arbitrary index set, such that the following properties hold:

- The sets U_i cover M:

$$\bigcup_{i \in \mathcal{I}} U_i = M \,. \tag{1.1.2}$$

- Any two charts (U_i, ϕ_i) and (U_j, ϕ_j) are compatible of class C^k.

We need to make a few remarks on this definition. First, what we have defined is also called a *real* atlas, since the target space of all functions ϕ_i is the real vector space $\mathbb{R}^n$. Second, a few particular types of atlases have their own names. In this lecture we will deal mostly with *smooth* or C^∞-atlases, by demanding that all transition functions are smooth, i.e., that they are continuous and infinitely often continuously differentiable. A less strict definition would have been that of a *topological* or C^0-atlas, where the transition functions only need to be continuous. However, in physics it is often convenient to assume that everything is smooth, and so we will stick to this assumption. We now further define:

Definition 1.1.5 (Maximal atlas)**.** An atlas $\mathcal{A}$ of class C^k on a set M is called *maximal* if there exists no further chart (U, ϕ) on M which is compatible of class C^k with all charts in $\mathcal{A}$ and which is not already contained in $\mathcal{A}$.

Note that any atlas $\mathcal{A}$ defines a maximal atlas $\bar{\mathcal{A}}$, which contains all charts which are compatible with all charts of $\mathcal{A}$. Finally, we define:

Definition 1.1.6 (Manifold)**.** A *manifold* of class C^k is a set M (its *space*) together with a maximal atlas $\mathcal{A}$ of class C^k on M.

Instead of calling the pair $(M, \mathcal{A})$ a manifold, it is also common to call M itself a manifold and to take $\mathcal{A}$ as implicitly defined. We will make use of this convention and explicitly write the atlas $\mathcal{A}$ only if it is needed.

One may ask why we want a *maximal* atlas in this definition. The answer becomes clear if we ask the question when two manifolds are the same. If we would simply say that a manifold is a set M together with an atlas $\mathcal{A}$, then two different atlases would by definition yield two different manifolds. However, this is usually not what we want, since there is no unique way how to specify an atlas. Instead, we want two manifolds to be "the same" if the atlases defining their geometry are compatible. There are two equivalent solutions to this problem:

- If we demand that two (non-maximal) atlases $\mathcal{A}, \mathcal{A}'$ define the same manifold structure on M if they are compatible, we can define a manifold as a set M together with an equivalence class of compatible atlases. Any atlas then uniquely defines an equivalence class, and thus uniquely defines a manifold. Compatible atlases belong to the same equivalence class, and hence define the same manifold.
- We can use the fact that two atlases are compatible if and only if their completions to maximal atlases agree, $\bar{\mathcal{A}} = \bar{\mathcal{A}}'$. Therefore, an equivalence class of compatible atlases is essentially the same as a maximal atlas, and so we can define a manifold as a set M together with a maximal atlas. Any (non-maximal) atlas can uniquely be completed to a maximal atlas, and these completions agree for compatible atlases, hence yield the same manifold.

For practical purposes, in order to uniquely specify a particular manifold, it is therefore enough to provide a non-maximal atlas, which is then uniquely extended to either a maximal atlas or an equivalence class as described above. In the most simple cases, even a single chart can be sufficient to cover the whole manifold. We make use of this fact in the examples below.

Example 1.1.1 (Euclidean space). The space $M = \mathbb{R}^n$ with atlas

$$\mathcal{A} = \{(\mathbb{R}^n, \mathrm{id}_{\mathbb{R}^n})\} \tag{1.1.3}$$

given by a single chart (U, ϕ), where $U = \mathbb{R}^n$ covers the whole space and $\phi = \mathrm{id}_{\mathbb{R}^n}$ is the identity function, is a smooth manifold of dimension n.

Example 1.1.2 (Complex numbers). The space $M = \mathbb{C}$ of complex numbers with atlas $\mathcal{A} = \{(U, \phi)\}$, where $U = \mathbb{C}$ and

$$\begin{array}{rccc} \phi & : & \mathbb{C} & \to & \mathbb{R}^2 \\ & & x + iy & \mapsto & (x, y) \end{array} \tag{1.1.4}$$

is a smooth manifold of dimension 2.

Example 1.1.3 (Quaternions). The space $M = \mathbb{H}$ of quaternions with atlas $\mathcal{A} = \{(U, \phi)\}$, where $U = \mathbb{H}$ and

$$\begin{array}{rcccc} \phi & : & \mathbb{H} & \to & \mathbb{R}^4 \\ & & x + iy + jz + ku & \mapsto & (x, y, z, u) \end{array} \tag{1.1.5}$$

is a smooth manifold of dimension 4.

Example 1.1.4 (Punctured space). The "punctured" space $M = \mathbb{R}^n \setminus \{(0, \ldots, 0)\}$ with atlas $\mathcal{A} = \{(U, \phi)\}$, where $U = M$ and $\phi = \mathrm{id}_M$, is a smooth manifold of dimension n.

Of course, also for the manifolds above one can find an infinite number of further charts, which are compatible with the single chart given above, and thus belong to the same maximal atlas, hence the same manifold. For example, one may simply consider a smaller open domain $V \subset U$, and restrict ϕ to $\phi|_V : V \to \mathbb{R}^n$, or compose ϕ with a bijective function $\psi : \mathbb{R}^n \to \mathbb{R}^n$, such that ψ and its inverse are sufficiently often differentiable of class C^k, and consider $\psi \circ \phi$ instead of ϕ. In general, however, a single chart is not sufficient to cover the whole manifold. An important example, which we encounter often in these lecture notes, is the following.

Example 1.1.5 (Circle). Consider the set $M = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 = 1\}$ (the unit circle), as well as the function

$$\begin{array}{rcccc} \psi & : & \mathbb{R} & \to & \mathbb{R}^2 \\ & & u & \mapsto & (\cos u, \sin u) \end{array} . \tag{1.1.6}$$

Clearly, for all $u \in \mathbb{R}$ one has $\psi(u) \in M$, since $\cos^2 u + \sin^2 u = 1$. Note that ψ is not invertible, since

$$\psi(u) = \psi(u + 2\pi n) \tag{1.1.7}$$

for all $u \in \mathbb{R}$ and $n \in \mathbb{Z}$. However, if we restrict ψ to an open interval $(a, b) \subset \mathbb{R}$ with $0 < b - a \leq 2\pi$, then we find that for each $p \in U_{ab} = \psi((a, b))$ there is a unique $u \in (a, b)$ such that $p = \psi(u)$, and thus a function $\psi_{ab} : U_{ab} \to \mathbb{R}$ which is uniquely defined by

$$\psi \circ \phi_{ab} = \mathrm{id}_{U_{ab}}, \quad \phi_{ab}(p) \in (a, b) \text{ for all } p \in U_{ab} . \tag{1.1.8}$$

This function ϕ_{ab} is injective, since for each $u \in (a, b)$, there is only one $p = \psi(u) \in U_{ab}$ such that $\phi_{ab}(p) = u$. Also the image $\phi_{ab}(U_{ab}) = (a, b)$ is open. Hence, (U_{ab}, ϕ_{ab}) is a chart. This is illustrated in figure 1.3.

Note that a single chart cannot be sufficient, since $\phi_{ab}(U_{ab}) \subset \mathbb{R}$ must be open, and so we cannot include endpoints to close the circle. We need at least two charts, and so we can choose for example the most common charts constructed from the intervals $(0, 2\pi)$ and $(-\pi, \pi)$, for which we have the domains

$$U_1 = M \setminus \{(1, 0)\}, \quad U_2 = M \setminus \{(-1, 0)\}, \tag{1.1.9}$$

so both charts lack exactly one point $(\pm 1, 0)$ on the circle. Their overlap is thus given by

$$U = U_1 \cap U_2 = M \setminus \{(1, 0), (-1, 0)\}. \tag{1.1.10}$$

To check compatibility, we calculate

$$\phi_1(U) = (0, \pi) \cup (\pi, 2\pi), \quad \phi_2(U) = (-\pi, 0) \cup (0, \pi), \tag{1.1.11}$$

which are both open. Further, we have the transition functions

$$\phi_{12} : \phi_2(U) \to \phi_1(U), u \mapsto \begin{cases} u + 2\pi & \text{if } u \in (-\pi, 0), \\ u & \text{if } u \in (0, \pi), \end{cases} \tag{1.1.12}$$

and

$$\phi_{21} : \phi_1(U) \to \phi_2(U), u \mapsto \begin{cases} u & \text{if } u \in (0, \pi), \\ u - 2\pi & \text{if } u \in (\pi, 2\pi), \end{cases} \tag{1.1.13}$$

which is visualized in figure 1.4. Note that despite the apparent discontinuity, these are smooth on their domains, which consist of two connected components, and so the location of the apparent discontinuity is not part of the domain. Hence, the two charts are compatible of class C^∞. Finally, $U_1 \cup U_2 = M$, and thus these two charts form an atlas of class C^∞ of M, and the dimension is 1.

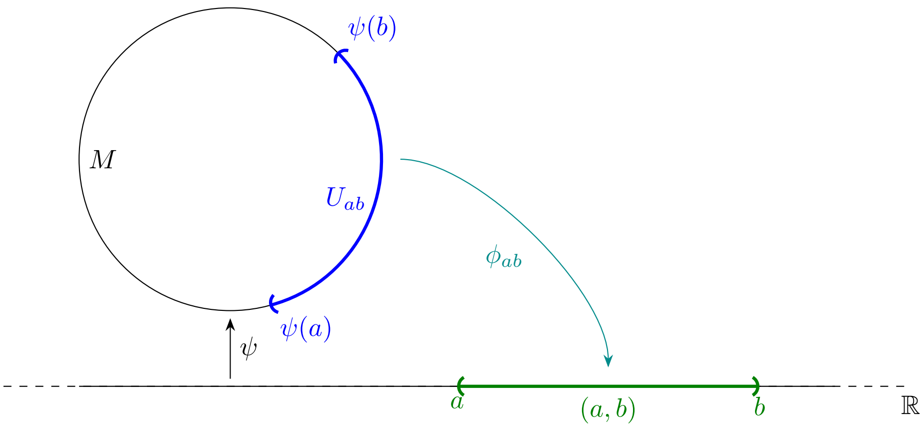


Figure 1.3: A chart of the circle (see example 1.1.5).

The examples above have in common that we started with a set M which is defined as a subset (or all of) a Euclidean, complex or quaternionic space, and we chose charts (U, ϕ) such that the functions ϕ are smooth functions on U. This often leaves the intuitive picture that a manifold *is* a subset of such as space, and that the charts are merely used in order to compare the local geometry to that of $\mathbb{R}^n$. However, this is not necessarily the case, since in the definition of a

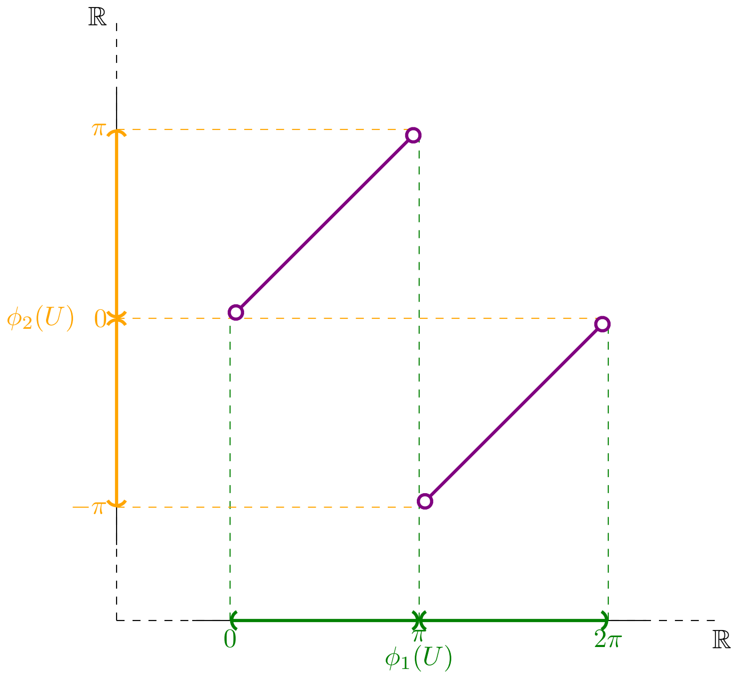


Figure 1.4: Transition functions for two charts on the circle (see example 1.1.5).

manifold, M can be any abstractly defined set, and the manifold structure is fully determined by the charts, even without defining M as or considering it as a subset of some other space. To illustrate this fact, we may return to the example of the circle again.

Example 1.1.6 (Circle as quotient space). Consider the set

$$M = \mathbb{R}/2\pi\mathbb{Z} = \{[x], x \in \mathbb{R}\}, \quad [x] = \left\{y \in \mathbb{R}, \frac{x-y}{2\pi} \in \mathbb{Z}\right\} \tag{1.1.14}$$

of equivalence classes $[x]$ of real numbers $x \in \mathbb{R}$, where we consider two numbers as equivalent when their difference is an integer multiple of 2π. Clearly, there is a one-to-one correspondence between elements of M and the circle we have encountered in the previous example 1.1.5, and we can equivalently define an atlas using this definition of M together with the maps $\tilde{\phi}_{ab}(x) = [x]$, using the same notation as above. This yields the same manifold S^1 without referring to the intuitive picture of a circle embedded in a plane.

We finally remark that in the literature one also encounters slightly different definitions of an atlas and a manifold. Often the space M is a priori assumed to be equipped with a *topology*, and that this topology is Hausdorff, i.e., for any two distinct points $x, y \in M$ there exist open subsets $U \ni x$ and $V \ni y$ of M such that $U \cap V = \emptyset$. We will not assume any a priori topology on M, since we can define the topology from the atlas itself as the coarsest topology on M such that all charts are continuous. Note that this topology will in general not be Hausdorff. However, this will not be relevant for most of the examples we consider, and follows the treatment in [Lan85, ch. II, § 1].

1.2 Charts and coordinates

A special role is given to charts in the application of differential geometry to physics. In this context, a chart together with an assignment of names to the components of $\mathbb{R}^n$ is also called a set of (local) *coordinates*, while a transition function is also called a *change of coordinates*. There are different ways to label coordinates. One possibility is to give an explicit name to each coordinate, such as (x, y, z) or (r, θ, φ) for a chart of a three-dimensional manifold. Another common possibility is to write coordinates as indexed quantities, such as $(x^a, a = 1, \ldots, 3)$, where here the coordinates are named (x^1, x^2, x^3). It is conventional to use upper indices for coordinates - these must not be confused with powers!

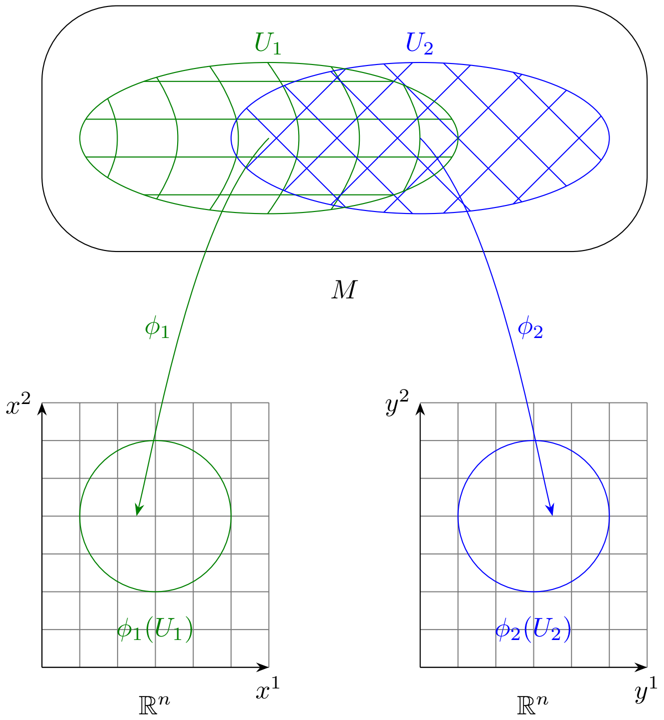


Figure 1.5: Once we give names to the components of $\mathbb{R}^n$, a chart associates *coordinates* to the points of its domain. This association, of course, depends on the chart. On the intersection $U_1 \cap U_2$ of two charts we have usually different coordinates defined by these charts.

A similar notation is used for transition functions. Let the coordinates of the chart (U_i, ϕ_i) be denoted by (x^a) and those of (U_j, ϕ_j) by (x'^a), where $a = 1, \ldots, n$. The transition function ϕ_{ji} is then commonly written as $x'(x)$, and specified in terms of the coordinate functions

$$x'^1(x^1, \ldots, x^n), \ldots, x'^n(x^1, \ldots, x^n)\,. \tag{1.2.1}$$

The requirement that a transition function must be smooth is then expressed by the requirement that all component functions must be continuous, that they must be infinitely often partially differentiable and that all partial derivatives are continuous.

One should be careful here, because in the physics literature one often finds coordinates corresponding to charts which cover *almost*, but not all of M. An example is the description of the two-dimensional sphere S^2 by latitude $-\pi/2 < \theta < \pi/2$ and longitude $0 < \varphi < 2\pi$, which does not include the poles and the zero meridian. It is also conventional to "cure" this problem by redefining the coordinate range in the form $-\pi/2 \leq \theta \leq \pi/2$ and $0 \leq \varphi < 2\pi$, but this is not even a chart anymore, since it does not define a function onto an open subset of $\mathbb{R}^2$! The missing / added points here are called "coordinate singularities". One can work with this description, but one must pay attention to all the possible illnesses that occur at coordinate singularities and know how to deal with them (namely, by using proper charts). The following example gives such charts for the sphere S^2.

Example 1.2.1 (Sphere). Consider a sphere of radius 1 centered around the origin embedded in Euclidean space $\mathbb{R}^3$. Using Cartesian coordinates x^1, x^2, x^3 on $\mathbb{R}^3$, the set is given by

$$M = \{(x^1, x^2, x^3) \in \mathbb{R}^3, (x^1)^2 + (x^2)^2 + (x^3)^2 = 1\} \,. \tag{1.2.2}$$

We now construct two charts of the sphere as follows. Consider first the set

$$U_1 = M \setminus \{(0, 0, -1)\} \,, \tag{1.2.3}$$

i.e., the sphere without the south pole $p_1 = (0, 0, -1)$. Given a point $p = (x^1, x^2, x^3) \in U_1$, there exists a unique line passing through p and p_1. This line intersects the plane $x^3 = 0$ exactly once in the point

$$\left(\frac{x^1}{1 + x^3}, \frac{x^2}{1 + x^3}, 0\right) \,. \tag{1.2.4}$$

This allows us to define a function

$$\begin{array}{cccc} \phi_1 & : & U_1 & \to & \mathbb{R}^2 \\ & & (x^1, x^2, x^3) & \mapsto & (v^1, v^2) = \left(\frac{x^1}{1+x^3}, \frac{x^2}{1+x^3}\right) \end{array} \,, \tag{1.2.5}$$

thus defining coordinates (v^1, v^2). One easily checks that $\phi_1(U_1) = \mathbb{R}^2$, and hence is open, so that (U_1, ϕ_1) is a chart. Similarly, we can remove the north pole $p_2 = (0, 0, 1)$ instead, to construct a chart consisting of the open set

$$U_2 = M \setminus \{(0, 0, 1)\} \,, \tag{1.2.6}$$

together with the function

$$\begin{array}{cccc} \phi_2 & : & U_2 & \to & \mathbb{R}^2 \\ & & (x^1, x^2, x^3) & \mapsto & (\tilde{v}^1, \tilde{v}^2) = \left(\frac{x^1}{1-x^3}, \frac{x^2}{1-x^3}\right) \end{array} \,. \tag{1.2.7}$$

We then check the compatibility of the two charts. First note that the images

$$\phi_1(U_1 \cap U_2) = \phi_2(U_1 \cap U_2) = \mathbb{R}^2 \setminus \{(0, 0)\} \tag{1.2.8}$$

of the intersection $U_1 \cap U_2$ are open sets. Finally, consider the transition function

$$\begin{array}{cccc} \phi_{21} & : & \mathbb{R}^2 \setminus \{(0,0)\} & \to & \mathbb{R}^2 \setminus \{(0,0)\} \\ & & (v^1, v^2) & \mapsto & (\tilde{v}^1, \tilde{v}^2) = \left(\frac{v^1}{(v^1)^2+(v^2)^2}, \frac{v^2}{(v^1)^2+(v^2)^2}\right) \end{array} \,. \tag{1.2.9}$$

One easily checks that this is smooth. The same holds for its inverse ϕ_{12}, and so the two charts (U_1, ϕ_1) and (U_2, ϕ_2) are compatible of class C^∞. They cover M, i.e., $U_1 \cup U_2 = M$, and so they define a smooth atlas on M. Hence, we obtain a smooth manifold of dimension 2. The manifold is usually denoted S^2.

In this lecture course we will use coordinates whenever it is necessary, which is the case for explicit calculations of examples (and which is also the most important application of coordinates in physics). Sometimes we will introduce a particular set of coordinates, sometimes we will simply assume that some set of coordinates is given, which we do not specify any further. But most of the time, whenever it is possible, we will avoid the use of coordinates.

1.3 Maps

The second very important notion we need is that of a *map* between manifolds. We define:

Definition 1.3.1 (Map). Let M, N be manifolds of class C^k. A *map* of class C^k from M to N is a function $f : M \to N$ such that for each point $p \in M$ exist charts (U, ϕ) of M and (V, χ) on N such that:

1. $p \in U$ and $f(U) \subset V$.
2. The function $\chi \circ f \circ \phi^{-1} : \phi(U) \to \chi(V)$ is of class C^k.

We denote the space of all maps of class C^k between M and N by $C^k(M, N)$.

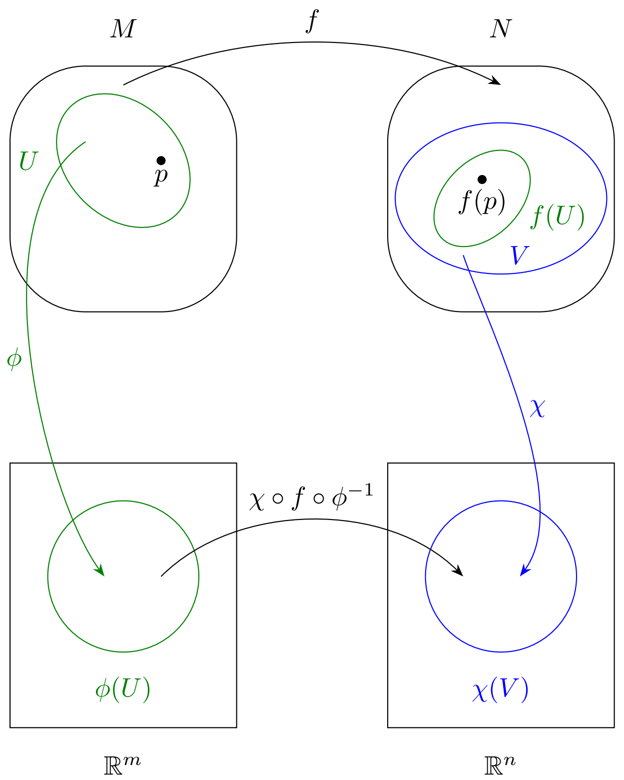


Figure 1.6: Illustration of a map of class C^k and the charts that appear in its definition 1.3.1.

Instead of "map" also the term "mapping" is often found in the literature.

One often expresses maps with the help of coordinates (charts). Using the notation from the definition above, denote the components or $\mathbb{R}^m$ by $(x^i, i = 1, \ldots, m)$ and the components of $\mathbb{R}^n$ by $(y^\mu, \mu = 1, \ldots, n)$, where m and n are the respective dimensions of M and N. Consider a point $p \in M$ and charts (U, ϕ) of M and (V, χ) of N such that $p \in U$ and $f(U) \subset V$. Then there exists a function

$$\chi \circ f \circ \phi^{-1} : \phi(U) \to \mathbb{R}^n \,, \tag{1.3.1}$$

where $\phi(U) \subset \mathbb{R}^m$. This function may be regarded as the coordinate representation of the map f. It assigns to a tuple $(x^1, \ldots, x^m) \in \phi(U)$ a tuple $(y^1, \ldots, y^n)$, so one may write it in the

form
$$(\chi \circ f \circ \phi^{-1})(x^1, \ldots, x^m) = (y^1, \ldots, y^n) . \tag{1.3.2}$$
Often in the literature the charts are omitted from the notation, and so one often writes f instead of $\chi \circ f \circ \phi^{-1}$ in the formula above; however, note that this is not strictly correct. The map f assigns elements of N to elements of M. These elements can be represented by tuples in $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively, with the help of a chart, but they are *not* identical to these tuples, and f does not assign tuples to tuples. The coordinate expression of f depends on the choice of the charts ϕ and χ.

To illustrate the notion of a map, and how to show its differentiability, let us consider the following simple class of maps.

Definition 1.3.2 (Constant map)**.** Let M, N be manifolds and $q \in N$. The *constant map* with image q is the map $f_q : M \to N$ defined by $f_q(p) = q$ for all $p \in M$.

Now it is easy to show the following.

Theorem 1.3.1. *The constant map $f_q : M \to N, p \mapsto q$ is a smooth map for all smooth manifolds M, N and $q \in N$.*

Proof. Let (V, χ) be a chart around $q \in N$. For any $p \in M$, we can choose a chart (U, ϕ). Clearly, this satisfies
$$f_q(U) = \{q\} \subset V , \tag{1.3.3}$$
by choice of (V, χ). We then consider the function
$$\chi \circ f_q \circ \phi^{-1} \quad : \quad \begin{array}{ccc} \phi(U) & \to & \chi(V) \\ x & \mapsto & \chi(q) \end{array} , \tag{1.3.4}$$
which is now a constant function between subsets of Euclidean spaces. However, any constant function between Euclidean spaces is infinitely often differentiable, and thus smooth. ■

Often one has to compose maps. The following statement guarantees that this is indeed allowed:

Theorem 1.3.2. *Let L, M, N be manifolds of class C^k and $f : L \to M$ and $g : M \to N$ maps of class C^k between them. Then $g \circ f : L \to N$ is a map of class C^k.*

Proof. We denote $l = \dim L$, $m = \dim M$ and $n = \dim N$. From the fact that both f and g are of class C^k follows that for every point $p \in L$ there exists charts as follows:

1. (U, ϕ) of L and (V, χ) of M such that:
   (a) $p \in U$,
   (b) $f(U) \subset V$,
   (c) $\chi \circ f \circ \phi^{-1} : \phi(U) \to \chi(V)$ is of class C^k;
2. (U', ϕ') of M and (V', χ') of N such that:
   (a) $f(p) \in U'$,
   (b) $g(U') \subset V'$,
   (c) $\chi' \circ g \circ \phi'^{-1} : \phi'(U') \to \chi'(V')$ is of class C^k.

Note that straight from the definition of smoothness of f and g, the two charts (V, χ) and (U', ϕ') of M are not guaranteed to be the same; however, they are guaranteed to be compatible of class C^k, since M is a manifold of class C^k. This means in particular that $\chi(V \cap U') \subset \mathbb{R}^m$ is open. Since $\chi \circ f \circ \phi^{-1}$ is of class C^k, and thus in particular continuous, also its preimage is continuous, i.e., the set

$$\begin{aligned}
\left(\chi \circ f \circ \phi^{-1}\right)^{-1}(\chi(V \cap U')) &= \{u \in \phi(U), \chi(f(\phi^{-1}(u))) \in \chi(V \cap U')\} \\
&= \{u \in \phi(U), f(\phi^{-1}(u)) \in V \cap U'\} \\
&= \{u \in \phi(U), f(\phi^{-1}(u)) \in U'\} \\
&= \{\phi(q), q \in U, f(q) \in U'\} \\
&= \phi(U \cap f^{-1}(U'))\,,
\end{aligned} \tag{1.3.5}$$

where in the first line we have used that χ is injective and can thus be omitted, in the second line that $f(\phi^{-1})(u) \in f(U) \subset V$, and finally that ϕ is a bijection from U to $\phi(U)$. Now we can then define $\tilde{U} = U \cap f^{-1}(U')$, as well as $\tilde{\phi} = \phi|_{\tilde{U}}$. We have already shown that

$$\tilde{\phi}(\tilde{U}) = \phi(\tilde{U}) = \phi(U \cap f^{-1}(U')) \tag{1.3.6}$$

is open; further, $\tilde{\phi}$ is injective and compatible of class C^k with ϕ, as it is simply a restriction. Hence, $(\tilde{U}, \tilde{\phi}$ is another chart of L. Now we have

$$g(f(\tilde{U})) \subset g(f(f^{-1}(U'))) = g(U') \subset V'\,. \tag{1.3.7}$$

We can thus consider the function

$$\chi' \circ g \circ f \circ \tilde{\phi}^{-1} = \underbrace{\chi' \circ g \circ \phi'^{-1}} \circ \underbrace{\phi' \circ \chi^{-1}} \circ \underbrace{\chi \circ f \circ \tilde{\phi}^{-1}}\,. \tag{1.3.8}$$

Here we see that the two outer parts are of class C^k since f and g are of class C^k, and the inner part is of class C^k, since the two charts on M are compatible of class C^k. Hence, also $g \circ f$ is of class C^k. The construction is illustrated in figure 1.7. ■

We finally introduce a particularly useful type of map:

Definition 1.3.3 (Diffeomorphism)**.** A map $f : M \to N$ which is bijective and whose inverse $f^{-1} : N \to M$ is again a map, is called a *diffeomorphism*. If such a diffeomorphism exists, the manifolds M, N are called *diffeomorphic*.

To discuss a simple example, we define the following map.

Definition 1.3.4 (Identity map)**.** Let M be a manifold of class C^k. The map $\mathrm{id}_M : M \to M, p \mapsto p$ is called the *identity* on M.

One may already assume that the identity is a diffeomorphism. We will show this as follows.

Theorem 1.3.3. *Let M be a manifold of class C^k. The identity id_M is a diffeomorphism of class C^k.*

Proof. Let $p \in M$, and pick a chart (U, ϕ) with $p \in U$. Obviously, one has $\mathrm{id}_M(U) = U$. In this chart, the identity is expressed by $\phi \circ \mathrm{id}_M \circ \phi^{-1} = \mathrm{id}_{\phi(U)}$, which is the identity on a subset $\phi(U)$ of Euclidean space. Since this is arbitrarily often differentiable, it follows that id_M is of class C^k. Further, the identity is bijective with inverse $\mathrm{id}_M^{-1} = \mathrm{id}_M$, and its inverse is thus obviously also of class C^k. Hence, id_M is a diffeomorphism of class C^k. ■

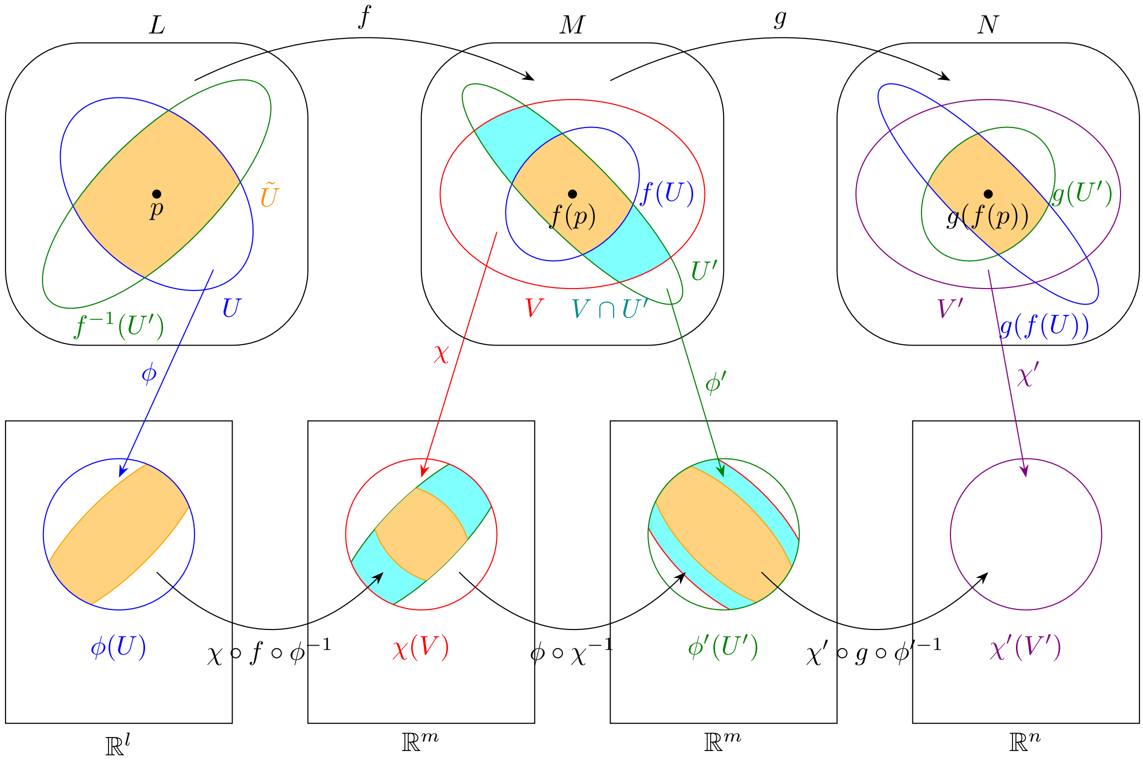


Figure 1.7: Illustration of the composition of two maps of class C^k and the charts that appear in theorem 1.3.2.

Note that we concluded on class C^k here only, and not on C^∞. The reason is that one may also choose different charts on the domain and codomain (even though these are the same manifold in this case), and for a C^k-manifold these are only required to be compatible of order C^k. Hence, only the differentiability up to C^k can be guaranteed independently of the choice of the charts. We also remark that there exist many more diffeomorphisms on a manifold, as we will see later. However, different manifolds are in general not diffeomorphic. In fact, manifolds are often considered as being "the same" if and only if they are diffeomorphic. This is possible because being diffeomorphic is an equivalence relation.

Another common type of map, which is sufficiently notable to deserve its own name, is obtained when the domain is the set $\mathbb{R}$ of real numbers:

Definition 1.3.5 (Curve)**.** A *curve* of class C^k on a manifold M is a map $\gamma : \mathbb{R} \to M$ of class C^k.

If one fixes a chart (U, ϕ) of M with coordinates (x^a), one may denote the components of $(\phi \circ \gamma)(t)$ by $(\phi \circ \gamma)^a(t)$. Also here the less lengthy, but not strictly correct notation $\gamma^a(t)$ is conventionally used. Smoothness of the curve γ then simply means that the components γ^a must be smooth functions.

We also point out particular maps, whose codomain is given either by the real or complex numbers:

Definition 1.3.6 (Real function)**.** A *real function* of class C^k on a manifold M is a map $f : M \to \mathbb{R}$ of class C^k.

Definition 1.3.7 (Complex function)**.** A *complex function* of class C^k on a manifold M is a map $f : M \to \mathbb{C}$ of class C^k.

These are most commonly used. One often finds the notation $C^k(M)$ as a shorthand for either $C^k(M, \mathbb{R})$ or $C^k(M, \mathbb{C})$. However, one may easily get confused whether this should denote the space of real or complex functions, and so we will avoid this notation.

Note that in physics one usually works with smooth (C^∞) manifolds and maps. Hence, in the following we will mostly omit the explicit "of class C^k" or "of class C^∞", and assume that maps and manifolds are smooth, unless we explicitly say otherwise.

1.4 Germs

There are situations in which only the local properties of a function are relevant, which are described by the behavior of that function in the neighborhood of a particular point. Any other function which behaves identically around the same point would then have the same local properties. In order to work with such local properties in a mathematically rigorous way, one therefore introduces a few notions, starting with the following.

Definition 1.4.1 (Local map)**.** Let M, N be manifolds of class C^k and $p \in M$. A *local map* of class C^k at p is a pair (U, f), where $U \subset M$ is an open set containing p and $f : U \to N$ is a map. The set of all local maps of class C^k around p is denoted $C^k_p(M, N)$.

This definition takes one step towards capturing only the local properties of a function f around a point p. We only need to care about its values in some neighborhood U, and it does not even have to be defined outside of U. However, there are still two issues which make working with $C^k_p(M, N)$ less convenient:

- Given a local function $(U, f) \in C^k_p(M, N)$, there are (in general infinitely) many other local functions which have exactly the same local properties at p. For example, one may simply choose another open subset $U' \subset U$ such that $p \in U'$ and consider the restriction $f|_{U'}$. Then $(U', f|_{U'})$ clearly describes the same local behavior around p. Also any other local function (V, g), which differs from f only outside some neighborhood of p, shares the same local behavior.
- Often the target manifold N carries some additional structure, such as that of a group or a vector space, and then the space $C^k(M, N)$ of global functions inherits this structure, by carrying out operations pointwise on M. This inheritance is lost for local functions, since they are, in general, defined on different domains. For example, one may add two real functions $f, g \in C^k(M, \mathbb{R})$ on a manifold, but this is not possible for two local functions $(U, f), (V, g) \in C^k_p(M, \mathbb{R})$ defined on different domains $U \neq V$.

In order to solve these two issues, one introduces another notion as follows.

Definition 1.4.2 (Germ)**.** Let M, N be manifolds of class C^k, $p \in M$ and $(U, f) \in C_p^k(M, N)$ a local function of class C^k at p. The *germ* of (U, f) at p is the set of all local functions $(V, g) \in C_p^k(M, N)$ for which there exists an open set $W \subseteq U \cap V$ containing p such that the restrictions of f and g to W agree, $f|_W = g|_W$. The set of all such germs of class C^k at p is denoted $\mathcal{O}_p^k(M, N)$.

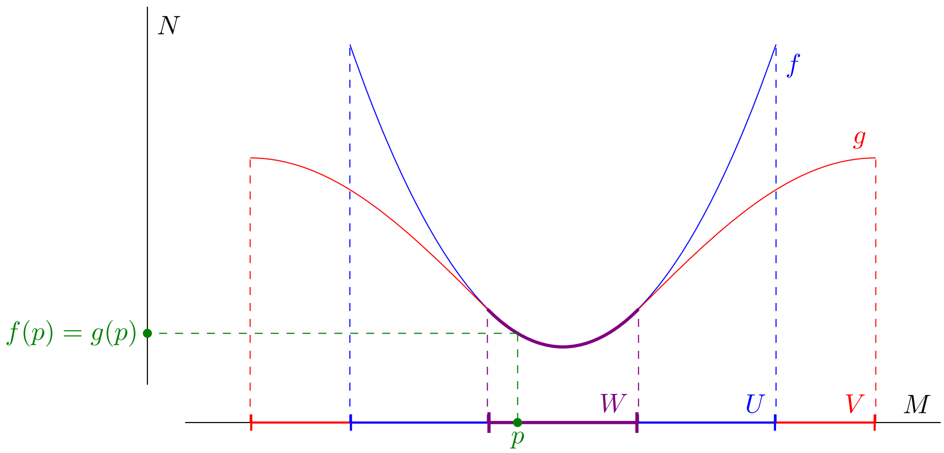


Figure 1.8: Two local functions (U, f) and (V, g) define the same germ at p if and only if there exists an open set $W \subseteq U \cap V$ containing p on which f and g agree, $f|_W = g|_W$. Since $p \in W$, one has in particular also $f(p) = g(p)$.

In other words, one introduces an equivalence relation $\sim$ on the space $C_p^k(M, N)$ of local functions, such that $(U, f) \sim (V, g)$ if and only if there exists an open set $W \subseteq U \cap V$ containing p such that $f|_W = g|_W$. The germ of (U, f) is then the equivalence class

$$[U, f] = \{(V, g) \in C_p^k(M, N) \,|\, (U, f) \sim (V, g)\} . \tag{1.4.1}$$

Of course, one still has to check that this is indeed an equivalence relation, but this is simple and we omit the proof here. We can thus take the set of equivalence classes

$$\mathcal{O}_p^k(M, N) = C_p^k(M, N)/\sim = \{[U, f] \,|\, (U, f) \in C_p^k(M, N)\} . \tag{1.4.2}$$

Note that instead of writing $[U, f]$ it is also conventional to write just $[f]$, since the domain U on which f is defined is not important; it only must be open and contain the point p, but otherwise does not have any influence on the local properties of f at p we are aiming to capture. Also the notation $\mathbf{f}$ instead of $[f]$ is sometimes encountered. Another commonly used notation emphasizes the point p by writing $[f]_p$.

The notion of germs now indeed addresses the two points mentioned above:

- A local function (U, f) defines the same germ as its restriction to a smaller domain $U' \subset U$, as long as also U' is open and contains the point p: $(U, f) \sim (U', f|_{U'})$. Hence, to describe the local properties of f around p, we can consider any open domain around p, no matter how small. This explains why the original domain U is not important and often omitted in the notation $[f]$.
- If the target space N has some algebraic structure, such as that of a group or vector space, then this structure is now inherited by the space of germs. For example, if N is equipped

with a binary operation $\circledast : N \times N \to N$, one may carry this operation to the space of germs as follows. Given two germs $[f], [g] \in \mathcal{O}_p^k(M, N)$ at $p \in M$, one picks representatives $(U, f), (V, g) \in C_p^k(M, N)$. Then, on the intersection $U \cap V$, one applies pointwise the operation $\circledast$ to $f|_{U\cap V}$ and $g|_{U\cap V}$. Finally, one defines $[f] \circledast [g]$ as the germ $[f_{U\cap V} \circledast g_{U\cap V}]$. After checking that this operation does not depend on the choice of the representatives, one has obtained a binary operation $\circledast : \mathcal{O}_p^k(M, N) \times \mathcal{O}_p^k(M, N) \to \mathcal{O}_p^k(M, N)$.

The local properties of f around p are now fully described by its germ $[f]$. For example, the image $f(p)$ does not depend on the choice of the representative. Also "derivatives" (a notion which we yet have to define, unless we resort to charts and functions on $\mathbb{R}^n$) at p only depend on the choice of the germ, but not on the choice of the representative. This includes in particular jets, which we will encounter in chapter 21.

1.5 Partitions of unity

Definition 1.5.1 (Partition of unity)**.** A *partition of unity* on a manifold M is a set R of smooth functions $\rho : M \to [0, 1]$ such that for each $p \in M$ only a finite number of function values are non-zero and their sum equals 1.

Chapter 2

Fiber bundles and sections

2.1 Product manifolds and projections

Given manifolds M and N with atlases $\mathcal{A}_M$ and $\mathcal{A}_N$, one can easily construct another manifold as follows:

Definition 2.1.1 (Product manifold)**.** Let M and N be manifolds of dimensions m and n with atlases

$$\mathcal{A}_M = \{(U_i, \phi_i), i \in \mathcal{I}\}\,, \tag{2.1.1a}$$

$$\mathcal{A}_N = \{(V_j, \chi_j), j \in \mathcal{J}\}\,. \tag{2.1.1b}$$

On the Cartesian product

$$M \times N = \{(p, q) | p \in M, q \in N\} \tag{2.1.2}$$

define an atlas $\mathcal{A}_{M \times N}$ of dimension $m + n$ with charts (W_{ij}, ψ_{ij}) as follows:

- The sets W_{ij} are given by $W_{ij} = U_i \times V_j$.
- The functions $\psi_{ij} : W_{ij} \to \mathbb{R}^{m+n}$ are given by $\psi_{ij}(p, q) = (\phi_i(p), \chi_j(q))$.

The completion of this atlas to a maximal atlas then turns $M \times N$ into a manifold, called the *product manifold* (or *direct product*).

One can easily check that this is indeed an atlas. We also see immediately the dimension of the product manifold, which follows from the way the charts are constructed:

Theorem 2.1.1. *The dimension of a product manifold is given by* $\dim M \times N = \dim M + \dim N$.

The product manifold comes with a set of useful maps:

Definition 2.1.2 (Projection map)**.** Let M and N be manifolds and $M \times N$ their direct product. The maps $\mathrm{pr}_1 : M \times N \to M, (p, q) \mapsto p$ and $\mathrm{pr}_2 : M \times N \to N, (p, q) \mapsto q$ are called the *projections* onto the first and second factor, respectively.

Again, it is easy to check that the projections are indeed smooth maps.

We also take a brief look at the coordinates one can use on a product manifold. Given coordinates (x^i) on M and (y^μ) on N, corresponding to charts (U, ϕ) and (V, χ), the corresponding coordinates for the product chart (W, ψ) as constructed above are simply (x^i, y^μ). We illustrate this with a few examples.

Example 2.1.1. Let $M = \mathbb{R}$ the line and $N = S^1$ the circle. Their direct product is the cylinder $\mathbb{R} \times S^1$.

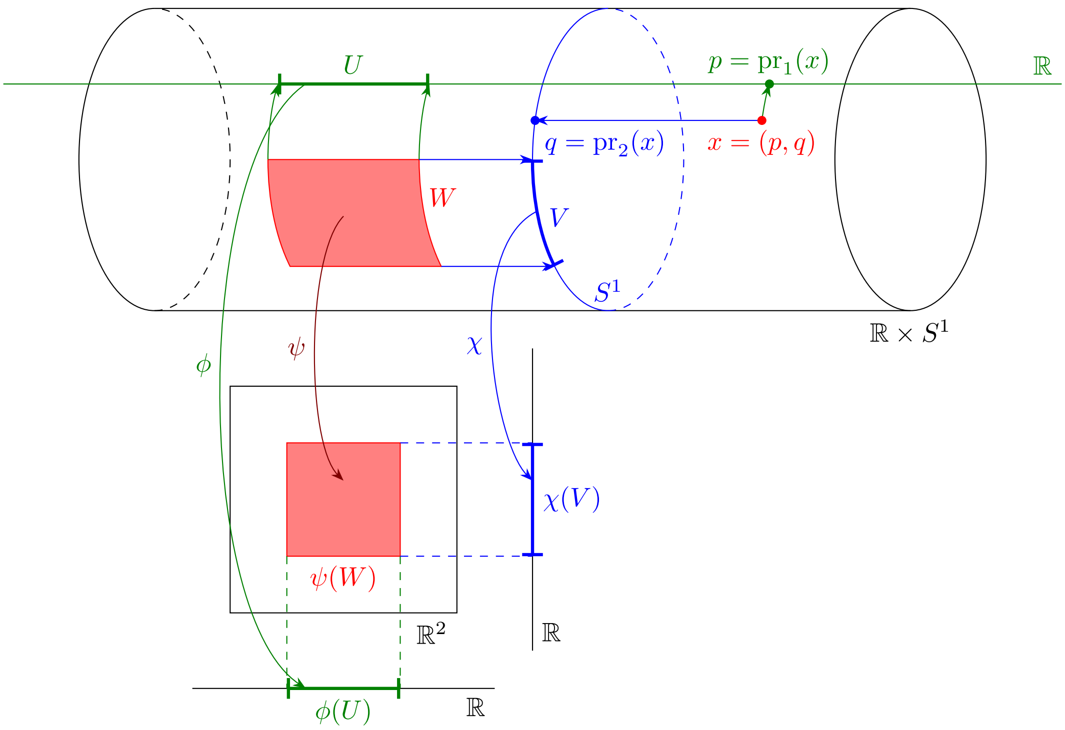


Figure 2.1: The (infinite) cylinder $\mathbb{R} \times S^1$ and its charts and projection maps.

Example 2.1.2. Let S^1 be the circle. The n-fold direct product $S^1 \times \ldots \times S^1$ is the n-dimensional torus T^n.

2.2 Fiber bundles

In the last section we have introduced the direct product of manifolds. We now discuss an important concept, called a *fiber bundle*, which can be viewed as a local version of a product manifold. Recall that in the case of the direct product $M \times N$ of two manifolds we have projections pr_1 and pr_2 onto each factor. One can show that the preimage of a point $p \in M$ under pr_1 is again a manifold which is diffeomorphic to N. We write: $\mathrm{pr}_1^{-1}(p) \cong N$. Of course the construction of the direct product is symmetric in M and N, so that also $\mathrm{pr}_2^{-1}(q) \cong M$ for $q \in N$.

For a fiber bundle, only one half of this is true. It consists of a manifold E called the *total space*, another manifold B called the *base space* and a map $\pi : E \to B$ called the *projection* or *bundle map*, such that for any $p \in B$ the preimage $\pi^{-1}(p)$ is diffeomorphic to a manifold F called the *fiber*. In addition, we need a condition which guarantees that the total space of the fiber bundle "locally looks like" a direct product. We define:

Definition 2.2.1 (Fiber bundle)**.** A *fiber bundle* (E, B, π, F) consists of manifolds E, B, F and a surjective map $\pi : E \to B$, such that for any $p \in B$ there exists an open set $U \subset B$ containing p and a diffeomorphism $\phi : \pi^{-1}(U) \to U \times F$ such that the diagram

$$\begin{array}{ccc} \pi^{-1}(U) & \xrightarrow{\phi} & U \times F \\ \downarrow \pi & \swarrow \mathrm{pr}_1 & \\ U & & \end{array} \tag{2.2.1}$$

commutes. The pair (U, ϕ) is called a *local trivialization*.

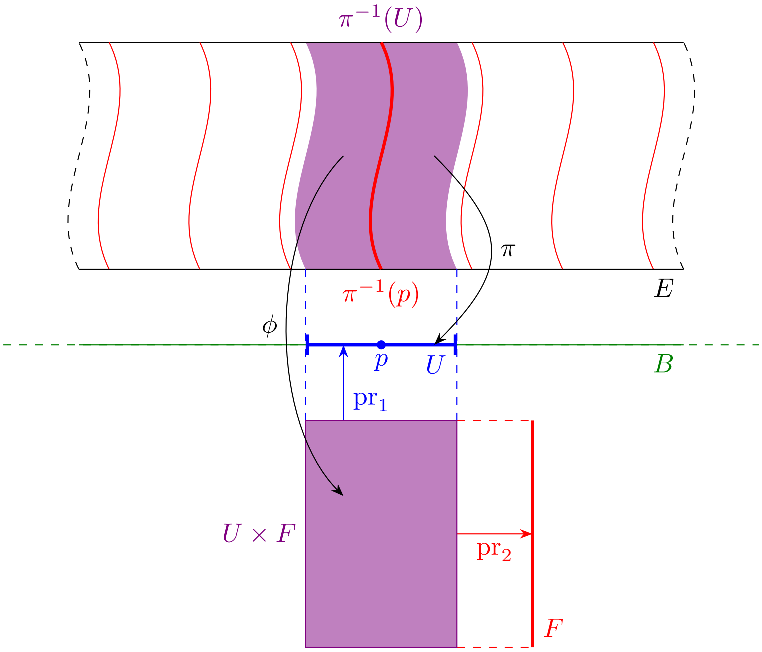


Figure 2.2: Illustration of a fiber bundle (E, B, π, F). For every $p \in B$ there exists $U \ni p$ such that $\pi^{-1}(U)$ is diffeomorphic to $U \times F$.

Another common notation for a fiber bundle is the "function notation" $\pi : E \to B$, when the fiber manifold F is clear from the context. We will also make use of this notation, as it turns out to be rather convenient. Sometimes, with a slight abuse of terminology, also the total space E is simply called a fiber bundle, when the base manifold and the projection are known. We will use this kind of terminology only for very special cases. Finally, in the latter case it is also common to denote the fiber $\pi^{-1}(p) \subset E$ over $p \in B$ by E_p.

We do not need to explicitly demand that $\pi^{-1}(p) \cong F$ for any $p \in B$, because this follows from the definition given above. To see this, note that $\mathrm{pr}_1^{-1}(p) \cong F$, as for any direct product. Since ϕ is a diffeomorphism, it follows that also $\pi^{-1}(p) = \phi^{-1}(\mathrm{pr}_1^{-1}(p)) \cong F$. Also the following relation for the dimension of a fiber bundle follows immediately from its definition:

Theorem 2.2.1. *For a fiber bundle* (E, B, π, F) *the dimensions of the manifolds satisfy* $\dim E = \dim B + \dim F$.

A particular class of fiber bundles, which we encounter frequently, is the following.

Definition 2.2.2 (Trivial fiber bundle)**.** The *trivial fiber bundle* with base manifold B and fiber F is the fiber bundle (E, B, π, F), where the total space is the product manifold $E = B \times F$ with projection $\pi = \mathrm{pr}_1$ onto the first factor.

Of course, not every fiber bundle is trivial. A well known example is the following.

Example 2.2.1 (Möbius strip). Let $\tilde{U}_1 = (0, 2\pi) \times (-1, 1)$, $\tilde{U}_2 = (-\pi, \pi) \times (-1, 1)$ and the functions

$$\begin{array}{rcccc} \tilde{\phi}_i & : & \tilde{U}_i & \to & \mathbb{R}^3 \\ & & (t, s) & \mapsto & \left(\left(R + Ws\cos\frac{t}{2}\right)\cos t, \left(R + Ws\cos\frac{t}{2}\right)\sin t, Ws\sin\frac{t}{2}\right) \end{array}, \tag{2.2.2}$$

$i = 1, 2$, with constants $0 < W < R$. Let further $E = \tilde{\phi}_1(\tilde{U}_1) \cup \tilde{\phi}_2(\tilde{U}_2)$. It is easy to show that E carries the structure of a two-dimensional manifold, and that an atlas is given by the charts $(U_i = \tilde{\phi}_i(\tilde{U}_i), \phi_i = \tilde{\phi}_i^{-1})$. This manifold is called the *Möbius strip*.

Now consider the function

$$\begin{array}{rcccc} \tilde{\pi} & : & \{x \in \mathbb{R}^3 | x_1^2 + x_2^2 > 0\} & \to & \mathbb{R}^2 \\ & & (x_1, x_2, x_3) & \mapsto & \left(\frac{x_1}{\sqrt{x_1^2+x_2^2}}, \frac{x_2}{\sqrt{x_1^2+x_2^2}}\right) \end{array} \tag{2.2.3}$$

Looking at the compositions

$$\begin{array}{rcccc} \tilde{\pi} \circ \tilde{\phi}_i & : & \tilde{U}_i & \to & \mathbb{R}^2 \\ & & (t, s) & \mapsto & (\cos t, \sin t) \end{array} \tag{2.2.4}$$

one can see that the restriction of $\tilde{\pi}$ to E defines a smooth map $\pi : E \to S^1 = \{x \in \mathbb{R}^2 | x_1^2 + x_2^2 = 1\}$ and that the preimages $\pi^{-1}(p)$ are diffeomorphic to $(-1, 1)$. One can show that $(E, S^1, \pi, (-1, 1))$ is a non-trivial fiber bundle.

2.3 Sections

When dealing with fiber bundles we often work with maps $\sigma : B \to E$ which assign to each point $p \in B$ on the base manifold a point $\sigma(p) \in \pi^{-1}(p) \subset E$ on the fiber over p. These maps are called *sections*, or *cross sections*, and are defined as follows.

Definition 2.3.1 (Section)**.** A *(global) section* of a fiber bundle (E, B, π, F) is a map $\sigma : B \to E$ such that $\pi \circ \sigma = \mathrm{id}_B$.

We usually omit the word global, and simply speak of sections. It should be noted that not every bundle admits global sections - there are bundles for which no global sections exist. This is often the case, for example, for frame bundles, which we will cover later. However, every bundle admits local sections, which we define as follows.

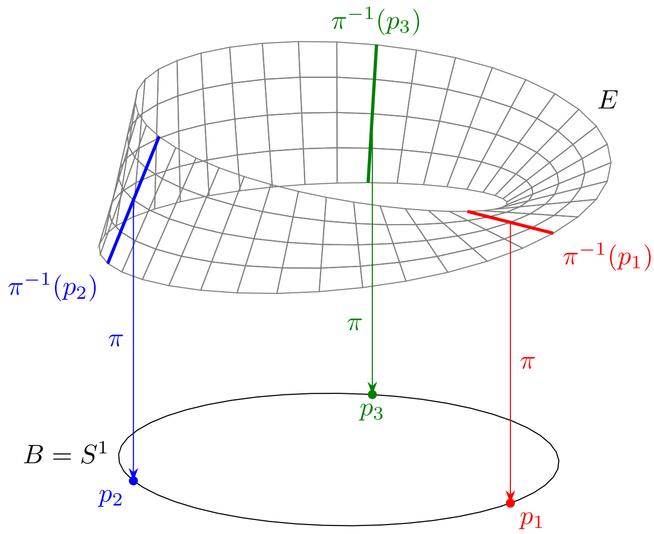


Figure 2.3: Möbius strip E as a fiber bundle with fiber $F = (-1, 1)$ over the base manifold $B = S^1$.

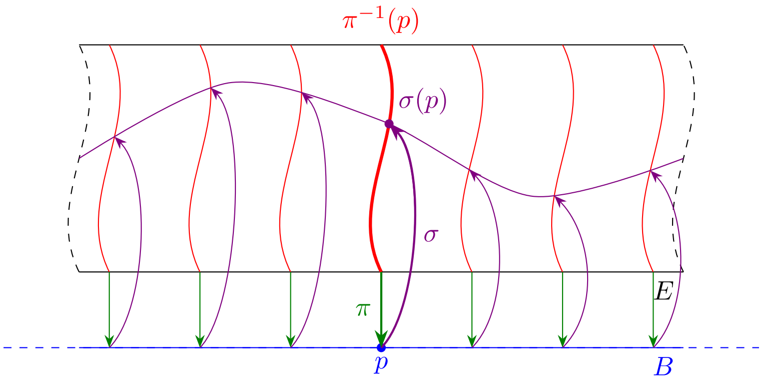


Figure 2.4: Illustration of a section σ, which assigns to each point $p \in B$ a point $\sigma(p) \in \pi^{-1}(p)$ of the fiber $\pi^{-1}(p)$ over p.

Definition 2.3.2 (Local section)**.** A *local section* of a fiber bundle (E, B, π, F) is pair (U, σ), composed of an open subset $U \subset B$ and a map $\sigma : U \to E$ defined on U such that $\pi \circ \sigma = \mathrm{id}_U$.

The set of all sections of a bundle is often denoted with the letter Γ, and we introduce the following notation, which will turn out to be useful later:

- The set of all global sections of a fiber bundle (E, B, π, F) we denote $\Gamma(E, B, \pi, F)$, or simply $\Gamma(E)$ if it is clear which are the other ingredients of the fiber bundle. Another convenient notation we will use is $\Gamma(\pi)$, if E and B are the known domain and codomain of π.
- The set of all local sections of (E, B, π, F) will be denoted $\Gamma_|(E, B, \pi, F)$ or simply $\Gamma_|(E)$ or $\Gamma_|(\pi)$.

- The set of all local sections of (E, B, π, F) with fixed domain $U \subset B$ will be denoted $\Gamma_{|U}(E, B, \pi, F)$ or simply $\Gamma_{|U}(E)$ or $\Gamma_{|U}(\pi)$.
- The set of all local sections of (E, B, π, F) whose domain contains $p \in B$ will be denoted $\Gamma_p(E, B, \pi, F)$ or simply $\Gamma_p(E)$ or $\Gamma_p(\pi)$.

These different sets are obviously related by

$$\Gamma(E) = \Gamma_{|B}(E)\,, \quad \Gamma_p(E) = \bigcup_{U \subset B, p \in U} \Gamma_{|U}(E)\,, \quad \Gamma_{|}(E) = \bigcup_{U \subset B} \Gamma_{|U}(E) = \bigcup_{p \in B} \Gamma_p\,. \tag{2.3.1}$$

As we mentioned before, there are bundles which do not admit any global sections, and so have $\Gamma(E) = \emptyset$. However, the situation is different for local sections, as we will see now.

Theorem 2.3.1. *For every fiber bundle (E, B, π, F) there exist local sections around any point $p \in B$, so that $\Gamma_p(E) \neq \emptyset$.*

Proof. By definition, a fiber bundle admits a local trivialization (U, ϕ) around any point $p \in B$, where $U \subset B$ contains p and $\phi : \pi^{-1}(U) \to U \times F$ is a diffeomorphism satisfying $\mathrm{pr}_1 \circ \phi = \pi$. Choose $f \in F$, and define

$$\begin{array}{rcccc} \sigma & : & U & \to & E \\ & & x & \mapsto & \phi^{-1}(x, f) \end{array}. \tag{2.3.2}$$

This is obviously smooth, since ϕ^{-1} and the inclusion $x \mapsto (x, f)$ of U in $U \times F$ are smooth maps. Further, it satisfies

$$(\pi \circ \sigma)(x) = \pi(\phi^{-1}(x, f)) = \mathrm{pr}_1(x, f) = x\,, \tag{2.3.3}$$

and so $\pi \circ \sigma = \mathrm{id}_U$. Hence, it is a local section whose domain contains p. ■

We remark that in the previous proof, instead of choosing a single point $f \in F$ in the fiber, one could also have chosen an arbitrary function $F : U \to F$, and defined $\sigma(x) = \phi^{-1}(x, F(x))$. In fact, this is just how any section can be expressed within a local trivialization. Of course, a fiber bundle is in general only *locally* trivial, and so only the existence of local sections is guaranteed by this statement, but not that of global sections. For trivial fiber bundles, however, global sections always exist, and we can identify them as follows.

Theorem 2.3.2. *Let M, N be manifolds and $M \times N$ their product, i.e., the trivial fiber bundle $(M \times N, M, \mathrm{pr}_1, N)$. Then there exists a one-to-one correspondence between maps $\psi : M \to N$ and global sections $\sigma : M \to M \times N$.*

Proof. Given a map $\psi : M \to N$, one can construct a section

$$\begin{array}{rcccc} \sigma & : & M & \to & M \times N \\ & & p & \mapsto & (p, \psi(p)) \end{array}, \tag{2.3.4}$$

while for a section $\sigma : M \to M \times N$, one can construct a map $\psi = \mathrm{pr}_2 \circ \sigma$. One easily checks that this establishes the desired one-to-one correspondence. ■

A particularly simple case is the following.

Theorem 2.3.3. *Let M, N be manifolds and $M \times N$ their product, i.e., the trivial fiber bundle $(M \times N, M, \mathrm{pr}_1, N)$. For every $q \in N$, the constant section $M \to M \times N, p \mapsto (p, q)$ is a smooth section.*

Proof. Recall from theorem 1.3.1 that the constant map $f_q : M \to N, p \mapsto q$ is a smooth map. From the preceding theorem 2.3.2 then follows that the constant section $p \mapsto (p, f_q(p)) = (p, q)$ is a smooth section. ■

2.4 Induced charts and coordinates

Since E, B and F are manifolds, we could in principle use any kind of chart to describe them. However, it is often more convenient to use special charts on E, which we call *adapted* or *induced*. These can be constructed as follows.

Theorem 2.4.1. *Let (E, B, π, F) be a fiber bundle. Let (V, χ) be a chart on F and (U, ψ) a chart on B, such that there exists a local trivialization (U, ϕ). Then the pair (W, ω), where $W = \phi^{-1}(U \times V)$ and*

$$\omega(e) = (\psi(\mathrm{pr}_1(\phi(e))), \chi(\mathrm{pr}_2(\phi(e)))) \tag{2.4.1}$$

is a chart of dimension $\dim B + \dim F$ *on* E.

Proof. Let $m = \dim B$ and $n = \dim F$. Then $\psi(\mathrm{pr}_1(\phi(e))) \in \mathbb{R}^m$ and $\chi(\mathrm{pr}_2(\phi(e))) \in \mathbb{R}^n$, and so $\omega(e) \in \mathbb{R}^{m+n}$. Further, let $e' \in W$ with $e \neq e'$. Then also $\phi(e) \neq \phi(e')$, since ϕ is bijective, and hence

$$(\mathrm{pr}_1(\phi(e)), \mathrm{pr}_2(\phi(e))) = \phi(e) \neq \phi(e') = (\mathrm{pr}_1(\phi(e')), \mathrm{pr}_2(\phi(e'))) \,. \tag{2.4.2}$$

Since ψ and χ are charts, and therefore injective, it thus follows that

$$\omega(e) = (\psi(\mathrm{pr}_1(\phi(e))), \chi(\mathrm{pr}_2(\phi(e)))) \neq (\psi(\mathrm{pr}_1(\phi(e'))), \chi(\mathrm{pr}_2(\phi(e')))) = \omega(e') \,, \tag{2.4.3}$$

and so also ω is injective. Finally,

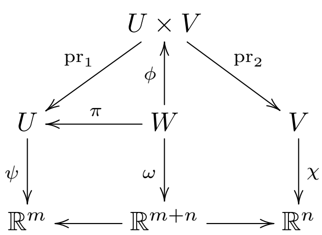
$$\omega(W) = \psi(U) \times \chi(V) \subset \mathbb{R}^{m+n} \tag{2.4.4}$$

is open in the product topology, since $\psi(U)$ and $\chi(V)$ are open. Hence, (W, ω) is a chart of dimension $m + n$. ■

The type of chart constructed above also carries its own name:

Definition 2.4.1 (Induced chart). The chart (W, ω) constructed in theorem 2.4.1 is called an *induced chart*.

To illustrate this definition, let $\dim B = m$, $\dim F = n$ and consider the following diagram:



(2.4.5)

The lower half of this diagram shows the charts $\psi : U \to \mathbb{R}^m$ and $\chi : V \to \mathbb{R}^n$. To understand its upper half, recall from definition 2.2.1 that a local trivialization (U, ϕ) consists of an open set $U \subset B$ and a diffeomorphism $\phi : \pi^{-1}(U) \to U \times F$. Since $V \subset F$, we have $U \times V \subset U \times F$, and so its preimage under ϕ satisfies

$$W = \phi^{-1}(U \times V) \subset \phi^{-1}(U \times F) = \pi^{-1}(U) \subset E \,. \tag{2.4.6}$$

The map ϕ now assigns to each $e \in W$ a pair (u, v) with $u = \pi(e) \in U$ and $v \in V$. Via the charts on B and F, we then obtain a pair $(\psi(u), \chi(v)) \in \mathbb{R}^m \times \mathbb{R}^n = \mathbb{R}^{m+n}$. This is the function ω introduced in definition 2.4.1.

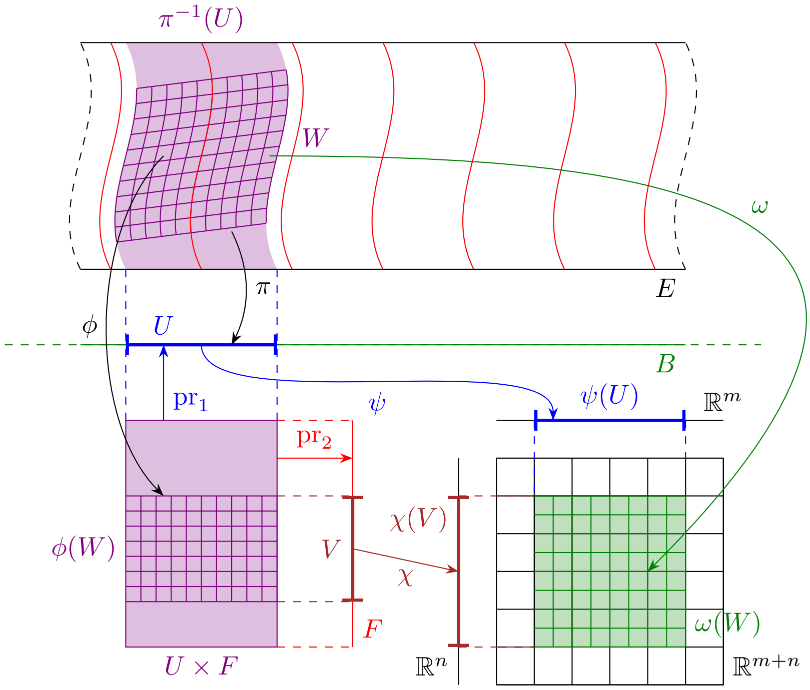


Figure 2.5: Induced coordinates on a fiber bundle (E, B, π, F). Note that the "vertical" coordinate lines follow the fibers in E, but there is no restriction on the location of the "horizontal" coordinate lines; the latter depends on the choice of the local trivialization (U, ϕ).

We also remark that in the definition we have chosen the open set $U \subset B$ used in the chart (U, ψ) and the local trivialization (U, ϕ) to be the same. Note that it is always possible to find such a chart around an arbitrary point $p \in B$. By definition of a manifold, around each point p there exists some chart $(\tilde{U}, \tilde{\psi})$ with $p \in \tilde{U}$. Further, by the definition of a fiber bundle there also exists a local trivialization $(\bar{U}, \bar{\phi})$ such that $p \in \bar{U}$. Now we simply define $U = \tilde{U} \cap \bar{U}$, as well as $\psi = \tilde{\psi}|_U$ and $\phi = \bar{\phi}|_U$. One easily checks that (U, ψ) is a chart and (U, ϕ) is a local trivialization. Of course, $p \in U$.

The main advantage of using induced charts and the corresponding coordinates on the total space E becomes apparent if we consider sections of a fiber bundle. Recall from definition 2.3.1 that a section is a map $f : B \to E$ such that $\pi \circ f = \mathrm{id}_B$. Using charts (U, ψ) on B and a corresponding induced chart (W, ω) on E, which comes from a suitable chart (V, χ) on F, we can express f in coordinates by looking at

$$\omega \circ f|_U \circ \psi^{-1} : \psi(U) \to \mathbb{R}^{m+n} \,. \tag{2.4.7}$$

Now by the definition of ω we have for $x \in \psi(U)$:

$$\begin{aligned}
(\omega \circ f \circ \psi^{-1})(x) &= ((\psi \circ \underbrace{\mathrm{pr}_1 \circ \phi}_{=\pi} \circ f \circ \psi^{-1})(x), (\chi \circ \mathrm{pr}_2 \circ \phi \circ f \circ \psi^{-1})(x)) \\
&= ((\psi \circ \underbrace{\pi \circ f}_{=\mathrm{id}_B} \circ \psi^{-1})(x), (\chi \circ \mathrm{pr}_2 \circ \phi \circ f \circ \psi^{-1})(x)) \\
&= ((\underbrace{\psi \circ \mathrm{id}_B \circ \psi^{-1}}_{=\mathrm{id}_{\psi(U)}})(x), (\chi \circ \mathrm{pr}_2 \circ \phi \circ f \circ \psi^{-1})(x)) \\
&= (x, (\chi \circ \mathrm{pr}_2 \circ \phi \circ f \circ \psi^{-1})(x)) \,.
\end{aligned} \tag{2.4.8}$$

We see that the first m components are already fully determined by the fact that f is a section, and we only need to specify the remaining n components. This becomes even more clear if we express the section explicitly in coordinates. Denoting the components of $\mathbb{R}^m$ by $(x^i, i = 1,\ldots,m)$, the components of $\mathbb{R}^n$ by $(y^\mu, \mu = 1,\ldots,n)$ and the components of $\mathbb{R}^{m+n}$ by $(z^A, A = 1,\ldots,m+n)$, instead of writing the section in generic coordinates as

$$\omega \circ f \circ \psi^{-1} : (x^i) \mapsto (z^A)\,, \tag{2.4.9}$$

which involves $m + n$ functions for all z^A, the choice of adapted coordinates $(z^A) = (x^i, y^\mu)$ allows us to write

$$\omega \circ f \circ \psi^{-1} : (x^i) \mapsto (x^i, y^\mu)\,, \tag{2.4.10}$$

which requires only n functions for y^μ to be given, since the functions for x^i are identities.

Of course, a single chart is in general not sufficient to cover the whole manifold E, and so we will need a whole atlas instead, which is comprised of mutually compatible charts which cover E. To achieve this with the induced chart construction, we first show compatibility as follows.

Theorem 2.4.2. *Let (E, B, π, F) be a fiber bundle. Let $(V_1, \chi_1), (V_2, \chi_2)$ be charts on F and $(U_1, \psi_1), (U_2, \psi_2)$ charts on B, such that there exist local trivializations $(U_1, \phi_1), (U_2, \phi_2)$. Then the corresponding induced charts $(W_1, \omega_1), (W_2, \omega_2)$ are compatible.*

Proof. We will assume $W_1 \cap W_2 \neq \emptyset$, since otherwise they are trivially compatible. For $e \in W_1 \cap W_2$ we have

$$\omega_1(e) = (\psi_1(\mathrm{pr}_1(\phi_1(e))), \chi_1(\mathrm{pr}_2(\phi_1(e))))\,, \tag{2.4.11a}$$
$$\omega_2(e) = (\psi_2(\mathrm{pr}_1(\phi_2(e))), \chi_2(\mathrm{pr}_2(\phi_2(e))))\,. \tag{2.4.11b}$$

►...◄ ■

Given atlases $\mathcal{A}_B$ on B and $\mathcal{A}_F$ on F, it is now straightforward to show that one can construct an atlas on E from the induced charts as follows.

Theorem 2.4.3. *Let (E, B, π, F) be a fiber bundle. The induced charts defined from atlases $\mathcal{A}_B$ and $\mathcal{A}_F$ together with the local trivializations form an atlas of E.*

Proof. ►...◄ ■

Of course, since E is a manifold, it is already equipped with an atlas by definition 1.1.6, and so we finally need to show that the original atlas on E is compatible with this induced atlas, so that both describe the same manifold.

Theorem 2.4.4. *Let (E, B, π, F) be a fiber bundle. The induced atlas defined from atlases $\mathcal{A}_B$ and $\mathcal{A}_F$ together with the local trivializations is compatible with $\mathcal{A}_E$.*

Proof. ►...◄ ■

2.5 Construction from trivializations

Theorem 2.4.4 has an interesting consequence: it shows that we can fully reconstruct the manifold structure on E, defined by its unique maximal atlas, from any induced atlas, and hence from the atlases on B and F and the local trivializations alone, even if we are not given an atlas on E beforehand. Of course, this constructing requires some additional input, since the definition of a local trivialization (U, ϕ) entails the condition that ϕ is a diffeomorphism, which requires that E is a manifold, and also the definition of the fiber bundle itself requires that the total space E is a manifold and π is a map. If we do not assume a manifold structure on the set E to be given a priori, we can only speak of bijective or surjective functions, and we need to replace this assumption by imposing certain conditions on them to guarantee that they will again yield an atlas on E. This leads to the following construction.

Theorem 2.5.1. *Let B, F be manifolds, E a set and $\pi : E \to B$ a surjective function. Further, let $(U_i, i \in \mathcal{I})$ be an open cover, such that for each $i \in \mathcal{I}$ there exists a bijective function $\phi_i : \pi^{-1}(U_i) \to U_i \times F$ with $\mathrm{pr}_1 \circ \phi_i = \pi$, such that for any $i, j \in \mathcal{I}$ with $U_{ij} = U_i \cap U_j \neq \emptyset$ the function $\phi_{ij} = \phi_i \circ \phi_j^{-1} : U_{ij} \times F \to U_{ij} \times F$ is a diffeomorphism. Then E is equipped with a unique manifold structure such that (E, B, π, F) is a fiber bundle.*

Proof. ►. . . ◄ ■

2.6 Construction from transition functions

In the definition 2.2.1 of a fiber bundle (E, B, π, F) we started with a given total space E, and demanded the existence of local trivializations as a condition. In practice, however, it is not always easy to prove the existence of such maps. Also one often encounters the situation that one does not a priori know the total space manifold E, but only some "local information" which specifies how fibers must be "glued" together to a total space, and then needs to construct E along with the fiber bundle structure from this local information. In the following, we explain this construction.

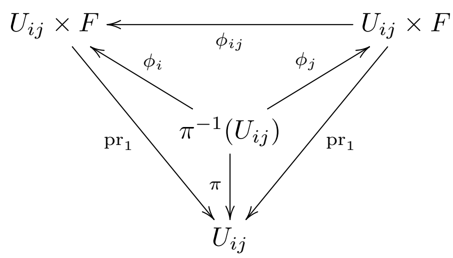
Recall that given a fiber bundle (E, B, π, F), for every $p \in B$, there exists a local trivialization (U, ϕ), with $p \in U \subset B$ and $\phi : \pi^{-1}(U) \to U \times F$ a diffeomorphism satisfying $\mathrm{pr}_1 \circ \phi = \pi$. From the fact that every point $p \in B$ must be contained in some local trivialization follows that the open sets $U \subset B$ on which these trivializations are defined cover B. This is similar to the definition 1.1.4 of an atlas, where one also covers a manifold, but with charts. Another similarity is the fact that in general one cannot use the same set $U = B$ for all points p, unless the fiber bundle is trivial; in general it will be necessary to use several, distinct local trivializations, and so it makes sense to label them as (U_i, ϕ_i) with an index $i \in \mathcal{I}$ in some index set $\mathcal{I}$, like we did with charts. The definition of a fiber bundle states that one can find local trivializations to cover B, hence

$$\bigcup_{i \in \mathcal{I}} U_i = B\,. \tag{2.6.1}$$

Also, again similarly to the case of charts, these local trivializations satisfy a certain notion of compatibility by construction. To see this, consider two local trivializations (U_i, ϕ_i) and (U_j, ϕ_j) with $U_{ij} = U_i \cap U_j \neq \emptyset$. Like with charts, one can then consider the maps

$$\begin{array}{cccccc} \phi_{ij} & : & U_{ij} \times F & \to & U_{ij} \times F \\ & & (p, f) & \mapsto & \phi_i(\phi_j^{-1}(p, f)) \end{array} . \tag{2.6.2}$$

Now recalling the property of local trivializations to preserve the base point p, we can construct the following diagram:



(2.6.3)

To clarify this diagram, recall that a local trivialization ϕ_i satisfies $\mathrm{pr}_1 \circ \phi_i = \pi$, and so $\pi(\phi_i^{-1}(p, f)) = p$. Consequently, for $p \in U_{ij}$, we have $\phi_i^{-1}(p, f) \in \pi^{-1}(U_{ij})$. Conversely, choosing $e \in \pi^{-1}(U_{ij})$, and hence $\pi(e) \in U_{ij}$, we have $\mathrm{pr}_1(\phi_i(e)) = \pi(e) \in U_{ij}$, and thus $\phi_i(e) \in U_{ij} \times F$. Taking these two statements together, we see that ϕ_i bijectively maps $\pi^{-1}(U_{ij})$ to $U_{ij} \times F$. Smoothness follows from the condition that a local trivialization is smooth. The same holds for ϕ_j, and so the upper part of the diagram, including the map ϕ_{ij}, is well-defined.

As discussed above, the left and right triangles in this diagram commute by the properties of local trivialization, $\mathrm{pr}_1 \circ \phi_i = \pi$ and analogously for ϕ_j. The upper triangle commutes by the definition of ϕ_{ij}. Hence, also the outer triangle commutes, i.e., we have $\mathrm{pr}_1 \circ \phi_{ij} = \mathrm{pr}_1$, and so ϕ_{ij} preserves the base point p. From the construction of ϕ_{ij} further follows

$$(\phi_{i_1 i_2} \circ \cdots \phi_{i_{n-1} i_n} \circ \phi_{i_n i_1})(p, f) = (\phi_{i_1} \circ \phi_{i_2}^{-1} \circ \phi_{i_2} \circ \cdots \circ \phi_{i_n}^{-1} \circ \phi_{i_n} \circ \phi_{i_1}^{-1})(p, f) = (p, f) \tag{2.6.4}$$

for all $(p, f) \in U_{i_1 \cdots i_n} \times F$, where we used the notation

$$U_{i_1 \cdots i_n} = \bigcap_{k=1}^{n} U_{i_k} \,. \tag{2.6.5}$$

It is instructive to consider this relation for the lowest values of n. For $n = 1$ it states that

$$\phi_{ii}(p, f) = (p, f) \quad \Leftrightarrow \quad \phi_{ii} = \mathrm{id}_{U_i \times F} \tag{2.6.6}$$

for all $i \in \mathcal{I}$. Setting $n = 2$, we have

$$(\phi_{ij} \circ \phi_{ji})(p, f) = (p, f) \quad \Leftrightarrow \quad \phi_{ij} = \phi_{ji}^{-1} \tag{2.6.7}$$

for all $i, j \in \mathcal{I}$. Finally, for $n = 3$ we find

$$(\phi_{ij} \circ \phi_{jk} \circ \phi_{ki})(p, f) = (p, f) \quad \Leftrightarrow \quad (\phi_{ij} \circ \phi_{jk})|_{U_{ijk} \times F} = \phi_{ki}^{-1}|_{U_{ijk} \times F} = \phi_{ik}|_{U_{ijk} \times F} \tag{2.6.8}$$

for all $i, j, k \in \mathcal{I}$. We do not need to consider higher values of n, since the last relation allows us to conclude from any n to $n + 1$. Hence, if we aim to construct local trivializations from the transition functions alone, we will need these three conditions. This brings us to the following construction.

Theorem 2.6.1. *Let B and F be manifolds, $(U_i, i \in \mathcal{I})$ an open cover of B and $\phi_{ij} : U_{ij} \times F \to U_{ij} \times F$ diffeomorphisms such that*

$$\mathrm{pr}_1 \circ \phi_{ij} = \mathrm{pr}_1 \,, \tag{2.6.9a}$$

$$\phi_{ii} = \mathrm{id}_{U_i \times F} \,, \tag{2.6.9b}$$

$$\phi_{ij} = \phi_{ji}^{-1} \,, \tag{2.6.9c}$$

$$(\phi_{ij} \circ \phi_{jk})|_{U_{ijk} \times F} = \phi_{ik}|_{U_{ijk} \times F} \tag{2.6.9d}$$

for all $i, j, k \in \mathcal{I}$. Then

$$E = \biguplus_{i \in \mathcal{I}} U_i \times F \Big/ \sim \,, \tag{2.6.10}$$

where

$$(i, p, f) \sim (i', p', f') \quad \Leftrightarrow \quad (p', f') = \phi_{i'i}(p, f) \,, \tag{2.6.11}$$

is equipped with a unique manifold structure such that (E, B, π, F), with

$$\begin{array}{cccc} \pi & : & E & \to & B \\ & & [i, p, f] & \mapsto & p \end{array} \,, \tag{2.6.12}$$

is a fiber bundle.

Proof. We first need to show that the relation $\sim$ defined above is indeed an equivalence relation, so that we can define E as the quotient set, constituted by equivalence classes. Hence, we check:

1. Reflexivity: $\phi_{ii}(p, f) = (p, f)$ and thus $(i, p, f) \sim (i, p, f)$.
2. Symmetry: If $(i, p, f) \sim (i', p', f')$, then $(p', f') = \phi_{i'i}(p, f)$. Hence,

   $$(p, f) = \phi_{i'i}^{-1}(p', f') = \phi_{ii'}(p', f') \,, \tag{2.6.13}$$

   and thus $(i', p', f') \sim (i, p, f)$.

3. Transitivity: If $(i,p,f) \sim (i',p',f')$ and $(i',p',f') \sim (i'',p'',f'')$, then

$$(p'',f'') = \phi_{i''i'}(p',f') = (\phi_{i''i'} \circ \phi_{i'i})(p,f) = \phi_{i''i}(p,f)\,, \tag{2.6.14}$$

and thus $(i,p,f) \sim (i'',p'',f'')$.

Hence, E is well-defined. Further, one has

$$p' = \mathrm{pr}_1(p',f') = \mathrm{pr}_1(\phi_{i'i}(p,f)) = \mathrm{pr}_1(p,f) = p \tag{2.6.15}$$

whenever $(i,p,f) \sim (i',p',f')$, and thus also π is well-defined, since it does not depend on the representative of the equivalence class $[i,p,f]$. Moreover, for each $p \in B$ there exists $i \in \mathcal{I}$ such that $p \in U_i$, since the U_i cover B. For each $i \in \mathcal{I}$, we can then define

$$\begin{array}{cccc} \tilde{\phi}_i & : & U_i \times F & \to & \pi^{-1}(U_i) \\ & & (p,f) & \mapsto & [i,p,f] \end{array}\,. \tag{2.6.16}$$

Also this is well-defined, since

$$\pi(\tilde{\phi}_i(p,f)) = \pi([i,p,f]) = p = \mathrm{pr}_1(p,f) \in U_i \tag{2.6.17}$$

by construction. It is surjective, since

$$\pi^{-1}(U_i) = \{[i,p,f], p \in U_i, f \in F\}\,, \tag{2.6.18}$$

and injective, since $[i,p,f] = [i,p',f']$ if and only if $(p,f) = (p',f')$, which follows from the fact that $\phi_{ii} = \mathrm{id}_{U_i \times F}$. Hence, we can take its inverse $\phi_i = \tilde{\phi}_i^{-1}$, and find that it satisfies

$$(\mathrm{pr}_1 \circ \phi_i)([i,p,f]) = \mathrm{pr}_1(p,f) = p = \pi([i,p,f])\,, \tag{2.6.19}$$

and so $\mathrm{pr}_1 \circ \phi_i = \pi$. Finally, we have that

$$(\phi_i \circ \phi_j^{-1})(p,f) = \phi_i([j,p,f]) = \phi_i([i,p,\mathrm{pr}_2(\phi_{ij}(p,f))]) = \phi_{ij}(p,f)\,, \tag{2.6.20}$$

which follows from

$$(j,p,f) \sim (i,p,\mathrm{pr}_2(\phi_{ij}(p,f)))\,, \tag{2.6.21}$$

and so $\phi_i \circ \phi_j^{-1} = \phi_{ij}$, which we assumed to be a diffeomorphism. Hence, the functions ϕ_i satisfy the conditions of theorem 2.4.4, and so the statement of the theorem follows. ▶...◀ ■

2.7 Bundle morphisms

We can say that a fiber bundle (E,B,π,F) equips the total space E with some additional structure, by dividing it into fibers, such that it locally looks like a product manifold. When we consider maps between the total spaces of two fiber bundles, we are mostly interested in maps which preserve this structure, i.e., we want points that reside on the same fiber to be mapped again to points residing on the same fiber. A map that satisfies this property is called a *bundle morphism* and defined as follows:

Definition 2.7.1 (Bundle morphism)**.** Let (E_1,B_1,π_1,F_1) and (E_2,B_2,π_2,F_2) be fiber bundles. A *bundle morphism* (or *bundle map*) is a map $\theta : E_1 \to E_2$ such that there exists a map $\vartheta : B_1 \to B_2$ for which the diagram

$$\begin{array}{ccc} E_1 & \xrightarrow{\theta} & E_2 \\ {\scriptstyle \pi_1}\downarrow & & \downarrow{\scriptstyle \pi_2} \\ B_1 & \xrightarrow[\vartheta]{} & B_2 \end{array} \tag{2.7.1}$$

commutes. The bundle morphism θ is then said to *cover* ϑ.

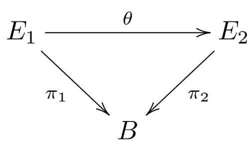
We see that a bundle morphism actually comprises two maps, since a second map ϑ relating the base spaces also appeared in this definition. However, we do not need to specify ϑ explicitly, since it is already uniquely determined by θ. To see this, recall from definition 2.2.1 that the projection of a fiber bundle, and hence also π_1, is surjective. This means that for any $p \in B_1$ we can find at least one $e \in E_1$ such that $\pi_1(e) = p$. The commutativity of the diagram (2.7.1) then implies that

$$\vartheta(p) = (\vartheta \circ \pi_1)(e) = (\pi_2 \circ \theta)(e) \tag{2.7.2}$$

is uniquely determined. Note that the definition also implies that θ must preserve the fibers, i.e., points $e, e' \in E_1$ with $\pi_1(e) = p = \pi_1(e')$ are mapped to points $\theta(e), \theta(e') \in E_2$ with $\pi_2(\theta(e)) = \vartheta(p) = \pi_2(\theta(e'))$.

Sometimes we need maps which do not only preserve the fibers as discussed above, but maps between bundles over the same base space that also preserve the base point. This condition then leads to the following definition.

Definition 2.7.2 (Bundle morphism covering the identity)**.** Let (E_1, B, π_1, F_1) and (E_2, B, π_2, F_2) be fiber bundles over a common base manifold B. A *bundle morphism covering the identity* is a bundle morphism $\theta : E_1 \to E_2$ covering the identity map $\vartheta = \mathrm{id}_B$ on B, i.e., a map θ such that the diagram



(2.7.3)

commutes.

Note that in general neither of the maps θ, ϑ is required to be bijective. Also there is no particular requirement on the fibers or the base manifolds of the two bundles. However, if one poses an additional condition on the bundle morphism, such that it possesses an inverse, this also implies further compatibility conditions for the two bundles. Let us therefore define:

Definition 2.7.3 (Bundle isomorphism)**.** A *bundle isomorphism* is a bijective bundle morphism whose inverse is also a bundle morphism. If a bundle isomorphism between two bundles exists, these bundles are called *isomorphic*.

It is now easy to check the following statement:

Theorem 2.7.1. *Let θ be a bundle isomorphism covering a map ϑ. Then both θ and ϑ are diffeomorphisms.*

Proof. This immediately follows from the fact that θ and ϑ are maps by definition of a bundle morphism, which in this case must be bijective and whose inverses are also maps, so that they are diffeomorphisms. ■

Another helpful fact is the following.

Theorem 2.7.2. *Let (E_i, B_i, π_i, F_i) with $i = 1, \ldots, 3$ be fiber bundles and $\theta : E_1 \to E_2$ and $\theta' : E_2 \to E_3$ be bundle morphisms (isomorphisms) covering $\vartheta : B_1 \to B_2$ and $\vartheta' : B_2 \to B_3$. Then $\theta' \circ \theta : E_1 \to E_3$ is a bundle morphism (isomorphism) covering $\vartheta' \circ \vartheta : B_1 \to B_3$.*

Proof. ▶ . . . ◀ ■

Finally, we take a look at the relation between bundle morphisms and sections. In particular, we have the following statement.

Theorem 2.7.3. *Let (E_1, B, π_1, F_1) and (E_2, B, π_2, F_2) be fiber bundles over a common base manifold B, $\theta : E_1 \to E_2$ a bundle morphism covering the identity and $\sigma : B \to E_1$ a section. Then also $\theta \circ \sigma : B \to E_2$ is a section.*

Proof. First note that if σ and θ are smooth maps, then also $\theta \circ \sigma$ is a smooth maps. Further, from the fact that θ is a bundle morphism covering the identity follows

$$\pi \circ \theta \circ \sigma = \pi \circ \sigma = \mathrm{id}_B \,, \tag{2.7.4}$$

and so $\theta \circ \sigma$ is a section. ■

2.8 Fibered product

There are different possibilities to construct new fiber bundles from given ones, some of which we will study in the following sections. One of the most elementary is the following. Given two fiber bundles over a common base manifold, one may construct another fiber bundle as follows:

Definition 2.8.1 (Fibered product)**.** Let (E_1, B, π_1, F_1) and (E_2, B, π_2, F_2) be two fiber bundles over a common base manifold B. Their *fibered product* is the fiber bundle (E, B, π, F) over B, where:

- The fiber is given by the product manifold $F = F_1 \times F_2$.
- The total space is the set

$$E = E_1 \times_B E_2 = \{(e_1, e_2) \in E_1 \times E_2 \,|\, \pi_1(e_1) = \pi_2(e_2)\} \subset E_1 \times E_2 \,. \tag{2.8.1}$$

- The atlas of the total space is constructed from the induced charts from definition 2.4.1, where the charts (V, χ) are the charts of the product manifold $F = F_1 \times F_2$ as given in definition 2.1.1.
- The projection is given by the map

$$\begin{array}{rccc} \pi & : & E_1 \times_B E_2 & \to & B \\ & & (e_1, e_2) & \mapsto & \pi(e_1, e_2) = \pi_1(e_1) = \pi_2(e_2) \end{array} \,. \tag{2.8.2}$$

The construction is illustrated in figure 2.6. At each point $p \in B$, the fiber $\pi^{-1}(p)$ is given by the product $\pi_1^{-1}(p) \times \pi_2^{-1}(p)$. To understand the geometry of the total space E constituted by these fibers, it is instructive to construct the local trivializations of this bundle, which are then used to construct induced charts on E. For $p \in B$, one can find local trivializations (U_1, ϕ_1) and (U_2, ϕ_2) of the two factor bundles. Since U_1 and U_2 are open sets, also their intersection $U = U_1 \cap U_2$ is open. To construct a local trivialization of (E, B, π, F) with set U, one still need a diffeomorphism $\phi : \pi^{-1}(U) \to U \times F$. Since $F = F_1 \times F_2$ is a product manifold, this is obtained from the known trivializations as

$$\phi(e_1, e_2) = (\pi(e_1, e_2), (\mathrm{pr}_2 \circ \phi_1)(e_1), (\mathrm{pr}_2 \circ \phi_2)(e_2)) \,. \tag{2.8.3}$$

Here pr_2 denotes the projection onto the second factor in $U \times F_1$ in the first occurrence, and in $U \times F_2$ in the second occurrence.

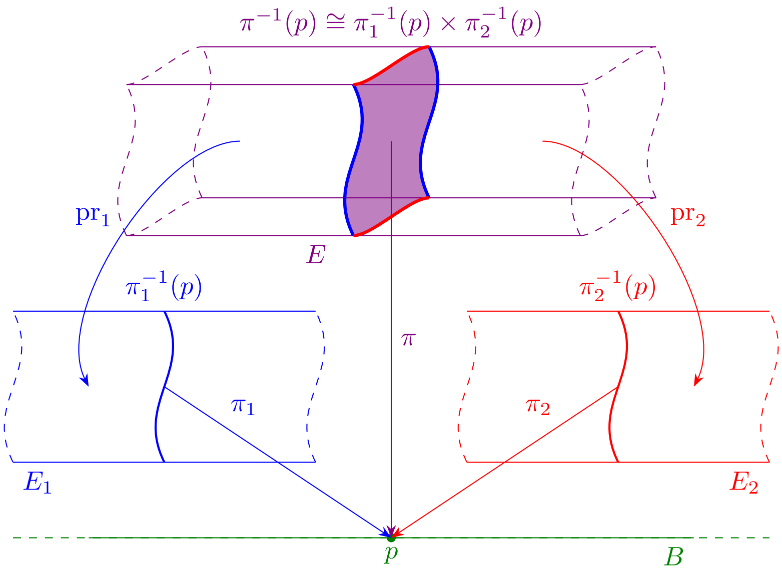


Figure 2.6: Illustration of a fibered product. Over each point $p \in B$, the fiber $\pi^{-1}(p)$ is the product manifold $\pi_1^{-1}(p) \times \pi_2^{-1}(p)$.

Recalling that the total space of a fibered product is as subset of the Cartesian product, $E \subset E_1 \times E_2$, one can restrict the projections maps defined by the latter to maps $\mathrm{pr}_1 : E \to E_1$ and $\mathrm{pr}_2 : E \to E_2$. Denoting the elements of $E \subset E_1 \times E_2$ as pairs (e_1, e_2) (keeping in mind that they share a common base point in B), these projections have the known form

$$\mathrm{pr}_1(e_1, e_2) = e_1\,, \quad \mathrm{pr}_2(e_1, e_2) = e_2 \tag{2.8.4}$$

and relate the total spaces of the fibered product and its constituent factors. This operation plays a similar role for fiber bundles as the projection from a product manifold onto its factor manifolds. One finds that these maps have a few useful properties. To reveal these, it is helpful to first notice that there are two more fiber bundles hidden in the construction above.

Theorem 2.8.1. *Given a fibered product $E = E_1 \times_B E_2$, the tuples $(E, E_1, \mathrm{pr}_1, F_2)$ and $(E, E_2, \mathrm{pr}_2, F_1)$ are fiber bundles.*

Proof. Here we show the proof for $(E, E_1, \mathrm{pr}_1, F_2)$; the proof for $(E, E_2, \mathrm{pr}_2, F_1)$ proceeds analogously. First note that $\mathrm{pr}_1 : E \to E_1$ is obviously surjective. We will show its smoothness along with the construction of the local trivializations. For this purpose, let $e_1 \in E_1$ and $p = \pi_1(e_1) \in B$. Then there exists $U \subset B$ with $p \in U$ such that there is a local trivialization (U, ϕ) of E of the form (2.8.3), with $\phi : \pi^{-1}(U) \to U \times F_1 \times F_2$, as well as a local trivialization $\phi_1 : \pi_1^{-1}(U) \to U \times F_1$. Now define $V = \pi_1^{-1}(U) \subset E_1$. Note that $e_1 \in V$ and that $\pi^{-1}(U) = \mathrm{pr}_1^{-1}(V)$. On this set, we can define a diffeomorphism

$$\begin{array}{cccc} \psi & : & \pi^{-1}(U) & \to & V \times F_2 \\ & & (e_1, e_2) & \mapsto & (e_1, \mathrm{pr}_3(\phi(e_1, e_2))) \end{array}. \tag{2.8.5}$$

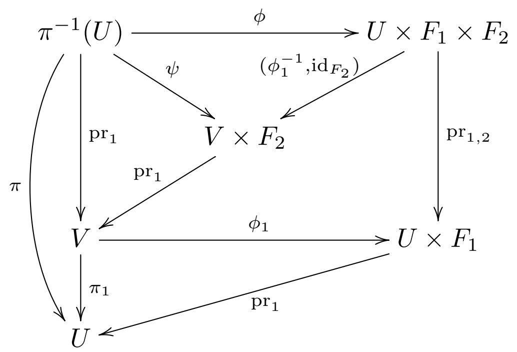
Using the trivialization ϕ_1, we have

$$\mathrm{pr}_{1,2}(\phi(e_1, e_2)) = \phi_1(e_1)\,, \tag{2.8.6}$$

and so we can also write ψ as

$$\psi = (\phi_1^{-1}, \mathrm{id}_{F_2}) \circ \phi\,. \tag{2.8.7}$$

The relations between the involved maps is visualized in the following commutative diagram:



(2.8.8)

We see that the upper left triangle between $\pi^{-1}(U)$, $V \times F_1$ and V commutes, showing that (V, ψ) is a local trivialization. This also shows that $\mathrm{pr}_1 : \pi^{-1}(U) \to V$ is smooth, since it is composed of the smooth maps $\psi : \pi^{-1}(U) \to V \times F_2$ and $\mathrm{pr}_1 : V \times F_2 \to V$. Hence, $(E, E_1, \mathrm{pr}_1, F_2)$ is a fiber bundle. ■

Since there are several fiber bundles involved in this construction, it is natural to expect that these are related by certain bundle morphisms. It turns out that this is indeed the case. The first class of bundle morphisms is the following.

Theorem 2.8.2. *Given a fibered product $E = E_1 \times_B E_2$, the maps $\mathrm{pr}_1 : E \to E_1$ and $\mathrm{pr}_2 : E \to E_2$ defined above are bundle morphisms from (E, B, π, F) to (E_1, B, π_1, F_1) and (E_2, B, π_2, F_2), respectively, covering the identity on B.*

Proof. The smoothness of pr_1 and pr_2 was was already shown in theorem 2.8.1, so that here we restrict ourselves to proving the commutativity of the diagram given in definition 2.7.2. This simply follows from the constructing of the projection $\pi : E \to B$ of the fibered product, which satisfies

$$\pi(e_1, e_2) = \pi_1(e_1) = \pi_1(\mathrm{pr}_1(e_1, e_2)), \quad \pi(e_1, e_2) = \pi_2(e_2) = \pi_2(\mathrm{pr}_2(e_1, e_2)) \tag{2.8.9}$$

for all $(e_1, e_2) \in E$, and hence

$$\pi = \pi_1 \circ \mathrm{pr}_1, \quad \pi = \pi_2 \circ \mathrm{pr}_2 . \tag{2.8.10}$$

■

This result shows once more the aforementioned similarity of a fibered product to a product manifold. While the latter is canonically equipped with maps which project onto the constituting factors, in the case of a fibered product, these are promoted to bundle morphisms, and therefore preserve the fiber bundle structure over B. Two more bundle morphisms are the following.

Theorem 2.8.3. *Given a fibered product $E = E_1 \times_B E_2$, the maps $\mathrm{pr}_1 : E \to E_1$ and $\mathrm{pr}_2 : E \to E_2$ defined above are bundle morphisms from $(E, E_2, \mathrm{pr}_2, F_1)$ to (E_1, B, π_1, F_1), covering $\pi_2 : E_2 \to B$, and from $(E, E_1, \mathrm{pr}_1, F_2)$ to (E_2, B, π_2, F_2), covering $\pi_1 : E_1 \to B$, respectively.*

Proof. Again we will not show the smoothness of the involved maps, as it was shown in theorem 2.8.1. To show the commutativity of the diagram from definition 2.7.1, note that by construction we have

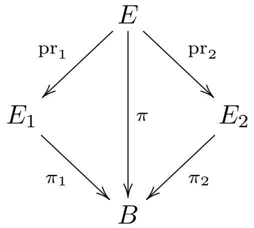
$$\pi_1(\mathrm{pr}_1(e_1, e_2)) = \pi_1(e_1) = \pi_2(e_2) = \pi_2(\mathrm{pr}_2(e_1, e_2)) \tag{2.8.11}$$

for all $(e_1, e_2) \in E$, and hence

$$\pi_1 \circ \mathrm{pr}_1 = \pi_2 \circ \mathrm{pr}_2 . \tag{2.8.12}$$

Identifying $\mathrm{pr}_1 : E \to E_1$ as map between total spaces, $\mathrm{pr}_2 : E \to E_2$ and $\pi_1 : E_1 \to B$ as bundle projections and $\pi_2 : E_2 \to B$ as map between base manifolds, the first part of the statement follows. The second part is obtained in full analogy by exchanging $1 \leftrightarrow 2$. ■

The two previous statements are easily understood by means of the diagram



(2.8.13)

which commutes by construction of the fibered product. By identifying the appropriate maps in this diagram either as bundle projections or as maps between bundles, its commutativity yields the proofs of the aforementioned statements.

The projection maps are also useful for understanding the structure of sections of a fibered product. It turns out that every section of (E, B, π, F) can equivalently be expressed as a pair of sections of the two factor bundles. This is formalized as follows.

Theorem 2.8.4. *There is a one-to-one correspondence between sections $\sigma : B \to E$ of a fibered product $E = E_1 \times_B E_2$ and pairs (σ_1, σ_2) of sections, where $\sigma_1 : B \to E_1$ is a section of E_1 and $\sigma_2 : B \to E_2$ is a section of E_2.*

Proof. Given a section $\sigma : B \to E$, one can easily obtain sections $\sigma_1 = \mathrm{pr}_1 \circ \sigma : B \to E_1$ and $\sigma_2 = \mathrm{pr}_2 \circ \sigma : B \to E_2$. These are sections, since pr_1 and pr_2 are bundle morphisms covering the identity.

Conversely, given sections $\sigma_1 : B \to E_1$ and $\sigma_2 : B \to E_2$, one can construct a section $\sigma : B \to E$ by defining

$$\sigma(p) = (\sigma_1(p), \sigma_2(p)) \tag{2.8.14}$$

for every $p \in B$. This is an element of E, since

$$\pi_1(\sigma_1(p)) = p = \pi_2(\sigma_2(p)) \,, \tag{2.8.15}$$

which follows from the fact that σ_1 and σ_2 are sections. Finally, one easily checks that this construction is the inverse of the construction of σ_1 and σ_2 given above. ■

▶Explain section in induced coordinates◀

2.9 Pullback bundles

We have seen that in order to specify a bundle morphism, it is sufficient to specify only one map θ between the total spaces of two bundles, since the map $\tilde{\theta}$ covered by θ is uniquely defined by the commutativity of the diagram (2.7.1). However, there are particular bundle morphisms which are uniquely defined by a map relating only the base manifolds. One class of such bundles, which we will encounter later, is called *natural bundles*. Another construction which yields this property is the following.

Definition 2.9.1 (Pullback bundle)**.** Let (E, B, π, F) be a fiber bundle and M a manifold together with a map $\psi : M \to B$. The *pullback bundle* (or *induced bundle*) of (E, B, π, F) along ψ is the fiber bundle $(\psi^*E, M, \psi^*\pi, F)$, where

$$\psi^*E = \{(m, e) \in M \times E, \psi(m) = \pi(e)\} \tag{2.9.1}$$

and

$$\begin{array}{rcccc} \psi^*\pi & : & \psi^*E & \to & M \\ & & (m, e) & \mapsto & m \end{array} . \tag{2.9.2}$$

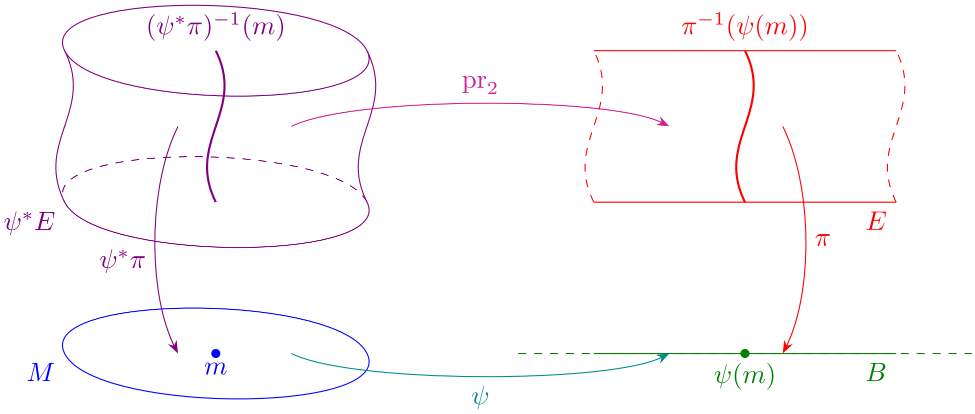


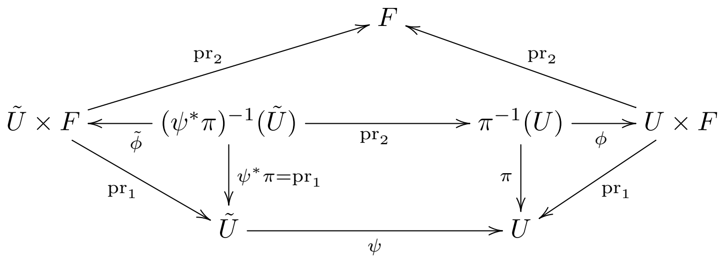
Figure 2.7: Illustration of a pullback bundle. Over each point $m \in M$, the fiber $(\psi^*\pi)^{-1}(m)$ is isomorphic to the fiber $\pi^{-1}(\psi(m))$. This isomorphism is mediated by the projection $\mathrm{pr}_2 : \psi^*E \to E$, which is defined from $\psi^*E \subset M \times E$.

The geometric picture behind this construction is that for every base point $m \in M$, the fiber $(\psi^*E)_m$ over m is a copy of the fiber $E_{\psi(m)}$ over $\psi(m)$, which is "pulled back" along ψ. In order to describe the manifold structure on ψ^*E, one most easily constructs the local trivializations of the pullback bundle, which then define an atlas via the induced charts as shown in section 2.4. For $m \in M$, let (U, ϕ) be a local trivialization of (E, B, ϕ, F) such that $\psi(m) \in U$. Then define $\tilde{U} = \psi^{-1}(U) \subset M$ and

$$\begin{array}{cccc} \tilde{\phi} & : & (\psi^*\pi)^{-1}(\tilde{U}) & \to & \tilde{U} \times F \\ & & (m, e) & \mapsto & (m, \mathrm{pr}_2(\phi(e))) \end{array} . \tag{2.9.3}$$

This construction can be visualized in the following commutative diagram:

(2.9.4)



To clarify this, recall that an element $(m, e) \in (\psi^*\pi)^{-1}(\tilde{U})$ is a pair such that $m \in \tilde{U}$ and $e \in E_{\psi(m)}$, i.e., $e \in E$ with $\pi(e) = \psi(m)$. The latter equality, together with the properties $\mathrm{pr}_2(m, e) = e$ and

$$(\psi^*\pi)(m, e) = \mathrm{pr}_1(m, e) = m \tag{2.9.5}$$

explains the lower central rectangle of the diagram (2.9.4). The lower right triangle commutes, since ϕ is a local trivialization by assumption. The lower left triangle commutes by the definition (2.9.3) of $\tilde{\phi}$, which leaves the first component unchanged, such that $\mathrm{pr}_1 \circ \tilde{\phi} = \mathrm{pr}_1$. Finally, the upper triangle also commutes due to the definition of $\tilde{\phi}$, since for the second component holds

$$(\mathrm{pr}_2 \circ \tilde{\phi})(m, e) = (\mathrm{pr}_2 \circ \phi)(e) = (\mathrm{pr}_2 \circ \phi \circ \mathrm{pr}_2)(m, e) . \tag{2.9.6}$$

It follows that $\tilde{\phi}$ is smooth, since it is constituted by smooth maps. To see that it is a diffeomorphism, one may explicitly construct its inverse

$$\begin{array}{cccc} \tilde{\phi}^{-1} & : & \tilde{U} \times F & \to & (\psi^*\pi)^{-1}(\tilde{U}) \\ & & (m, v) & \mapsto & (m, \phi^{-1}(\psi(m), v)) \end{array} . \tag{2.9.7}$$

One first has to check that this is well-defined. For $(m, v) \in \tilde{U} \times F$ holds $(\psi(m), v) \in U \times F$ by the definition of $\tilde{U}$. Hence, we can apply ϕ^{-1} to obtain an element of $\pi^{-1}(U)$. Now obviously holds

$$\pi(\phi^{-1}(\psi(m), v)) = \mathrm{pr}_1(\psi(m), v) = \psi(m)\,, \tag{2.9.8}$$

so that $(m, \phi^{-1}(\psi(m), v)) \in (\psi^*\pi)^{-1}(\tilde{U})$, as required. To check that it is indeed the inverse of $\tilde{\phi}$, we calculate

$$\begin{aligned}(\tilde{\phi} \circ \tilde{\phi}^{-1})(m, v) &= \tilde{\phi}(m, \phi^{-1}(\psi(m), v)) \\ &= (m, \mathrm{pr}_2(\phi(\phi^{-1}(\psi(m), v)))) \\ &= (m, \mathrm{pr}_2(\psi(m), v)) \\ &= (m, v)\,.\end{aligned} \tag{2.9.9}$$

We must also check the opposite direction, which reads

$$\begin{aligned}(\tilde{\phi}^{-1} \circ \tilde{\phi})(m, e) &= \tilde{\phi}^{-1}(m, \mathrm{pr}_2(\phi(e))) \\ &= (m, \phi^{-1}(\psi(m), \mathrm{pr}_2(\phi(e)))) \\ &= (m, \phi^{-1}(\pi(e), \mathrm{pr}_2(\phi(e)))) \\ &= (m, \phi^{-1}(\mathrm{pr}_1(\phi(e)), \mathrm{pr}_2(\phi(e)))) \\ &= (m, \phi^{-1}(\phi(e))) \\ &= (m, e)\,.\end{aligned} \tag{2.9.10}$$

Hence, we have indeed constructed the inverse of $\tilde{\phi}$. Finally, since $\tilde{\phi}^{-1}$ is constructed from smooth maps, it is also smooth. This finally proves that $\tilde{\phi}$ is a diffeomorphism, so that $(\tilde{U}, \tilde{\phi})$ is a trivialization. Since we can construct such a pair $(\tilde{U}, \tilde{\phi})$ around any point $m \in M$ from a local trivialization (U, ϕ) around $\psi(m) \in B$, it follows that $(\psi^*E, M, \psi^*\pi, F)$ is indeed a fiber bundle.

Pullback bundles are interesting also in physics because they allow studying maps whose domain is one manifold M, but whose codomain is the total space E of a fiber bundle over a different manifold B. A typical example occurs when one discusses parallel transport, where $M = \mathbb{R}$, and one is interested in elements of the fibers of E only over the image of a curve $\gamma : M \to E$. These are described by sections of the corresponding pullback bundle γ^*E. This turns out to be a special case of the following statement.

Theorem 2.9.1. *Let (E, B, π, F) be a fiber bundle and M a manifold together with a map $\psi : M \to B$, and denote the pullback bundle by $(\psi^*E, M, \psi^*\pi, F)$. Then there is a one-to-one correspondence between sections σ of $(\psi^*E, M, \psi^*\pi, F)$ and maps $\hat{\sigma} : M \to E$ satisfying $\pi \circ \hat{\sigma} = \psi$.*

Proof. Let first $\sigma : M \to \psi^*E$ be a section. By definition of the pullback bundle it satisfies

$$\psi = \psi \circ \psi^*\pi \circ \sigma = \psi \circ \mathrm{pr}_1 \circ \sigma = \pi \circ \mathrm{pr}_2 \circ \sigma\,. \tag{2.9.11}$$

We may hence define $\hat{\sigma} = \mathrm{pr}_2 \circ \sigma : M \to E$, and obtain the desired map. Conversely, given such a map $\hat{\sigma}$, one may construct

$$\begin{array}{cccc}\sigma & : & M & \to & \psi^*E \\ & & m & \mapsto & (m, \hat{\sigma}(m))\end{array}\,. \tag{2.9.12}$$

The latter is indeed an element of ψ^*E, since

$$(\pi \circ \hat{\sigma})(m) = \psi(m)\,. \tag{2.9.13}$$

Also it is clear that σ is a section. One now easily checks that the two assignments $\sigma \mapsto \hat{\sigma}$ and $\hat{\sigma} \mapsto \sigma$ described above are inverses of each other. ■

The appearance of the pullback bundle can be seen also in the following two constructions.

Definition 2.9.2 (Pullback section)**.** Let (E, B, π, F) be a fiber bundle and M a manifold together with a map $\psi : M \to B$, and denote the pullback bundle by $(\psi^* E, M, \psi^* \pi, F)$. Given a section $\sigma : B \to E$ of E, its *pullback* is defined as the section

$$\begin{array}{cccc} \psi^* \sigma & : & M & \to & \psi^* E \\ & & m & \mapsto & (m, \sigma(\psi(m))) \end{array} . \tag{2.9.14}$$

This is well-defined, since

$$(\pi \circ \sigma \circ \psi)(m) = \psi(m) , \tag{2.9.15}$$

so that $(m, \sigma(\psi(m))) \in \psi^* E$, and obviously satisfies $(\psi^* \pi) \circ (\psi^* \sigma) = \mathrm{id}_M$, so that it is indeed a section. This construction now allows us to work with objects in the fibers over B as if they had been defined over M instead. One may also go in the opposite direction, as we will show next.

Theorem 2.9.2. *Let (E, B, π, F) be a fiber bundle and M a manifold together with a map $\psi : M \to B$, and denote the pullback bundle by $(\psi^* E, M, \psi^* \pi, F)$. The projection*

$$\begin{array}{ccccc} \mathrm{pr}_2 & : & \psi^* E & \to & E \\ & & (m, e) & \mapsto & e \end{array} \tag{2.9.16}$$

is a bundle morphism covering ψ.

Proof. Obviously pr_2 is smooth, since projections are smooth. To see that it is a bundle morphism, one calculates

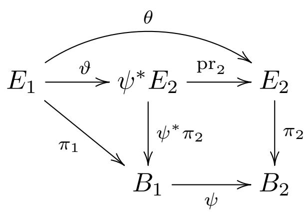
$$(\pi \circ \mathrm{pr}_2)(m, e) = \pi(e) = \psi(m) = (\psi \circ \mathrm{pr}_1)(m, e) = (\psi \circ \psi^* \pi)(m, e) , \tag{2.9.17}$$

where $\pi(e) = \psi(m)$ follows from $(m, e) \in \psi^* E$. ■

Finally, pullback bundles allow us to relate general bundle morphisms to bundle morphisms covering the identity as follows.

Theorem 2.9.3. *Let (E_1, B_1, π_1, F_1) and (E_2, B_2, π_2, F_2) be fiber bundles and $\psi : B_1 \to B_2$ a map. Then there exists a one to one correspondence between bundle morphisms $\theta : E_1 \to E_2$ covering ψ and bundle morphisms $\vartheta : E_1 \to \psi^* E_2$ covering id_{B_1}.*

Proof. We first visualize the maps and spaces mentioned above in the following diagram.



(2.9.18)

Assuming that ϑ is a bundle morphism, the lower left triangle of this diagram commutes. Also the lower right rectangle commutes due to theorem 2.9.2. Since both ϑ and pr_2 are bundle morphisms, also their composition $\theta = \mathrm{pr}_2 \circ \vartheta$ is a bundle morphism. This can also be seen directly from

$$\begin{aligned} \pi_2 \circ \theta &= \pi_2 \circ \mathrm{pr}_2 \circ \vartheta \\ &= \psi \circ (\psi^* \pi_2) \circ \vartheta \\ &= \psi \circ \pi_1 . \end{aligned} \tag{2.9.19}$$

Conversely, given a bundle morphism θ on can construct ϑ as

$$\vartheta(e) = (\pi_1(e), \theta(e)) . \tag{2.9.20}$$

The right hand side is an element of the fiber $(\psi^*\pi_2)^{-1}(\pi_1(e))$, since

$$(\pi_2 \circ \mathrm{pr}_2)(\pi_1(e), \theta(e)) = (\pi_2 \circ \theta)(e) = (\psi \circ \pi_1)(e) = (\psi \circ \mathrm{pr}_1)(\pi_1(e), \theta(e)) \tag{2.9.21}$$

and

$$(\psi^*\pi_2)(\pi_1(e), \theta(e)) = \mathrm{pr}_1(\pi_1(e), \theta(e)) = \pi_1(e) . \tag{2.9.22}$$

Thus, ϑ obtained from this construction is a bundle morphism covering the identity. One now easily checks that these two constructions are in one to one correspondence. ■

Chapter 3

Vector bundles

3.1 Vector bundles

Often we encounter fiber bundles whose fibers are not just manifolds, but also carry additional structure. The most common structure is that of a (real or complex) vector space. In this case the fiber bundle is called a *vector bundle*. To keep things simple for now, we will restrict ourselves to real vector bundles, which are defined as follows.

Definition 3.1.1 (Vector bundle)**.** A (real) *vector bundle* of rank $k \in \mathbb{N}$ is a fiber bundle $(E, B, \pi, \mathbb{R}^k)$ such that for all $p \in B$ the fiber $E_p = \pi^{-1}(p)$ is a real vector space of dimension k and such that the restrictions of the local trivializations $\phi : \pi^{-1}(U) \to U \times \mathbb{R}^k$ to a fiber E_p for $p \in U$ are vector space isomorphisms from E_p to $\{p\} \times \mathbb{R}^k$.

One may ask why we want the local trivializations to restrict to vector space isomorphisms. This is a typical example for a very common situation that we have two different structures, here that of a manifold and that of a vector space, which we want to be *compatible*. In this case it guarantees that on every fiber E_p for $p \in B$, which is both a manifold diffeomorphic to $\mathbb{R}^k$ and a vector space isomorphic to $\mathbb{R}^k$, both

1. the scalar multiplication $\cdot : \mathbb{R} \times E_p \to E_p$
2. and the addition $+ : E_p \times E_p \to E_p$

are smooth maps. Further, it guarantees that if $p, p' \in B$ are "close to each other", then:

1. The zero elements of the vector spaces E_p and $E_{p'}$ are "close to each other".
2. If $v \in E_p$ and $v' \in E_{p'}$ are "close to each other", then also λv and $\lambda v'$ are "close to each other" for any $\lambda \in \mathbb{R}$.
3. If $v \in E_p$ is "close to" $v' \in E_{p'}$ and $w \in E_p$ is "close to" $w' \in E_{p'}$, then also $v + w$ and $v' + w'$ are "close to each other".

Of course, we need to define what we mean by being "close to each other", and why this is necessary. The necessity will become clear later, when we discuss sections of vector bundles. One consequence of the aforementioned "closeness" is the following helpful statement.

Theorem 3.1.1. *For every vector bundle $(E, B, \pi, \mathbb{R}^k)$ the functions*

$$\begin{array}{cccc} + & : & E \times_B E & \to & E \\ & & (e_1, e_2) & \mapsto & e_1 + e_2 \end{array} \tag{3.1.1}$$

and

$$\begin{array}{cccc} \cdot & : & (B \times \mathbb{R}) \times_B E & \to & E \\ & & ((\pi(e), \lambda), e) & \mapsto & \lambda e \end{array} \tag{3.1.2}$$

are fiber bundle morphisms covering the identity on B.

Proof. For $p \in B$, let $U \subset B$ with $p \in U$ such that there exists a chart (U, ψ) of B with $\psi : U \to \mathbb{R}^n$ and a local trivialization (U, ϕ) of E with $\phi : \pi^{-1}(U) \to U \times \mathbb{R}^k$. This allows us to define an induced chart $(\tilde{U}, \tilde{\psi})$ of E with $\tilde{U} = \pi^{-1}(U)$ and

$$\tilde{\psi}(e) = (\psi(\pi(e)), \mathrm{pr}_2(\phi(e))) \in \mathbb{R}^n \times \mathbb{R}^k \cong \mathbb{R}^{n+k}\,. \tag{3.1.3}$$

We will use these components to check that the two functions given above are smooth maps:

First, we construct the set

$$V = (\pi \times_B \pi)^{-1}(U) = \{(e_1, e_2) \in E \times E, \pi(e_1) = \pi_(e_2) \in U\}\,, \tag{3.1.4}$$

on which we define an induced chart (V, χ) by

$$\begin{array}{cccc} \chi & : & V & \to & \mathbb{R}^{n+2k} \\ & & (e_1, e_2) & \mapsto & (\psi(\pi(e_1)), \mathrm{pr}_2(\phi(e_1)), \mathrm{pr}_2(\phi(e_2))) \end{array}\,. \tag{3.1.5}$$

On this chart, the addition takes the form

$$\begin{array}{cccc} \tilde{\psi} \circ + \circ \chi^{-1} & : & \chi(V) & \to & \tilde{\psi}(\tilde{U}) \\ & & (u, v, w) & \mapsto & (u, v + w) \end{array}\,, \tag{3.1.6}$$

where we used the fact that ϕ restricts to a vector space isomorphism on each fiber, and hence

$$\mathrm{pr}_2(\phi(e_1 + e_2)) = \mathrm{pr}_2(\phi(e_1)) + \mathrm{pr}_2(\phi(e_2))\,, \tag{3.1.7}$$

while the base point remains unchanged,

$$\pi(e_1 + e_2) = \pi(e_1) = \pi(e_2)\,. \tag{3.1.8}$$

The function $(u, v, w) \mapsto (u, v + w)$ is smooth, and so it follows that addition is smooth.

We proceed analogously with the multiplication. Let

$$W = \{((b, \lambda), e) \in (B \times \mathbb{R}) \times E, b = \pi(e) \in U\}\,, \tag{3.1.9}$$

and define an induced chart (W, ω) by

$$\begin{array}{cccc} \omega & : & W & \to & \mathbb{R}^{n+k+1} \\ & & ((\pi(e), \lambda), e) & \mapsto & (\psi(\pi(e)), \lambda, \mathrm{pr}_2(\phi(e))) \end{array}\,. \tag{3.1.10}$$

On this chart, the scalar multiplication becomes

$$\begin{array}{cccc} \tilde{\psi} \circ \cdot \circ \omega^{-1} & : & \omega(W) & \to & \tilde{\psi}(\tilde{U}) \\ & & (u, \lambda, v) & \mapsto & (u, \lambda v) \end{array}\,, \tag{3.1.11}$$

where we once again used the fiber-wise linearity of ϕ to conclude

$$\mathrm{pr}_2(\phi(\lambda e)) = \lambda\, \mathrm{pr}_2(\phi(e))\,, \tag{3.1.12}$$

as well as the fact that these operations do not change the base point. Since $(u, \lambda, v) \mapsto (u, \lambda v)$ is smooth, we find that also scalar multiplication is smooth.

Finally, we have seen that both for addition and scalar multiplication the base point is unchanged. Hence, both operations define bundle morphisms covering the identity on B. ∎

We continue with a motivating example.

Example 3.1.1 (Möbius strip as a vector bundle). In the last section we discussed the Möbius strip as a fiber bundle $(M, S^1, \pi, (-1,1))$. However, the open interval $(-1,1)$ and the real line $\mathbb{R}$ are diffeomorphic, one-dimensional manifolds, so that one can also view the Möbius strip as a fiber bundle $(M, S^1, \pi, \mathbb{R})$, which one may call an "infinite Möbius strip". This can be seen most easily by changing the charts from our previous definition 2.2.1 such that $\tilde{U}_1 = (0, 2\pi) \times \mathbb{R}$, $\tilde{U}_2 = (-\pi, \pi) \times \mathbb{R}$ and the functions

$$\begin{array}{rccl} \tilde{\phi}_i & : & \tilde{U}_i & \to & \mathbb{R}^3 \\ & & (t,s) & \mapsto & \left(\left(R + \frac{Ws}{\sqrt{1+s^2}}\cos\frac{t}{2}\right)\cos t, \left(R + \frac{Ws}{\sqrt{1+s^2}}\cos\frac{t}{2}\right)\sin t, \frac{Ws}{\sqrt{1+s^2}}\sin\frac{t}{2}\right) \end{array} . \tag{3.1.13}$$

On each fiber $\pi^{-1}(p) \cong \mathbb{R}$ one has the usual structure of the one-dimensional vector space $\mathbb{R}$.

3.2 Induced charts and fiber coordinates

The properties given in the definition of vector bundles allow the construction of very convenient induced charts from the local trivializations and an atlas on the base manifold B alone, since the fiber $F = \mathbb{R}^k$ is canonically equipped with an atlas containing the single chart $(V, \chi) = (\mathbb{R}^k, \mathrm{id}_{\mathbb{R}^k})$. Using the same notation as in definition 2.4.1, the diagram (2.4.5) therefore reduces to

$$\begin{array}{ccccc} \pi^{-1}(U) & \xrightarrow{\phi} & U \times \mathbb{R}^k & & \\ \downarrow{\scriptstyle\pi} & \swarrow{\scriptstyle \mathrm{pr}_1} & & \searrow{\scriptstyle \mathrm{pr}_2} & \\ U & & & & \mathbb{R}^k \\ \downarrow{\scriptstyle\psi} & & & & \\ \mathbb{R}^m & & & & \end{array} \tag{3.2.1}$$

For the induced chart (W, ω) we thus have $W = \phi^{-1}(U \times \mathbb{R}^k) = \pi^{-1}(U)$ and

$$\omega(e) = (\psi(\pi(e)), \mathrm{pr}_2(\phi(e))) \tag{3.2.2}$$

for $e \in W$. Since ϕ restricts to isomorphisms of vector spaces on every fiber E_p, the same holds for the combined function

$$E_p \xrightarrow{\phi} \{p\} \times \mathbb{R}^k \xrightarrow{\mathrm{pr}_2} \mathbb{R}^k \tag{3.2.3}$$

This means that for $v, w \in E_p$ and $\lambda, \mu \in \mathbb{R}$ we have

$$(\mathrm{pr}_2 \circ \phi)(\lambda v + \mu w) = \lambda(\mathrm{pr}_2 \circ \phi)(v) + \mu(\mathrm{pr}_2 \circ \phi)(w) \,. \tag{3.2.4}$$

Denoting the coordinates on B by $(x^i, i = 1, \ldots, m)$ and the coordinates on $F = \mathbb{R}^k$ by $(y^\mu, \mu = 1, \ldots, k)$, so that the coordinates in the induced chart are (x^i, y^μ) and the coordinates x^i are constant along each fiber E_p, the vector space operations act on the fiber coordinates y^μ.

Another possible way to view the fiber coordinates on E_p is by realizing that $\mathbb{R}^k$ is equipped with a canonical basis $(\mathbb{e}_\mu, \mu = 1, \ldots, k)$, so that the coordinates y^μ are simply the components of an element $y = y^\mu \mathbb{e}_\mu$ with respect to this basis. Using the fact that local trivializations restrict to linear vector space isomorphisms on each fiber, this allows us to define the following notion:

Definition 3.2.1 (Coordinate basis)**.** Let $(E, B, \pi, \mathbb{R}^k)$ be a vector bundle, $p \in B$ and (U, ϕ) a local trivialization such that $p \in U$. Denote by $(\mathbb{e}_\mu, \mu = 1, \ldots, k)$ the canonical basis of $\mathbb{R}^k$. The *coordinate basis* on the fiber $E_p = \pi^{-1}(p)$ induced by the local trivialization (U, ϕ) is the basis $(\epsilon_\mu, \mu = 1, \ldots, k)$ defined by

$$\epsilon_\mu(p) = \phi^{-1}(p, \mathbb{e}_\mu) \in E_p\,. \tag{3.2.5}$$

With the help of the coordinate basis, we can now write every element of E_p in the form $y^\mu \epsilon_\mu(p)$, where y^μ are the fiber coordinates. Obviously the coordinate basis depends on the choice of the local trivialization. However, we know that any two bases of a vector space are related by a linear transformation. It follows that given two different trivializations, and hence two induced charts both defining a coordinate basis, these bases, and therefore the induced coordinates, must also be related by linear transformation. We conclude:

Theorem 3.2.1. *For the induced vector bundle charts the transition functions are linear transformations of the fiber coordinates.*

Proof. Let (U, ϕ) and $(\tilde{U}, \tilde{\phi})$ be local trivializations with $U \cap \tilde{U} \neq \emptyset$. Restricting the trivializations to their intersection, we have the commutative diagram

$$\begin{array}{ccccc} (U \cap \tilde{U}) \times \mathbb{R}^k & \xleftarrow{\ \phi\ } & \pi^{-1}(U \cap \tilde{U}) & \xrightarrow{\ \tilde{\phi}\ } & (U \cap \tilde{U}) \times \mathbb{R}^k\,. \\ & \searrow^{\mathrm{pr}_1} & \downarrow{\scriptstyle \pi} & \swarrow_{\mathrm{pr}_1} & \\ & & U & & \end{array} \tag{3.2.6}$$

We see that $\phi|_{\pi^{-1}(U\cap\tilde{U})}$ and $\tilde{\phi}|_{\pi^{-1}(U\cap\tilde{U})}$ are diffeomorphisms which both map $\pi^{-1}(U \cap \tilde{U})$ to $(U \cap \tilde{U}) \times \mathbb{R}^k$, and so we have a diffeomorphism

$$\tilde{\phi}|_{\pi^{-1}(U\cap\tilde{U})} \circ (\phi|_{\pi^{-1}(U\cap\tilde{U})})^{-1} : (U \cap \tilde{U}) \times \mathbb{R}^k \to (U \cap \tilde{U}) \times \mathbb{R}^k \tag{3.2.7}$$

and its inverse. Further restricting these to the fiber over a point p, we have the diagram

$$\{p\} \times \mathbb{R}^k \xleftarrow{\ \phi\ } E_p \xrightarrow{\ \tilde{\phi}\ } \{p\} \times \mathbb{R}^k \tag{3.2.8}$$

and the respective restrictions $\phi|_{E_p}$ and $\tilde{\phi}|_{E_p}$ are vector space isomorphisms. Hence, also $\tilde{\phi}_{E_p} \circ (\phi|_{E_p})^{-1}$ and its inverse are vector space isomorphisms, and thus linear transformations of the fiber coordinates at the point p. ■

This can also easily be seen if we write the corresponding induced fiber coordinates as y^μ and $\tilde{y}^\mu$, the bases as $\epsilon_\mu(p)$ and $\tilde{\epsilon}_\mu(p)$, as well as their duals as $\bar{\epsilon}^\mu(p)$ and $\bar{\tilde{\epsilon}}^\mu(p)$. From the relation

$$y^\mu \epsilon_\mu(p) = \tilde{y}^\mu \tilde{\epsilon}_\mu(p) \tag{3.2.9}$$

follows

$$y^\mu = \tilde{y}^\nu \langle \bar{\epsilon}^\mu(p), \tilde{\epsilon}_\nu(p) \rangle\,, \quad \tilde{y}^\mu = y^\nu \langle \bar{\tilde{\epsilon}}^\mu(p), \epsilon_\nu(p) \rangle\,, \tag{3.2.10}$$

where we wrote $\langle \bullet, \bullet \rangle$ for the canonical pairing between elements of E_p and its dual E_p^*.

3.3 Sections of vector bundles

Vector bundles always admit global sections, the most simple one given in the following example.

Definition 3.3.1 (Zero section)**.** Let $(E, B, \pi, \mathbb{R}^k)$ be a vector bundle. The *zero section* is the map $0 : B \to E$ which assigns to each $p \in B$ the zero element of the vector space $E_p = \pi^{-1}(p)$.

It is not difficult to show that the zero section is indeed a section. We will prove a more general statement here, from which also this property of the zero section follows.

Theorem 3.3.1. *The set of all sections of a (real) vector bundle is a (real) vector space, where scalar multiplication and addition are defined pointwise.*

Proof. Let f, g be sections of a vector bundle $(E, B, \pi, \mathbb{R}^k)$ with $\dim B = n$ and $\lambda, \mu \in \mathbb{R}$. We have to check that also the function $h = \lambda f + \mu g$ defined by

$$\begin{array}{cccc} h & : & B & \to & E \\ & & p & \mapsto & h(p) = \lambda f(p) + \mu g(p) \end{array} \tag{3.3.1}$$

is a smooth section, i.e., a smooth map such that $\pi \circ h = \mathrm{id}_B$. We first have to check that this function is well-defined. Since both f and g are sections, they satisfy $\pi \circ f = \pi \circ g = \mathrm{id}_B$. For any $p \in B$ we thus have $f(p) \in \pi^{-1}(p)$ and $g(p) \in \pi^{-1}(p)$. Since $\pi^{-1}(p)$ carries the structure of a vector space, there is a well-defined element $\lambda f(p) + \mu g(p) = h(p) \in \pi^{-1}(p)$, so that the function h is indeed well-defined. This also shows that $\pi \circ h = \mathrm{id}_B$.

We finally show that h is a smooth map. To see this, let (U, ϕ) be a local trivialization around some point $p \in B$, (U, ψ) a chart defined on the same open set $U \ni p$ and $(\pi^{-1}(U), \omega)$ the corresponding induced chart on E. The functions $\tilde{f} = \omega \circ f \circ \psi^{-1} : \psi(U) \to \mathbb{R}^{n+k}$ and $\tilde{g} = \omega \circ g \circ \psi^{-1} : \psi(U) \to \mathbb{R}^{n+k}$ are smooth, since f and g are smooth. By definition of the induced coordinates, they are of the form

$$\tilde{f}(x^1, \ldots, x^n) = (x^1, \ldots, x^n, f^1(x^1, \ldots, x^n), \ldots, f^k(x^1, \ldots, x^n)), \tag{3.3.2}$$

and analogously for $\tilde{g}$. We then define a function $\tilde{h} : \psi(U) \to \mathbb{R}^{n+k}$ by

$$\tilde{h}(x) = (x, \lambda \tilde{f}(x) + \mu \tilde{g}(x)), \tag{3.3.3}$$

where we wrote $x = (x^1, \ldots, x^n)$. This is smooth, since sums and multiples of smooth functions on $\mathbb{R}^n$ are smooth. Using the fact that in induced coordinates the linear operations on each fiber E_p are represented by the linear operations on the fiber coordinates, it is now easy to see that $h = \omega^{-1} \circ \tilde{h} \circ \psi$ is smooth on U. Finally, this construction holds for all $p \in B$, and thus h is a smooth map.

As an alternative proof, we can also make use of theorem 3.1.1. For this purpose, note that

$$\begin{array}{cccc} f_\lambda & : & B & \to & (B \times \mathbb{R}) \times_B E \\ & & p & \mapsto & ((p, \lambda), f(p)) \end{array} \tag{3.3.4}$$

and

$$\begin{array}{cccc} g_\mu & : & B & \to & (B \times \mathbb{R}) \times_B E \\ & & p & \mapsto & ((p, \mu), g(p)) \end{array} \tag{3.3.5}$$

are smooth sections due to theorem 2.8.4, since f, g and the constant sections $p \mapsto (p, \lambda)$ and $p \mapsto (p, \mu)$ of the trivial bundle $B \times \mathbb{R}$ are smooth sections due to theorem 2.3.3. Since $\cdot : (B \times \mathbb{R}) \times_B E \to E$ is a bundle morphism covering the identity, also $\lambda f = \cdot \circ f_\lambda$ and $\mu g = \cdot \circ g_\mu$ are smooth sections by theorem 2.7.3. It then follows that also

$$\begin{array}{cccc} (\lambda f, \mu g) & : & B & \to & E \times_B E \\ & & p & \mapsto & (\lambda f(p), \mu g(p)) \end{array} \tag{3.3.6}$$

is a smooth section, again using theorem 2.8.4. Finally using that $+ : E \times_B E \to E$ is a bundle morphism covering the identity, it follows again from theorem 2.7.3 that $\lambda f + \mu g = + \cdot (\lambda f, \mu g)$ is a smooth section. ∎

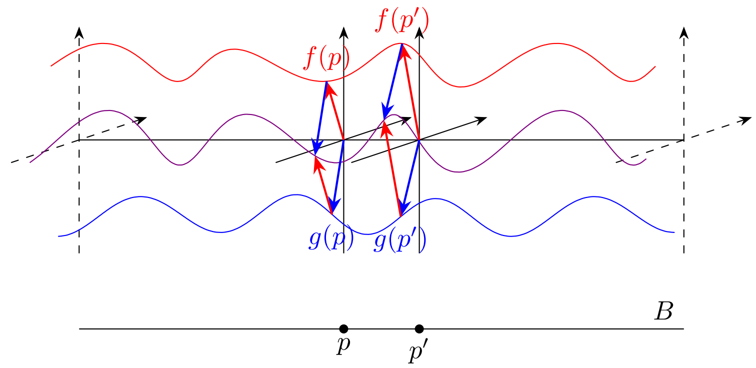


Figure 3.1: Compatibility between geometry and algebraic structure of a vector bundle: if for two "nearby" points p and p' both $f(p)$ is "near" $f(p')$ (i.e., f is a smooth section) and $g(p)$ is "near" $g(p')$ (i.e., g is a smooth section), then the same also holds for their linear combinations.

Now this is the precise notion of what we meant by being "close to each other" in the previous section. It means that if f, g are smooth sections of a vector bundle ("$f(p)$ is close to $f(p')$" if "p is close to p'" and the same for g), then also $\lambda f + \mu g$ is a smooth section ("$\lambda f(p) + \mu g(p)$ is close to $\lambda f(p') + \mu g(p')$") for any $\lambda, \mu \in \mathbb{R}$. In fact, we can even go one step further, and replace the constants μ, ν by functions $\alpha, \beta \in C^\infty(B, \mathbb{R})$, where also multiplication is then defined pointwise. This then leads to the following statement:

Theorem 3.3.2. *The set of all sections of a (real) vector bundle $(E, B, \pi, \mathbb{R}^k)$ is a module over $C^\infty(B, \mathbb{R})$, where scalar multiplication and addition are defined pointwise.*

Proof. The proof is very similar to the proof of theorem 3.3.1 above. First note that for functions $\alpha, \beta \in C^\infty(B, \mathbb{R})$ and sections $f, g \in \Gamma(E)$ the maps

$$\begin{array}{cccl} f_\alpha & : & B & \rightarrow & (B \times \mathbb{R}) \times_B E \\ & & p & \mapsto & ((p, \alpha(p)), f(p)) \end{array} \tag{3.3.7}$$

and

$$\begin{array}{cccl} g_\beta & : & B & \rightarrow & (B \times \mathbb{R}) \times_B E \\ & & p & \mapsto & ((p, \beta(p)), g(p)) \end{array} \tag{3.3.8}$$

are smooth sections. The remaining steps are then fully analogous to the proof of theorem 3.3.1. ∎

Since vector bundles are (a particular type of) fiber bundles, sections are most conveniently described in terms of induced coordinates (x^i, y^μ) on a chart (W, ω) of the total space E, where (x^i) are coordinates on a chart (U, ψ) the base manifold B. In these coordinates a section is expressed by an assignment $(x^i) \mapsto (x^i, y^\mu)$, where only the components y^μ must be specified, while the components x^i are determined by identity functions. However, in the case of a vector bundle we can use the fact that for each $p \in U$ the coordinates y^μ of $f(p)$ are simply its components in the coordinate basis ϵ_μ. Using the dual basis $\bar{\epsilon}^\mu$ of E_p^* these are given by $y^\mu = \bar{\epsilon}^\mu(f(p))$.

3.4 Vector bundle morphisms

In section 2.7 we have discussed bundle morphisms as maps between the total spaces of fiber bundles which preserve the fibers, i.e., which map elements belonging to the same fiber of one bundle to elements of the same fiber of another bundle. In the case of vector bundles, each fiber is equipped with the structure of a vector space, and we are usually interested in maps which also preserve this structure. We can define such maps as follows.

Definition 3.4.1 (Vector bundle morphism)**.** Let $(E_1, B_1, \pi_1, \mathbb{R}^{k_1})$ and $(E_2, B_2, \pi_2, \mathbb{R}^{k_2})$ be vector bundles. A *vector bundle morphism* (or *vector bundle homomorphism*) is a bundle morphism $\theta : E_1 \to E_2$ covering a map $\vartheta : B_1 \to B_2$ such that for each $p \in B_1$ the restriction of θ to the fiber $\pi_1^{-1}(p)$ is a linear function between the vector spaces $\pi_1^{-1}(p)$ and $\pi_2^{-1}(\vartheta(p))$.

As in the case of (general) bundle morphisms, we can define an isomorphism as an invertible morphism.

Definition 3.4.2 (Vector bundle isomorphism)**.** A *vector bundle isomorphism* is a bijective vector bundle morphism whose inverse is also a vector bundle morphism. If a vector bundle morphism between two vector bundles exists, these bundles are called *isomorphic*.

Finally, as in the case for general fiber bundles, also for vector bundles the following statement holds.

Theorem 3.4.1. *Let $(E_i, B_i, \pi_i, \mathbb{R}^{k_i})$ with $i = 1, \ldots, 3$ be vector bundles and $\theta : E_1 \to E_2$ and $\theta' : E_2 \to E_3$ be vector bundle homomorphisms (isomorphisms) covering $\vartheta : B_1 \to B_2$ and $\vartheta' : B_2 \to B_3$. Then $\theta' \circ \theta : E_1 \to E_3$ is a vector bundle homomorphism (isomorphism) covering $\vartheta' \circ \vartheta : B_1 \to B_3$.*

Proof. ▶. . . ◀ ■

Definition 3.4.3 (Rank)**.** Let (E_1, B_1, π_1, F_1) and (E_2, B_2, π_2, F_2) be vector bundles and $\theta : E_1 \to E_2$ a vector bundle morphism covering $\vartheta : B_1 \to B_2$. For $p \in B_1$, the *rank* of θ in p is the rank of the linear function $\theta_p : E_{1p} \to E_{2\vartheta(p)}$. A vector bundle morphism is of *constant rank* if it has the same rank in all points $p \in B_1$.

3.5 Line bundles

Definition 3.5.1 (Line bundle)**.** A *line bundle* over the field $\mathbb{F}$ is a vector bundle $(E, B, \pi, \mathbb{F})$ whose typical fiber is the field $\mathbb{F}$.

Theorem 3.5.1. *For every line bundle $(E, B, \pi, \mathbb{F})$ there exists a one-to-one correspondence between nowhere vanishing sections of E and vector bundle isomorphisms from the trivial line bundle $B \times \mathbb{F}$ to E.*

Proof. ▶. . . ◀ ■

3.6 Subbundles

Often it is necessary to consider not all elements of the total space of a vector bundle, but only elements which belong a particular subset of its total space, which again forms a vector bundle. Recall that given a vector space, such as the fiber E_p over a point $p \in B$ of a vector bundle $(E, B, \pi, \mathbb{R}^k)$, a linear subspace of E_p is a subset which is closed under the vector space operations of addition and scalar multiplication, and thus is again a vector space. In order to generalize this concept to vector bundles, one also needs to preserve their geometric properties, i.e., their differentiable structure. This gives rise to the following notion.

Definition 3.6.1 (Subbundle of a vector bundle)**.** Let $(E, B, \pi, \mathbb{R}^k)$ a vector bundle of rank k. A *subbundle* of $(E, B, \pi, \mathbb{R}^k)$ is a vector bundle $(S, B, \varsigma, \mathbb{R}^l)$ with $S \subset E$ and $\varsigma = \pi|_S$ of rank $l \leq k$ such that for every $p \in B$ there exists an open set $U \subset B$ with $p \in U$ and local sections $\sigma_1, \ldots, \sigma_l \in \Gamma_{|U}(E)$ such that for all $p' \in U$ the fiber $S_{p'} = \varsigma^{-1}(p')$ is a vector space of dimension l spanned by $\sigma_1(p'), \ldots, \sigma_l(p')$.

This definition conveys the idea that the vector spaces $S_p \subset E_p$ at each point $p \in B$ are not chosen arbitrarily, but "smoothly varying", in the sense that they are spanned by (smooth) local sections. If follows that also the local trivializations of E and S are related to each other. We formulate this statement as follows.

Theorem 3.6.1. *Let $(E, B, \pi, \mathbb{R}^k)$ a vector bundle of rank k. A vector bundle $(S, B, \varsigma, \mathbb{R}^l)$ with $S \subset E$ and $\varsigma = \pi|_S$ is a subbundle of rank l if and only if for each $p \in B$ there exists a local trivialization (U, ϕ) of E with $p \in U$ such that $(U, \tilde{\phi})$ with*

$$\begin{array}{rcccc} \tilde{\phi} & : & \varsigma^{-1}(U) & \to & U \times \mathbb{R}^l \\ & & e & \mapsto & (\phi^1(e), \ldots, \phi^l(e)) \end{array} \tag{3.6.1}$$

is a local trivialization of S.

Proof. We prove the statement in two steps. Let ►. . . ◄ ■

In the literature one sometimes finds a different definition of subbundles as vector, which are mapped into another bundle by means of an injective vector bundle morphism. This definition is closely related, and essentially equivalent up to isomorphism, to the definition we have given, as we shall see below.

Theorem 3.6.2. *Let $(E, B, \pi, \mathbb{R}^k)$ and $(\tilde{S}, B, \tilde{\varsigma}, \mathbb{R}^l)$ be vector bundles over a common base B and $\theta : S \to E$ an injective vector bundle morphism covering the identity on B. Then $(S, B, \varsigma, \mathbb{R}^l)$ with $S = \theta(\tilde{S})$ and $\varsigma = \pi|_S$ is a subbundle of $(E, B, \pi, \mathbb{R}^k)$.*

Proof. For $p \in B$, consider a local trivialization of $\tilde{S}$ on $U \subset B$ with $p \in U$, which is expressed as a local basis $(\epsilon_\mu, \mu = 1, \ldots, l)$. The basis elements $\epsilon_\mu : U \to \tilde{S}$ are local sections, which are linear independent at each $p' \in U$. Since $\theta : \tilde{S} \to E$ is injective, the compositions $\theta \circ \epsilon_\mu : U \to E$ are local sections which are also linear independent at each point. Hence, they span a l-dimensional subspace $S_{p'} \subset E_{p'}$, which agrees with the image $\theta(\tilde{S}_{p'})$, since θ is a vector bundle morphism and thus linear on each fiber. Hence, S is a subbundle. ■

3.7 Metrics on vector bundles

Chapter 4

Operations on vector bundles

4.1 Dual bundle

Given a vector space, there exist several canonical operations which allow the construction of further vector spaces, and we will now see that performing them on each fiber of a vector bundle yields a similar set of operations also on vector bundles. The most common such operation on vector spaces is taking its dual vector space, and it allows us to define the following notion.

Definition 4.1.1 (Dual vector bundle)**.** Let $(E, B, \pi, \mathbb{R}^k)$ be a vector bundle and denote the fiber over $p \in B$ as E_p. Its *dual bundle* is the bundle $(E^*, B, \bar{\pi}, \mathbb{R}^k)$ where the total space E^* is the union

$$E^* = \biguplus_{p \in B} E_p^* \,, \tag{4.1.1}$$

where E_p^* is the dual vector space of the fiber E_p, and the projection $\bar{\pi} : E^* \to B$ assigns to $\alpha \in E_p^*$ the base point p.

We still need to specify an atlas on E^* in order to define its manifold structure. Alternatively, we can provide the local trivializations of $(E^*, B, \bar{\pi}, \mathbb{R}^k)$, which then yield an induced atlas on E^* from an atlas on B. Here we will do the latter, and construct the local trivializations from those on $(E, B, \pi, \mathbb{R}^k)$. Recall that these are pairs (U, ϕ), where $U \subset B$ and $\phi : \pi^{-1}(U) \to U \times \mathbb{R}^k$ restricts to a linear isomorphism on every fiber E_p with $p \in U$. From this we can define a map $\tilde{\phi} : U \times \mathbb{R}^k \to \bar{\pi}^{-1}(U)$, which assigns to $(p, x) \in U \times \mathbb{R}^k$ the linear function

$$\begin{array}{rccc} \tilde{\phi}(p, x) & : & E_p & \to & \mathbb{R} \\ & & e & \mapsto & x \cdot (\mathrm{pr}_2 \circ \phi)(e) \end{array} . \tag{4.1.2}$$

This definition needs a few remarks. First, recall that $\mathrm{pr}_2 \circ \phi : E_p \to \mathbb{R}^k$ is a linear isomorphism. The Euclidean space $\mathbb{R}^k$ is equipped with an inner product $(x, y) \mapsto x \cdot y$, which is linear in each component and non-degenerate. Hence, for fixed $x \in \mathbb{R}^k$, the function $y \mapsto x \cdot y$ is linear, and therefore defines an element of $(\mathbb{R}^k)^* \cong \mathbb{R}^k$. Together with the map $\mathrm{pr}_2 \circ \phi$ we thus see that $\tilde{\phi}(p, x) \in E_p^*$. Using the Riesz representation theorem, one can show that this establishes a linear isomorphism

$$\begin{array}{rccc} \tilde{\phi}(p, \bullet) & : & \mathbb{R}^k & \to & E_p^* \\ & & x & \mapsto & \tilde{\phi}(p, x) \end{array} \tag{4.1.3}$$

of vector spaces. One can further show that $\tilde{\phi}$ is a diffeomorphism from $U \times \mathbb{R}^k$ to $\bar{\pi}^{-1}(U) \subset E^*$. Its inverse therefore yields a local trivialization $(U, \tilde{\phi}^{-1})$ of $(E^*, B, \bar{\pi}, \mathbb{R}^k)$. One easily checks that these are compatible with the definition of a vector bundle.

If E is equipped with adapted coordinates (x^i, y^μ) on a chart (W, ω) with $W = \pi^{-1}(U)$, which are induced by coordinates (x^i) on a chart (U, ψ) on B and a local trivialization, then the construction above yields coordinates (x^i, z_μ) on $W^* = \bar{\pi}^{-1}(U) \subset E^*$ such that (y^μ) and (z_μ) are dual vector space coordinates on each pair of fibers E_p and E_p^* (which is the reason for using lower indices on z_μ).

An equivalent possibility to arrive at the coordinates (x^i, z_μ) makes use of the induced coordinate basis on E as given by definition 3.2.1. Denoting the coordinate basis of E_p by (ϵ_μ), there exists a corresponding dual basis $(\bar{\epsilon}^\mu)$ on the dual vector space E_p^*. This is the coordinate basis induced by the local trivialization $(U, \tilde{\phi}^{-1})$ constructed above.

Recall that for any (finite-dimensional) real vector space V and its dual V^* there exists a *canonical pairing*, i.e., a function

$$\begin{array}{rcccc} \langle \bullet, \bullet \rangle & : & V^* \times V & \to & \mathbb{R} \\ & & (\alpha, e) & \mapsto & \langle \alpha, e \rangle = \alpha(e) \end{array}, \tag{4.1.4}$$

which is bilinear and non-degenerate. This of course also applies to the fibers of a vector bundle and its dual. Further exploiting their geometric relation, we can extend this notion to the whole bundles, and show the following.

Theorem 4.1.1. *Let $(E, B, \pi, \mathbb{R}^k)$ be a vector bundle and $(E^*, B, \bar{\pi}, \mathbb{R}^k)$ its dual bundle. Then the function*

$$\begin{array}{rcccc} \langle \bullet, \bullet \rangle & : & E^* \times_B E & \to & \mathbb{R} \\ & & (\alpha, e) & \mapsto & \langle \alpha, e \rangle = \alpha(e) \end{array} \tag{4.1.5}$$

is a smooth map.

Proof. For $p \in B$, let (U, ψ) be a chart of B with $p \in U$, as well as (U, ϕ) a local trivialization of E and $(U, \bar{\phi})$ the corresponding local trivialization of E^*. These define a local trivialization (U, Φ) of $E^* \times_B E$ such that

$$\Phi(\alpha, e) = (\pi(e), \mathrm{pr}_2(\bar{\phi}(\alpha)), \mathrm{pr}_2(\phi(e))) \tag{4.1.6}$$

For $e \in \pi^{-1}(U)$ and $\alpha \in \bar{\pi}^{-1}(U)$. The scalar product $\cdot : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}$ is smooth, and so is

$$(\alpha, e) \mapsto \mathrm{pr}_2(\bar{\phi}(\alpha)) \cdot \mathrm{pr}_2(\phi(e)) = \alpha(e) = \langle \alpha, e \rangle \,. \tag{4.1.7}$$

■

It is now clear that if we pre-compose this map with a section of $E^* \times_B E$, which is equivalent to a pair of sections of E^* and E, we obtain a function from B to $\mathbb{R}$. This leads us to the following definition.

Definition 4.1.2 (Canonical pairing). Let $(E, B, \pi, \mathbb{R}^k)$ be a vector bundle and $(E^*, B, \bar{\pi}, \mathbb{R}^k)$ its dual bundle. The *canonical pairing* $\langle \bullet, \bullet \rangle : \Gamma(E^*) \times \Gamma(E) \to C^\infty(B, \mathbb{R})$ between sections of E and E^* is the function that assigns to $f \in \Gamma(E)$ and $\upsilon \in \Gamma(E^*)$ the function

$$\begin{array}{rcccc} \langle \upsilon, f \rangle & : & B & \to & \mathbb{R} \\ & & p & \mapsto & \langle \upsilon(p), f(p) \rangle \end{array}. \tag{4.1.8}$$

It is clear that

$$\langle \upsilon, f \rangle = \langle \bullet, \bullet \rangle \circ (\upsilon, f) \tag{4.1.9}$$

is a composition of smooth maps, and hence also a smooth map.

With the help of the canonical pairing, it is now also possible to define the following notion.

Definition 4.1.3 (Dual vector bundle morphism). Let $(E_1, B_1, \pi_1, \mathbb{R}^{k_1})$ and $(E_2, B_2, \pi_2, \mathbb{R}^{k_2})$ be vector bundles and $\theta : E_1 \to E_2$ a vector bundle homomorphism covering a diffeomorphism $\vartheta : B_1 \to B_2$. The *dual vector bundle morphism* is the vector bundle morphism $\bar{\theta} : E_2^* \to E_1^*$ covering $\bar{\vartheta} = \vartheta^{-1} : B_2 \to B_1$ which is defined such that

$$\langle\bar{\theta}(\alpha), e\rangle = \langle\alpha, \theta(e)\rangle \tag{4.1.10}$$

for all $e \in E_1$ and $\alpha \in E_2^*$ with $\vartheta(\pi_1(e)) = \bar{\pi}_2(\alpha)$.

4.2 Direct sum

We continue with another important construction:

Definition 4.2.1 (Direct sum bundle). Let $(E_1, B, \pi_1, \mathbb{R}^{k_1})$ and $(E_2, B, \pi_2, \mathbb{R}^{k_2})$ be vector bundles and denote the fibers over $p \in B$ as E_{1p} and E_{2p}. Their *direct sum* (or *Whitney sum*) is the bundle $(E_1 \oplus E_2, B, \pi_1 \oplus \pi_2, \mathbb{R}^{k_1+k_2})$ where the total space $E_1 \oplus E_2$ is the union

$$E_1 \oplus E_2 = \biguplus_{p\in B} E_{1p} \oplus E_{2p}\,, \tag{4.2.1}$$

where $(E_1 \oplus E_2)_p = E_{1p} \oplus E_{2p}$ is the direct sum of the fiber vector spaces E_{1p} and E_{2p}, and the projection $\pi_1 \oplus \pi_2 : E_1 \oplus E_2 \to B$ assigns to $e \in (E_1 \oplus E_2)_p$ the base point p.

Note that the elements of each fiber $(E_1 \oplus E_2)_p$ are given by pairs (v, v') of elements $v \in E_{1p}$ and $v' \in E_{2p}$, where the vector space operations act on both components:

$$\lambda(v, v') + \mu(w, w') = (\lambda v + \mu w, \lambda v' + \mu w')\,. \tag{4.2.2}$$

Given a chart (U, ψ) on B and local trivializations (U, ϕ_1) of $(E_1, B, \pi_1, \mathbb{R}^{k_1})$ and (U, ϕ_2) of $(E_2, B, \pi_2, \mathbb{R}^{k_2})$, so that the latter induce charts (W_1, ω_1) and (W_2, ω_2), we can construct an induced chart (W, ω) and local trivialization (U, ϕ) of $(E_1 \oplus E_2, B, \pi_1 \oplus \pi_2, \mathbb{R}^{k_1+k_2})$ as follows. Let $W = (\pi_1 \oplus \pi_2)^{-1}(U) \subset E_1 \oplus E_2$ and define

$$\phi(v, v') = ((\pi_1 \oplus \pi_2)(v, v'), (\mathrm{pr}_2 \circ\phi_1)(v), (\mathrm{pr}_2 \circ\phi_2)(v'))\,, \tag{4.2.3}$$

as well as

$$\omega(v, v') = (\psi((\pi_1 \oplus \pi_2)(v, v')), (\mathrm{pr}_2 \circ\phi_1)(v), (\mathrm{pr}_2 \circ\phi_2)(v'))\,. \tag{4.2.4}$$

One easily checks that all required properties are satisfied.

Given coordinates (x^i) defined by chart (U, ψ) on B and corresponding introduced coordinates (x^i, y^μ) on $(E_1, B, \pi_1, \mathbb{R}^{k_1})$ and $(x^i, z^{\bar{\mu}})$ on $(E_2, B, \pi_2, \mathbb{R}^{k_2})$, the corresponding induced coordinates on $(E_1 \oplus E_2, B, \pi_1 \oplus \pi_2, \mathbb{R}^{k_1+k_2})$ simply take the form $(x^i, y^\mu, z^{\bar{\mu}})$.

Again we can make use of the induced coordinate bases. Denoting the coordinate bases of E_{1p} and E_{2p} by $(\epsilon_\mu, \mu = 1, \ldots, k_1)$ and $(\epsilon'_{\bar{\mu}}, \bar{\mu} = 1, \ldots, k_2)$, respectively, the induced basis of $E_{1p} \oplus E_{2p}$ is simply their union

$$(\epsilon_\mu, \epsilon'_{\bar{\mu}}) = \{\epsilon_1, \ldots, \epsilon_{k_1}, \epsilon'_1, \ldots, \epsilon'_{k_2}\} \tag{4.2.5}$$

with k_1+k_2 elements. Using this basis an element of $E_{1p}\oplus E_{2p}$ is uniquely written as $y^\mu\epsilon_\mu + z^{\bar{\mu}}\epsilon'_{\bar{\mu}}$.

We finally remark that there is a close relationship, but also a subtle difference between the fibered product $E_1 \times_B E_2$ and the direct sum $E_1 \oplus E_2$. We will first study the former, which we can formulate as follows.

Theorem 4.2.1. *Let $(E_1, B, \pi_1, \mathbb{R}^{k_1})$ and $(E_2, B, \pi_2, \mathbb{R}^{k_2})$ be vector bundles over a common base B. Then the map*

$$\begin{array}{ccccc} \oplus & : & E_1 \times_B E_2 & \to & E_1 \oplus E_2 \\ & & (v, v') & \mapsto & (v, v') \end{array} \tag{4.2.6}$$

is a fiber bundle isomorphism covering the identity.

Proof. In both cases the total spaces are constituted by pairs $(v, v') \in E_1 \times E_2$ with $\pi_1(v) = \pi_2(v')$. Hence, both spaces contain the same elements. Also the atlases and local trivializations agree, and so $E_1 \times_B E_2$ and $E_1 \oplus E_2$ are isomorphic fiber bundles, and the isomorphism is as given above. ■

The statement above implies that the total spaces of $E_1 \times_B E_2$ and $E_1 \oplus E_2$ are "the same", i.e., diffeomorphic manifolds, and that also the projections agree. The difference lies in the additional algebraic structure. The fibered product is an operation on general fiber bundles, and does not assign any algebraic structure to the resulting bundle; hence, $E_1 \times_B E_2$ is a fiber bundle, but not a vector bundle, even if E_1 and E_2 are vector bundles, because it does not imply any vector space operations on the pairs (v, v'). In contrast, $E_1 \oplus E_2$ is a vector bundle, since its definition equips every fiber with a particular vector space structure.

Theorem 4.2.2. *Let $(E_1, B, \pi_1, \mathbb{R}^{k_1})$ and $(E_2, B, \pi_2, \mathbb{R}^{k_2})$ be vector bundles over a common base B. Then the inclusion maps*

$$\begin{array}{ccccc} \iota_1 & : & E_1 & \to & E_1 \oplus E_2 \\ & & v & \mapsto & (v, 0) \end{array} \tag{4.2.7}$$

and

$$\begin{array}{ccccc} \iota_2 & : & E_2 & \to & E_1 \oplus E_2 \\ & & v' & \mapsto & (0, v') \end{array} \tag{4.2.8}$$

are vector bundle homomorphisms covering the identity.

Proof. ►...◄ ■

Theorem 4.2.3. *Let $(E_1, B, \pi_1, \mathbb{R}^{k_1})$ and $(E_2, B, \pi_2, \mathbb{R}^{k_2})$ be vector bundles over a common base B. Then the projection maps*

$$\begin{array}{ccccc} \mathrm{pr}_1 & : & E_1 \oplus E_2 & \to & E_1 \\ & & (v, v') & \mapsto & v \end{array} \tag{4.2.9}$$

and

$$\begin{array}{ccccc} \mathrm{pr}_2 & : & E_1 \oplus E_2 & \to & E_2 \\ & & (v, v') & \mapsto & v' \end{array} \tag{4.2.10}$$

are vector bundle homomorphisms covering the identity.

Proof. ►...◄ ■

4.3 Tensor product

Another possible construction is the following:

Definition 4.3.1 (Tensor product bundle)**.** Let $(E_1, B, \pi_1, \mathbb{R}^{k_1})$ and $(E_2, B, \pi_2, \mathbb{R}^{k_2})$ be vector bundles and denote the fibers over $p \in B$ as E_{1p} and E_{2p}. Their *tensor product* is the bundle $(E_1 \otimes E_2, B, \pi_1 \otimes \pi_2, \mathbb{R}^{k_1 k_2})$ where the total space $E_1 \otimes E_2$ is the union

$$E_1 \otimes E_2 = \biguplus_{p \in B} E_{1p} \otimes E_{2p} , \tag{4.3.1}$$

where $(E_1 \otimes E_2)_p = E_{1p} \otimes E_{2p}$ is the tensor product of the fiber vector spaces E_{1p} and E_{2p}, and the projection $\pi_1 \otimes \pi_2 : E_1 \otimes E_2 \to B$ assigns to $e \in (E_1 \otimes E_2)_p$ the base point p.

The construction of induced charts proceeds essentially in the same way as for the direct sum discussed in the previous section, and so we will not repeat it here. Instead, we will construct the coordinate bases. Given bases $(\epsilon_\mu, \mu = 1, \ldots, k_1)$ of E_{1p} and $(\epsilon'_{\bar{\mu}}, \bar{\mu} = 1, \ldots, k_2)$ of E_{2p}, a basis of $(E_1 \otimes E_2)_p$ is given by $(\epsilon_\mu \otimes \epsilon'_{\bar{\mu}})$, and thus has $k_1 k_2$ elements. Coordinates on $E_1 \otimes E_2$ therefore are of the form $(x^i, w^{\mu\bar{\mu}})$, where the basis expansion takes the form

$$w^{\mu\bar{\mu}} \epsilon_\mu \otimes \epsilon'_{\bar{\mu}} , \tag{4.3.2}$$

and (x^i) are coordinates on the base manifold B. It is thus straightforward to conclude the following.

Theorem 4.3.1. *The tensor product of two vector bundles of rank k_1 and k_2 is a vector bundle of rank $k_1 k_2$.*

Proof. ▶ . . . ◀ ■

Combining this notion with that of the dual bundle, one finds the following.

Theorem 4.3.2. *The dual $(E_1 \otimes E_2)^*$ of a tensor product bundle $E_1 \otimes E_2$ is canonically isomorphic to $E_1^* \otimes E_2^*$.*

Proof. ▶ . . . ◀ ■

Theorem 4.3.3. *Let $(E_1, B, \pi_1, \mathbb{R}^{k_1})$ and $(E_2, B, \pi_2, \mathbb{R}^{k_2})$ be vector bundles over a common base B. Then the map*

$$\begin{array}{cccc} \otimes & : & E_1 \times_B E_2 & \to & E_1 \otimes E_2 \\ & & (v, v') & \mapsto & v \otimes v' \end{array} \tag{4.3.3}$$

is a fiber bundle morphism covering the identity.

Proof. ▶ . . . ◀ ■

Often one considers the tensor product of multiple copies of the same bundle. Since this is a rather common construction, we give it its own name.

Definition 4.3.2 (Tensor power bundle)**.** Let $(E, B, \pi, \mathbb{R}^n)$ be a vector bundle of rank n. Its k'th *tensor power* is the bundle $(\bigotimes^k E, B, \bigotimes^k \pi, \mathbb{R}^{n^k})$ with

$$\bigotimes^k E = \underbrace{E \otimes \ldots \otimes E}_{k \text{ times}} . \tag{4.3.4}$$

Finally, it is often useful to also include powers of the dual vector bundle E^* discussed in section 4.1. Also here we introduce a convenient notation.

Definition 4.3.3 (Tensor bundle)**.** Let $(E, B, \pi, \mathbb{R}^n)$ be a vector bundle of rank n. Its *tensor bundle* of type (r, s) for $r, s \in \mathbb{N}$ is the tensor product bundle

$$E^r_s = \bigotimes^r E \otimes \bigotimes^s E^* \,. \tag{4.3.5}$$

Given coordinates (x^i, y^μ) on E corresponding to a coordinate basis (ϵ_μ), the corresponding coordinates on E^r_s can therefore be written in the form $(x^i, z^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_s})$ and correspond to a coordinate basis $\epsilon_{\mu_1} \otimes \ldots \otimes \epsilon_{\mu_r} \otimes \bar{\epsilon}^{\nu_1} \otimes \ldots \otimes \bar{\epsilon}^{\nu_s}$. From these basis expressions, it is also straightforward to arrive at the following statement.

Theorem 4.3.4. *The dual $(E^r_s)^*$ of a tensor bundle E^r_s is canonically isomorphic to E^s_r.*

Proof. This follows immediately from repeatedly applying theorem 4.3.2. ∎

We have already encountered a few special cases of tensor bundles. It follows directly from definition 4.3.3 that $E^1_0 \cong E$ and $E^0_1 \cong E^*$. Further, setting $r = s = 0$, we find the empty tensor product; this is simply the trivial line bundle $E^0_0 \cong B \times \mathbb{R}$. Finally, we will encounter another special case for $r = s = 1$ in section 4.6.

Sections of the tensor bundle E^r_s play an important role, and carry various operations. We will therefore discuss them in detail in chapter 5.

4.4 Exterior power

From the construction of the tensor power bundle shown in the one can easily derive similar notions. One of the most important, which will be relevant for the construction of differential forms in chapter 9, is the following.

Definition 4.4.1 (Exterior power bundle)**.** Let $(E, B, \pi, \mathbb{R}^n)$ be a vector bundle of rank n and denote the fiber over $p \in B$ as E_p. Its k'th *exterior power* is the bundle $(\Lambda^k E, B, \Lambda^k\pi, \mathbb{R}^{\binom{n}{k}})$ where the total space $\Lambda^k E$ is the union

$$\Lambda^k E = \bigcup_{p \in B} \Lambda^k E_p \,, \tag{4.4.1}$$

where $\Lambda^k E_p$ is the k'th exterior power of the fiber E_p, and the projection $\Lambda^k\pi : \Lambda^k E \to B$ assigns to $\alpha \in \Lambda^k E_p$ the base point p.

Recall from linear algebra that the exterior power $\Lambda^k V$ of a vector space V of dimension n with basis $(\epsilon_i, i = 1, \ldots, n)$ is spanned by the vectors

$$\epsilon_{i_1} \wedge \ldots \wedge \epsilon_{i_k} = \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) \epsilon_{i_{\sigma(1)}} \otimes \ldots \otimes \epsilon_{i_{\sigma(k)}} \,, \tag{4.4.2}$$

where the sum is taken over all permutations σ (elements of the symmetric group S_k permuting k objects) and $\operatorname{sgn}(\sigma)$ is the signature of σ. It follows that there are $\binom{n}{k}$ linearly independent vectors of this type, which constitute a basis

$$\{\epsilon_{i_1} \wedge \ldots \wedge \epsilon_{i_k}, 1 \leq i_1 < \ldots < i_k \leq n\} \tag{4.4.3}$$

of $\Lambda^k V$. If $V = E_p$ is a fiber of a vector bundle, we may choose for this purpose an induced coordinate basis. Following essentially the same construction as in the case of tensor product bundles and tensor powers, one constructs an induced basis of $\Lambda^k E_p$. It is thus clear that $(\Lambda^k E, B, \Lambda^k \pi, \mathbb{R}^{\binom{n}{k}})$ is indeed a vector bundle of rank

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}\,. \tag{4.4.4}$$

Its relation to the tensor power bundle is established by the following statement:

Theorem 4.4.1. *The exterior power bundle $(\Lambda^k E, B, \Lambda^k \pi, \mathbb{R}^{\binom{n}{k}})$ is a subbundle of the tensor power bundle $(\bigotimes^k E, B, \bigotimes^k \pi, \mathbb{R}^{n^k})$, and there exists a surjective vector bundle homomorphism $\theta^- : \bigotimes^k E \to \Lambda^k E$, which can be written in the coordinate basis as*

$$(x^i, y^{\mu_1 \cdots \mu_k} \epsilon_{\mu_1} \otimes \ldots \otimes \epsilon_{\mu_k}) \mapsto (x^i, y^{\mu_1 \cdots \mu_k} \epsilon_{\mu_1} \wedge \ldots \wedge \epsilon_{\mu_k})\,. \tag{4.4.5}$$

Proof. ►. . . ◄ ■

Using the definition (4.4.2) of the basis, we can also write the fiber coordinate of the image in the form

$$\begin{aligned}\frac{1}{k!} y^{\mu_1 \cdots \mu_k} \epsilon_{\mu_1} \wedge \ldots \wedge \epsilon_{\mu_k} &= \frac{1}{k!} y^{\mu_1 \cdots \mu_k} \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) \epsilon_{\mu_{\sigma(1)}} \otimes \ldots \otimes \epsilon_{\mu_{\sigma(k)}} \\ &= \frac{1}{k!} \left(\sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) y^{\mu_{\sigma(1)} \cdots \mu_{\sigma(k)}} \right) \epsilon_{\mu_1} \otimes \ldots \otimes \epsilon_{\mu_k} \\ &= y^{[\mu_1 \cdots \mu_k]} \epsilon_{\mu_1} \otimes \ldots \otimes \epsilon_{\mu_k}\,,\end{aligned} \tag{4.4.6}$$

where we introduced the notation

$$y^{[\mu_1 \cdots \mu_k]} = \frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) y^{\mu_{\sigma(1)} \cdots \mu_{\sigma(k)}}\,. \tag{4.4.7}$$

4.5 Symmetric power

In analogy to the exterior power of vector bundles, one may also consider their symmetric power, which is defined as follows.

Definition 4.5.1 (Symmetric power bundle). Let $(E, B, \pi, \mathbb{R}^n)$ be a vector bundle of rank n and denote the fiber over $p \in B$ as E_p. Its k'th *symmetric power* is the bundle $(\operatorname{Sym}^k E, B, \operatorname{Sym}^k \pi, \mathbb{R}^{\binom{n+k-1}{k}})$ where the total space $\operatorname{Sym}^k E$ is the union

$$\operatorname{Sym}^k E = \bigcup_{p \in B} \operatorname{Sym}^k E_p\,, \tag{4.5.1}$$

where $\operatorname{Sym}^k E_p$ is the k'th symmetric power of the fiber E_p, and the projection $\operatorname{Sym}^k \pi : \operatorname{Sym}^k E \to B$ assigns to $\alpha \in \operatorname{Sym}^k E_p$ the base point p.

In this case we may use the fact that the symmetric power $\mathrm{Sym}^k V$ of a vector space V of dimension n with basis $(\epsilon_i, i = 1, \ldots, n)$ is spanned by the vectors

$$\epsilon_{i_1} \odot \ldots \odot \epsilon_{i_k} = \sum_{\sigma \in S_k} \epsilon_{i_{\sigma(1)}} \otimes \ldots \otimes \epsilon_{i_{\sigma(k)}}\,, \tag{4.5.2}$$

where also in this case the sum is taken over all permutations σ. One can show that there are $\binom{n+k-1}{k}$ linearly independent vectors of this type, which constitute a basis

$$\{\epsilon_{i_1} \odot \ldots \odot \epsilon_{i_k}, 1 \leq i_1 \leq \ldots \leq i_k \leq n\} \tag{4.5.3}$$

of $\mathrm{Sym}^k V$. Applying this to the fibers of a vector bundle and making use of the induced coordinate basis, in the same way as done for the exterior power, one finds that $(\mathrm{Sym}^k E, B, \mathrm{Sym}^k \pi, \mathbb{R}^{\binom{n+k-1}{k}})$ is indeed a vector bundle of rank

$$\binom{n+k-1}{k} = \frac{(n+k-1)!}{k!(n-1)!}\,. \tag{4.5.4}$$

It is related to the tensor power bundle as follows:

Theorem 4.5.1. *The symmetric power bundle* $(\mathrm{Sym}^k E, B, \mathrm{Sym}^k \pi, \mathbb{R}^{\binom{n+k-1}{k}})$ *is a subbundle of the tensor power bundle* $(\bigotimes^k E, B, \bigotimes^k \pi, \mathbb{R}^{n^k})$*, and there exists a surjective vector bundle homomorphism* $\theta^+ : \bigotimes^k E \to \mathrm{Sym}^k E$*, which can be written in the coordinate basis as*

$$(x^i, y^{\mu_1 \cdots \mu_k} \epsilon_{\mu_1} \otimes \ldots \otimes \epsilon_{\mu_k}) \mapsto (x^i, y^{\mu_1 \cdots \mu_k} \epsilon_{\mu_1} \odot \ldots \odot \epsilon_{\mu_k})\,. \tag{4.5.5}$$

Proof. ▶ . . . ◀ ■

Using the definition (4.5.2) of the basis, we can also write the fiber coordinate of the image in the convenient form

$$\begin{aligned} \frac{1}{k!} y^{\mu_1 \cdots \mu_k} \epsilon_{\mu_1} \odot \ldots \odot \epsilon_{\mu_k} &= \frac{1}{k!} y^{\mu_1 \cdots \mu_k} \sum_{\sigma \in S_k} \epsilon_{\mu_{\sigma(1)}} \otimes \ldots \otimes \epsilon_{\mu_{\sigma(k)}} \\ &= \frac{1}{k!} \left(\sum_{\sigma \in S_k} y^{\mu_{\sigma(1)} \cdots \mu_{\sigma(k)}} \right) \epsilon_{\mu_1} \otimes \ldots \otimes \epsilon_{\mu_k} \\ &= y^{(\mu_1 \cdots \mu_k)} \epsilon_{\mu_1} \otimes \ldots \otimes \epsilon_{\mu_k}\,, \end{aligned} \tag{4.5.6}$$

where we introduced the notation

$$y^{(\mu_1 \cdots \mu_k)} = \frac{1}{k!} \sum_{\sigma \in S_k} y^{\mu_{\sigma(1)} \cdots \mu_{\sigma(k)}}\,. \tag{4.5.7}$$

4.6 Homomorphism and endomorphism bundles

Another possible way to obtain a vector space from two given vector spaces V_1, V_2 is by considering all linear functions from V_1 to V_2. Clearly, this is a vector space, denoted $\mathrm{Hom}(V_1, V_2)$, since any linear combination of linear functions is again a linear function, and the zero element of this vector space is the function which sends all elements of V_1 to the zero element of V_2. If these vector spaces are fibers of vector bundles, it appears natural to define a related notion also for the whole bundles. We thus define:

Definition 4.6.1 (Homomorphism bundle)**.** Let $(E_1, B, \pi_1, \mathbb{R}^{k_1})$ and $(E_2, B, \pi_2, \mathbb{R}^{k_2})$ be vector bundles and denote the fibers over $p \in B$ as E_{1p} and E_{2p}. Their *homomorphism bundle* is the bundle $(\mathrm{Hom}(E_1, E_2), B, \mathrm{Hom}(\pi_1, \pi_2), \mathbb{R}^{k_1 k_2})$ where the total space $\mathrm{Hom}(E_1, E_2)$ is the union

$$\mathrm{Hom}(E_1, E_2) = \bigcup_{p \in B} \mathrm{Hom}(E_{1p}, E_{2p}), \tag{4.6.1}$$

where $\mathrm{Hom}(E_1, E_2)_p = \mathrm{Hom}(E_{1p}, E_{2p})$ is the space of linear functions from E_{1p} to E_{2p}, and the projection $\mathrm{Hom}(\pi_1, \pi_2) : \mathrm{Hom}(E_1, E_2) \to B$ assigns to $e \in \mathrm{Hom}(E_1, E_2)_p$ the base point p.

Of course one still has to show that this is indeed a vector bundle, with the total space being a manifold and all requirements being satisfied. However, instead of doing this, we will make use of the following statement, which relates that we essentially already encountered and discussed this bundle before.

Theorem 4.6.1. *Let $(E_1, B, \pi_1, \mathbb{R}^{k_1})$ and $(E_2, B, \pi_2, \mathbb{R}^{k_2})$ be vector bundles. There exists a canonical vector bundle isomorphism from the homomorphism bundle $(\mathrm{Hom}(E_1, E_2), B, \mathrm{Hom}(\pi_1, \pi_2), \mathbb{R}^{k_1 k_2})$ to the tensor product bundle $(E_2 \otimes E_1^*, B, \pi_2 \otimes \bar{\pi}_1, \mathbb{R}^{k_1 k_2})$.*

Proof. ▶ . . . ◀ ■

We will not give a full proof here, and a sketch of the proof should be sufficient. The key ingredient is that there is a canonical vector space isomorphism from $\mathrm{Hom}(V_1, V_2)$ to $V_2 \otimes V_1^*$ for any pair of vector spaces. This isomorphism can be applied to each fiber of the bundles constructed above. One then easily checks that this indeed yields a vector bundle isomorphism. Hence, one may canonically identify these two bundles. Note that in the literature one also finds the opposite order $E_1^* \otimes E_2$ of the tensor product; also this is vector bundle isomorphic to $\mathrm{Hom}(E_1, E_2)$. However, he we prefer the former, since it will allow for a more intuitive component notation, as we shall see below.

Sections of homomorphism bundles have a few interesting properties. Note that such a section $f : B \to \mathrm{Hom}(E_1, E_2)$ assigns to each point $p \in B$ a linear function $f(p) : E_{1p} \to E_{2p}$. Given a section $v : B \to E_1$ of E_1 one may then define $fv : B \to E_2$ such that

$$\begin{array}{rcccc} fv & : & B & \to & E_2 \\ & & p & \mapsto & f(p)(v(p)) \end{array} . \tag{4.6.2}$$

It is not difficult to prove that fv is a section of E_2. Hence, sections of the homomorphism bundle allow to relate sections of different vector bundles. This can be put in formal terms in the following statement:

Theorem 4.6.2. *There is a one-to-one correspondence between vector bundle homomorphisms from $(E_1, B, \pi_1, \mathbb{R}^{k_1})$ to $(E_2, B, \pi_2, \mathbb{R}^{k_2})$ covering the identity id_B on B and sections of the homomorphism bundle $(\mathrm{Hom}(E_1, E_2), B, \mathrm{Hom}(\pi_1, \pi_2), \mathbb{R}^{k_1 k_2})$.*

Proof. From definition 3.4.1 follows that a vector bundle homomorphism covering the identity on a common base manifold B is a map $\theta : E_1 \to E_2$ that restricts to a linear map $\theta_p = \theta|_{E_{1p}} : E_{1p} \to E_{2p}$ on the fibers over every point $p \in B$. Hence, $\theta_p \in \mathrm{Hom}(E_{1p}, E_{2p})$, and so

$$\begin{array}{rcccc} \vartheta & : & B & \to & \mathrm{Hom}(E_1, E_2) \\ & & p & \mapsto & \vartheta(p) = \theta_p \end{array} \tag{4.6.3}$$

defines a section of the homomorphism bundle. Conversely, every such section $\vartheta \in \Gamma(\mathrm{Hom}(E_1, E_2))$ defines a vector bundle homomorphism $\theta : E_1 \to E_2$ covering the identity by defining $\theta(e) = \vartheta(\pi_1(e))e$. These two prescriptions are obviously inverse to each other.

To check the smoothness of the respective maps, consider local trivializations (U, ϕ_1) of E_1 and (U, ϕ_2) of E_2 on a common domain $U \subset B$. ►. . . ◄ ■

If the two bundles are the same, $E_1 = E_2$ and $\pi_1 = \pi_2$, then one has the following special case:

Definition 4.6.2 (Endomorphism bundle)**.** Let $(E, B, \pi, \mathbb{R}^k)$ a vector bundle. Its *endomorphism bundle* is the bundle $\mathrm{End}(E) = \mathrm{Hom}(E, E)$.

We finally also discuss coordinate bases of the homomorphism bundle. For this purpose, let $(\epsilon_\mu, \mu = 1, \ldots, k_1)$ and $(\epsilon'_{\bar{\mu}}, \bar{\mu} = 1, \ldots, k_2)$ denote the coordinate basis of the fibers E_{1p} and E_{2p} over a point $p \in B$. Then we can construct the basis $(\bar{\epsilon}^\mu)$ of the fiber E^*_{1p} of the dual bundle, as shown in section 4.1. Using the construction from section 4.3, we further obtain the basis $(\epsilon'_{\bar{\mu}} \otimes \bar{\epsilon}^\mu)$ of the fibers of the tensor product bundle, which, following theorem 4.6.1, is canonically isomorphic to the homomorphism bundle. Hence, we may use the same basis and the canonical isomorphism as a basis on the homomorphism bundle $\mathrm{Hom}(E_1, E_2)$.

Denoting the coordinates on B by (x^i), one may denote the coordinates on the homomorphism bundle by $(x^i, w^{\bar{\mu}}{}_\mu)$. Using the coordinate basis, an element of the fiber over the point with coordinates (x^i) is then expressed as $w^{\bar{\mu}}{}_\mu \epsilon'_{\bar{\mu}} \otimes \bar{\epsilon}^\mu$.

The coordinate expressions above are particularly useful for describing operations on the sections of the respective bundles, such as the aforementioned application (4.6.2) of a section of the homomorphism bundle to a section of the first constituting vector bundle. Writing these sections in coordinates as $(x^i) \mapsto (x^i, f^{\bar{\mu}}{}_\mu)$ and $(x^i) \mapsto (x^i, v^\mu)$, respectively, we can write fv as $(x^i) \mapsto (x^i, f^{\bar{\mu}}{}_\mu v^\mu)$, where we used the Einstein summation convention, i.e., we take the sum over the index μ. This follows from the action of the homomorphism,

$$
\begin{aligned}
fv &= \left(f^{\bar{\mu}}{}_\mu \epsilon'_{\bar{\mu}} \otimes \bar{\epsilon}^\mu\right) \cdot (v^\nu \epsilon_\nu) \\
&= f^{\bar{\mu}}{}_\mu v^\nu \epsilon'_{\bar{\mu}} \otimes (\bar{\epsilon}^\mu \cdot \epsilon_\nu) \\
&= f^{\bar{\mu}}{}_\mu v^\nu \epsilon'_{\bar{\mu}} \delta^\mu_\nu \\
&= f^{\bar{\mu}}{}_\mu v^\mu \epsilon'_{\bar{\mu}} \,.
\end{aligned}
\tag{4.6.4}
$$

This notation is also the reason why we choose to write the homomorphism bundle as $E_2 \otimes E^*_1$ and not in the opposite order, since it is reminiscent of multiplying a matrix f with a vector v. More of these constructions are discussed in chapter 5.

4.7 Quotient bundles

Chapter 5

Tensors

5.1 Tensor fields

Among the different operations on vector bundles we studied in chapter 4, the tensor power bundle 4.3.2, and more generally the tensor bundle 4.3.3, introduced in section 4.3 play an important role. We have seen that the symmetric and antisymmetric powers introduced in sections 4.4 and 4.5 constitute particular subbundles of tensor bundles, while the endomorphism bundle 4.6.2 is simply identified with E^1_1. We now study the sections of these bundles in detail. We first introduce a suitable name.

Definition 5.1.1 (Tensor field)**.** A *tensor field* of type (r, s) in a vector bundle $\pi : E \to B$ is a section of the tensor bundle E^r_s. The set of all tensor fields of type (r, s) in E is denoted $\Gamma(E^r_s)$.

In order to work with tensor fields, it is most convenient to introduce a local basis of the tensor bundle E^r_s. This is achieved most easily by using a basis (ϵ_a) of E. It follows that this basis induces a dual basis and $(\bar{\epsilon}^a)$ of E^*, which can then further be used to construct a basis on E^r_s. This basis is then given by the elements

$$\epsilon_{a_1} \otimes \ldots \otimes \epsilon_{a_r} \otimes \bar{\epsilon}^{b_1} \otimes \ldots \otimes \bar{\epsilon}^{b_s} , \tag{5.1.1}$$

where each index runs from 1 to $\dim B$, so that the basis has $(\dim B)^{r+s}$ elements. Any element $u \in E^r_s$ can be expanded using this basis in the form

$$u = u^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} \epsilon_{a_1} \otimes \ldots \otimes \epsilon_{a_r} \otimes \bar{\epsilon}^{b_1} \otimes \ldots \otimes \bar{\epsilon}^{b_s} , \tag{5.1.2}$$

with r upper and s lower indices. For a tensor field $T \in \Gamma(E^r_s)$, we have such an element for every $p \in B$, and so we can write it analogously locally as

$$T = T^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} \epsilon_{a_1} \otimes \ldots \otimes \epsilon_{a_r} \otimes \bar{\epsilon}^{b_1} \otimes \ldots \otimes \bar{\epsilon}^{b_s} , \tag{5.1.3}$$

where the components $T^{a_1 \cdots a_r}{}_{b_1 \cdots b_s}$ are smooth functions on a neighborhood $U \subset B$ on which the basis (ϵ_a) is defined.

►Point out special cases?◄

Recall that a real function $f \in C^\infty(B, \mathbb{R})$ is a smooth map $f : B \to \mathbb{R}$. However, each real function uniquely determines a smooth section σ of the trivial line bundle $B \times \mathbb{R}$, by setting

$$\begin{array}{rcccc} \sigma & : & B & \to & B \times \mathbb{R} \\ & & p & \mapsto & (p, f(p)) \end{array} , \tag{5.1.4}$$

and conversely, every such section defines a real function $f = \mathrm{pr}_2 \circ \sigma$. In other words, there is a canonical bijection such that $C^\infty(B, \mathbb{R}) \cong \Gamma(B \times \mathbb{R})$. Identifying $B \times \mathbb{R} \cong E^r_s$, we can thus interpret a real function as a tensor field of type $(0, 0)$ for any vector bundle.

5.2 Product of tensor fields

There are different ways to obtain tensor fields from simpler ones. A rather simple construction works as follows.

Definition 5.2.1 (Tensor field product)**.** Let $\pi : E \to B$ be a vector bundle and $T \in \Gamma(E^r_s)$ and $U \in \Gamma(E^t_u)$ be tensor fields. Their *tensor product* is a tensor field $T \otimes U \in \Gamma(E^{r+t}_{s+u})$ such that for each $p \in B$,

$$(T \otimes U)(p) = T(p) \otimes U(p)\,. \tag{5.2.1}$$

This definition can most easily be understood using a local basis (ϵ_a) of E. Let $T \in \Gamma(E^r_s)$ and $U \in \Gamma(E^t_u)$, and write $V = T \otimes U$. The tensor product is given by

$$\begin{aligned} T \otimes U &= \left(T^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} \epsilon_{a_1} \otimes \ldots \otimes \epsilon_{a_r} \otimes \bar{\epsilon}^{b_1} \otimes \ldots \otimes \bar{\epsilon}^{b_s}\right) \\ &\quad \otimes \left(U^{c_1 \cdots c_t}{}_{d_1 \cdots d_u} \epsilon_{c_1} \otimes \ldots \otimes \epsilon_{c_t} \otimes \bar{\epsilon}^{d_1} \otimes \ldots \otimes \bar{\epsilon}^{d_u}\right) \\ &= T^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} U^{c_1 \cdots c_t}{}_{d_1 \cdots d_u} \\ &\quad \epsilon_{a_1} \otimes \ldots \otimes \epsilon_{a_r} \otimes \bar{\epsilon}^{b_1} \otimes \ldots \otimes \bar{\epsilon}^{b_s} \otimes \epsilon_{c_1} \otimes \ldots \otimes \epsilon_{c_t} \otimes \bar{\epsilon}^{d_1} \otimes \ldots \otimes \bar{\epsilon}^{d_u} \end{aligned} \tag{5.2.2}$$

and yields the components

$$V^{a_1 \cdots a_r}{}_{b_1 \cdots b_s}{}^{c_1 \cdots c_t}{}_{d_1 \cdots d_u} = T^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} U^{c_1 \cdots c_t}{}_{d_1 \cdots d_u}\,. \tag{5.2.3}$$

One might be worried that the basis elements ϵ_a and $\bar{\epsilon}^a$ appear now in "mixed order", in contrast to the definition 4.3.3 of the tensor bundle. This is not a problem, since the tensor product bundles $E \otimes E^*$ and $E^* \otimes E$ are canonically isomorphic, so one can simply define a new tensor field $\tilde{V}$ such that

$$\tilde{V}^{a_1 \cdots a_r c_1 \cdots c_t}{}_{b_1 \cdots b_s d_1 \cdots d_u} = V^{a_1 \cdots a_r}{}_{b_1 \cdots b_s}{}^{c_1 \cdots c_t}{}_{d_1 \cdots d_u}\,. \tag{5.2.4}$$

However, this does *not* mean that changing the order of indices does not change the tensor field - V and $\tilde{V}$ carry the same information, but encoded differently. As another simple example, the tensor fields

$$V_{ab} \bar{\epsilon}^a \otimes \bar{\epsilon}^b \quad \text{and} \quad V_{ba} \bar{\epsilon}^a \otimes \bar{\epsilon}^b = V_{ab} \bar{\epsilon}^b \otimes \bar{\epsilon}^a \tag{5.2.5}$$

are (for general V_{ab}) not the same! This will be discussed in more detail in section 5.4.

There is another possibility to understand the tensor product of tensor fields. To see this, recall that the pair (T, U) with $T \in \Gamma(E^r_s)$ and $U \in \Gamma(E^t_u)$ constitutes a section of the fibered product $E^r_s \times_B E^t_u$, since for each $p \in B$, it defines a pair $(T(p), U(p))$ with $\pi^r_s(T(p)) = \pi^t_u(U(p)) = p$. This leads to the following statement.

Theorem 5.2.1. *The tensor field product* $\otimes : \Gamma(E^r_s) \times \Gamma(E^t_u) \to \Gamma(E^{r+t}_{s+u})$ *is induced by a fiber bundle morphism* $\theta^{r,t}_{s,u} : E^r_s \times_B E^t_u \to E^{r+t}_{s+u}$*, such that*

$$T \otimes U = \theta^{r,t}_{s,u} \circ (T, U) \tag{5.2.6}$$

for all $T \in \Gamma(E^r_s)$ *and* $U \in \Gamma(E^t_u)$*.*

Proof. ▶. . . ◀ ■

As discussed in section 4.2, the fibered product $E^r_s \times_B E^t_u$ is *not* equipped with a vector bundle structure, and so we can only refer to the map $\theta^{r,t}_{s,u}$ defined above as a finber bundle morphism. However, one may pose the question whether one can equip $E^r_s \times_B E^t_u$ with an additional structure, which allows to obtain further properties of $\theta^{r,t}_{s,u}$. A naive approach could be to introduce a vector bundle structure by replacing $E^r_s \times_B E^t_u$ with the Whitney sum $E^r_s \oplus E^t_u$, and ask whether this turns $\theta^{r,t}_{s,u}$ into a vector bundle homomorphism. However, it turns out that this is *not* the case. ►...◄

5.3 Contraction of tensor fields

After showing a way how to construct higher tensor fields from simpler ones, we also show a way how to obtain simpler tensor fields.

Definition 5.3.1 (Tensor field contraction)**.** Let $\pi : E \to B$ be a vector bundle and $\Gamma(E^r_s)$ the space of tensors of type (r, s) of E with $r, s \geq 1$. The *contraction* of the k'th and l'th tensor component, where $1 \leq k \leq r$ and $1 \leq l \leq s$, is the unique linear function

$$\mathrm{tr}^k_l : \Gamma(E^r_s) \to \Gamma(E^{r-1}_{s-1}), \tag{5.3.1}$$

such that

$$\mathrm{tr}^k_l(A \otimes B \otimes C \otimes D \otimes E \otimes F) = \langle B, E\rangle A \otimes C \otimes D \otimes F \tag{5.3.2}$$

for all $A \in \Gamma(E^{k-1}_0)$, $B \in \Gamma(E^1_0)$, $C \in \Gamma(E^{r-k}_0)$, $D \in \Gamma(E^0_{l-1})$, $E \in \Gamma(E^0_1)$, $F \in \Gamma(E^0_{s-l})$.

Also this construction is most easily illustrated using coordinates. Let $T \in \Gamma(E^r_s)$ a tensor field of type (r, s) of E. In a local basis it is expanded as

$$T = T^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} \epsilon_{a_1} \otimes \ldots \otimes \epsilon_{a_r} \otimes \bar{\epsilon}^{b_1} \otimes \ldots \otimes \bar{\epsilon}^{b_s} \,. \tag{5.3.3}$$

We then apply the relation (5.3.2) from definition 5.3.1 with

$$A = \epsilon_{a_1} \otimes \ldots \otimes \epsilon_{a_{k-1}} \,, \qquad B = \epsilon_{a_k} \,, \qquad C = \epsilon_{a_{k+1}} \otimes \ldots \otimes \epsilon_{a_r} \,, \tag{5.3.4a}$$

$$D = \bar{\epsilon}^{b_1} \otimes \ldots \otimes \bar{\epsilon}^{b_{l-1}} \,, \qquad E = \bar{\epsilon}^{b_l} \,, \qquad F = \bar{\epsilon}^{b_{l+1}} \otimes \ldots \otimes \bar{\epsilon}^{b_s} \,, \tag{5.3.4b}$$

which yields

$$\begin{aligned} &\mathrm{tr}^k_l \left(\epsilon_{a_1} \otimes \ldots \otimes \epsilon_{a_r} \otimes \bar{\epsilon}^{b_1} \otimes \ldots \otimes \bar{\epsilon}^{b_s}\right) = \\ &\qquad \underbrace{\langle \epsilon_{a_k}, \bar{\epsilon}^{b_l}\rangle}_{=\delta^{b_l}_{a_k}} \left(\epsilon_{a_1} \otimes \ldots \otimes \widehat{\epsilon_{a_k}} \otimes \ldots \otimes \epsilon_{a_r} \otimes \bar{\epsilon}^{b_1} \otimes \ldots \otimes \widehat{\bar{\epsilon}^{b_l}} \otimes \ldots \otimes \bar{\epsilon}^{b_s}\right) \,, \end{aligned} \tag{5.3.5}$$

where we used the "hat notation" to mark elements which are to be omitted from the tensor product. Now using the fact that tr^k_l is by definition linear, we can use the basis expansion (5.3.3) of the tensor field T to obtain Its contraction of the k'th and l'th component then simply takes the form

$$\begin{aligned} \mathrm{tr}^k_l T &= T^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} \mathrm{tr}^k_l \left(\epsilon_{a_1} \otimes \ldots \otimes \epsilon_{a_r} \otimes \bar{\epsilon}^{b_1} \otimes \ldots \otimes \bar{\epsilon}^{b_s}\right) \\ &= T^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} \delta^{b_l}_{a_k} \left(\epsilon_{a_1} \otimes \ldots \otimes \widehat{\epsilon_{a_k}} \otimes \ldots \otimes \epsilon_{a_r} \otimes \bar{\epsilon}^{b_1} \otimes \ldots \otimes \widehat{\bar{\epsilon}^{b_l}} \otimes \ldots \otimes \bar{\epsilon}^{b_s}\right) \,. \end{aligned} \tag{5.3.6}$$

In other words, the components of $\mathrm{tr}^k_l T$ are obtained simply by summation over the k'th upper and l'th lower indices.

►Introduce morphism $E^r_s \to E^{r-1}_{s-1}$.◄

5.4 Symmetry decomposition

Definition 5.4.1 (Transpose). ►...◄

5.5 Canonical tensor fields

Being a vector bundle, any tensor bundle, of course, possesses a canonical section, namely the zero section given in definition 3.3.1. However, there are tensor bundles which possess also another canonical sections. It turns out that this is the case for tensor bundles E^r_s with $r = s$. The existence of these sections becomes clear by realizing that these are simply tensor powers of the endomorphism bundle,

$$E^r_r \cong \bigotimes^r \operatorname{End}(E)\,. \tag{5.5.1}$$

We will therefore start with the latter. Here the existence of a canonical section arises from the fact that the space of endomorphisms over a vector space possesses a canonical unit element. This allows us to define the following notion.

Definition 5.5.1 (Unit section). Let $(E, B, \pi, \mathbb{R}^k)$ a vector bundle and $\operatorname{End}(E) \cong E^1_1$ its endomorphism bundle. The *unit section* $\delta \in \Gamma(\operatorname{End}(E))$ is the section defined by $\delta(p) = \operatorname{id}_{E_p}$ for all $p \in B$.

It follows from the definition of the homomorphism and endomorphism bundles that this is indeed a smooth section for every vector bundle E. This can be derived from the following statement:

Theorem 5.5.1. *The one-to-one correspondence given in theorem 4.6.2 relates the unit section $\delta \in \Gamma(\operatorname{End}(E))$ and the identity map $\operatorname{id}_E : E \to E$.*

Proof. Using the notation from the proof of theorem 4.6.2, we set $\theta = \operatorname{id}_E$. Then on every fiber E_p with $p \in B$, θ restricts to $\operatorname{id}_{E_p} = \delta(p)$, and so $\vartheta = \delta$. Conversely, if we start with $\vartheta = \delta$, then we find

$$\theta(e) = \delta(\pi(e))e = \operatorname{id}_{E_{\pi(e)}} e = e\,, \tag{5.5.2}$$

and so $\theta = \operatorname{id}_E$. ■

Hence, for any section $v \in \Gamma(E)$ we have $\delta v = \operatorname{id}_E \circ v = v$. We will encounter other incantations of the unit section later, e.g., as a vector-valued one-form in chapter 17 or as a tensor field, canonically identified with a vector bundle isomorphism via theorem 5.5.1, in various places.

Recall given induced coordinates on E, for the endomorphism bundle we have the basis $(\epsilon_\mu \otimes \bar{\epsilon}^\nu)$, constructed from the basis (ϵ_μ) of E and corresponding dual basis $(\bar{\epsilon}^\mu)$ of E^*. In this basis the unit section δ can be written as

$$\delta = \epsilon_\mu \otimes \bar{\epsilon}^\mu = \delta^\mu_\nu \epsilon_\mu \otimes \bar{\epsilon}^\nu\,, \tag{5.5.3}$$

so that its components are given by the Kronecker symbol

$$\delta^\mu_\nu = \begin{cases} 1 & \mu = \nu\,, \\ 0 & \mu \neq \nu\,. \end{cases} \tag{5.5.4}$$

This is the reason for introducing the notation δ in definition 5.5.1.

Given a canonical section of E^1_1, one can of course obtain sections of its tensor product bundles ►...◄

Chapter 6

Affine bundles

6.1 Affine bundles

One important feature of a vector space, and hence every fiber of a vector bundle, is the existence of a distinguished zero element. Often we will encounter bundles whose structure is similar to that of vector bundles, but for which no such distinguished element exists. In this case the fibers are not vector spaces, but only *affine* spaces. Consequently, we may define the following:

Definition 6.1.1 (Affine bundle)**.** A (real) *affine bundle* of rank $k \in \mathbb{N}$ is a fiber bundle $(A, B, \varpi, \mathbb{R}^k)$ such that for all $p \in B$ the fiber $A_p = \varpi^{-1}(p)$ is a real affine space of dimension k and such that the restrictions of the local trivializations $\varphi : \varpi^{-1}(U) \to U \times \mathbb{R}^k$ to a fiber A_p for $p \in U$ are affine space isomorphisms from A_p to $\{p\} \times \mathbb{R}^k$.

We know from vector bundles that the definition guarantees a compatibility of the algebraic and geometric structures, in the sense that we may extend the linear vector space operations, which are defined on each fiber, also to (smooth) sections, and we will obtain (smooth) sections again. A similar compatibility holds also for affine bundles. However, recall that for an affine space the operation of addition needs an element of an underlying vector space, and so for every fiber A_p we must have also a vector space, conventionally denoted $\vec{A}_p$, so that addition is defined as a function $+ : A_p \times \vec{A}_p \to A_p$. It is not surprising that the vector spaces $\vec{A}_p$ constitute the fibers of a vector bundle. To see this, we first define the following notion.

Definition 6.1.2 (Affine bundle modeled over a vector bundle)**.** A (real) *affine bundle* modeled over a vector bundle $(E, B, \pi, \mathbb{R}^k)$ is a fiber bundle $(A, B, \varpi, \mathbb{R}^k)$ such that:

1. For all $p \in B$ the fiber $A_p = \varpi^{-1}(p)$ is an affine space modeled over the vector space $E_p = \pi^{-1}(p)$, i.e., $E_p = \vec{A}_p$.
2. For each local trivialization (U, ϕ) of E there exists a local trivialization (U, φ) of A such that for all $p \in U$ the restrictions $\phi_p = (\mathrm{pr}_2 \circ \phi)|_{E_p} : E_p \to \mathbb{R}^k$ and $\varphi_p = (\mathrm{pr}_2 \circ \varphi)|_{A_p} : A_p \to \mathbb{R}^k$ satisfy

$$\varphi_p(a + e) = \varphi_p(a) + \phi_p(e) \tag{6.1.1}$$

for all $a \in A_p$ and $e \in E_p$.

This definition follows [KSM93, 6.22]; an alternative definition is given in [Sau89, def. 2.4.4], and we will show their equivalence below. However, note that yet another definition can be found in [GMS09, sec. 1.1.3] and [Sar13, sec. 1.2], which is *not* equivalent and, in fact, too weak to guarantee the relation between the affine bundle and the vector bundle which we will derive below[1].

One may wonder whether such an underlying vector bundle exists for every affine bundle according to definition 6.1.1 and whether it is unique. We find that both is true:

Theorem 6.1.1. *For every affine bundle $(A, B, \varpi, \mathbb{R}^k)$ there exists a unique vector bundle $(\vec{A}, B, \vec{\varpi}, \mathbb{R}^k)$ over which it is modeled.*

Proof. In order to show its existence, we will now explicitly construct this vector bundle. We start with the total space, which we denote by $\vec{A}$. From the fact that A is an affine bundle follows that for all $p \in B$, A_p is an affine space. Denoting by $\vec{A}_p$ the underlying vector space, we define

$$\vec{A} = \biguplus_{p\in B} \vec{A}_p\,, \tag{6.1.2}$$

together with the projection $\vec{\varpi}$ such that $\vec{\varpi}^{-1}(p) = \vec{A}_p$. To equip $\vec{A}$ with the structure of a manifold, we still need to specify an atlas. For this purpose it is enough to specify the local trivializations of the bundle $\vec{\varpi} : \vec{A} \to B$, and then to use induced charts. Here we use the fact that every element $e \in \vec{A}_p$ for some $p \in B$ can be written as $e = a_1 - a_2$, where $a_1, a_2 \in A_p$. Given a local trivialization (U, φ) of $\varpi : A \to B$, we may thus define

$$\phi : \vec{\varpi}^{-1}(U) \to U \times \mathbb{R}^k\,, \quad \mathrm{pr}_2(\phi(a_1 - a_2)) = \mathrm{pr}_2(\varphi(a_1)) - \mathrm{pr}_2(\varphi(a_2))\,, \tag{6.1.3}$$

One easily checks that ϕ is well-defined, i.e., independent of the choice of a_1, and depends only on the difference $a_1 - a_2$, which follows from the fact that φ reduces to an affine morphism on every fiber A_p. From the same fact further follows that ϕ reduces to a linear function on each fiber $\vec{A}_p$. Hence, it equips $\vec{A}$ with the structure of a vector bundle.

One now easily checks that, by construction, $\vec{A}$ has the two properties given in definition 6.1.2, since we chose both the fibers $\vec{A}_p$ and the local trivializations just to match these conditions. We also see that definition 6.1.2 uniquely specifies the fibers and a set of local trivializations, so that the vector bundle we constructed is indeed the unique vector bundle satisfying these conditions. ∎

Given the statement about the underlying vector bundle, which we will denote $\vec{A}$ in the following, we can now discuss the relation between these two bundles. We find that the following holds:

Theorem 6.1.2. *For every affine bundle $(A, B, \varpi, \mathbb{R}^k)$ the functions*

$$\begin{array}{cccc} + & : & A \times_B \vec{A} & \to & A \\ & & (a, e) & \mapsto & a + e \end{array} \tag{6.1.4}$$

and

$$\begin{array}{cccc} - & : & A \times_B A & \to & \vec{A} \\ & & (a_1, a_2) & \mapsto & a_1 - a_2 \end{array} \tag{6.1.5}$$

are fiber bundle morphisms covering the identity on B.

[1] The definition in [GMS09, sec. 1.1.3] and [Sar13, sec. 1.2] demands only that each fiber of the affine bundle is an affine space modeled over the corresponding fiber of the vector bundle, but does not make any reference to the geometry of the vector bundle. Without taking the geometry into account, any two vector bundles, which have the same fibers, would be treated alike - for example, a cylinder and a Möbius strip - and the uniqueness in theorem 6.1.1 does not hold. But the geometry is important when it comes to relating sections and maps on these bundles, since otherwise theorems like 6.3.3 or 6.4.1 would fail.

Proof. The proof is very similar to the proof of theorem 3.1.1. For $p \in B$, let $U \subset B$ with $p \in U$ such that there exists a chart (U, υ) of B with $\upsilon : U \to \mathbb{R}^n$, a local trivialization (U, ϕ) of A with $\phi : \varpi^{-1}(U) \to U \times \mathbb{R}^k$ and the unique local trivialization $(U, \vec{\phi})$ of $\vec{A}$ with $\vec{\phi} : \vec{\varpi}^{-1}(U) \to U \times \mathbb{R}^k$ defined such that

$$\mathrm{pr}_2(\phi(a + e)) = \mathrm{pr}_2(\phi(a)) + \mathrm{pr}_2(\vec{\phi}(e)) \tag{6.1.6}$$

for all $a \in \varpi^{-1}(U)$ and $e \in \vec{\varpi}^{-1}(U)$ with $\varpi(a) = \vec{\varpi}(e)$. This allows us to define induced charts (X, ψ) of A with $X = \varpi^{-1}(U)$ and

$$\psi(a) = (\upsilon(\varpi(a)), \mathrm{pr}_2(\phi(a))) \in \mathbb{R}^n \times \mathbb{R}^k \cong \mathbb{R}^{n+k}, \tag{6.1.7}$$

as well as $(\vec{X}, \vec{\psi})$ of $\vec{A}$ with $\vec{X} = \vec{\varpi}^{-1}(U)$ and

$$\vec{\psi}(e) = (\upsilon(\vec{\varpi}(e)), \mathrm{pr}_2(\vec{\phi}(e))) \in \mathbb{R}^n \times \mathbb{R}^k \cong \mathbb{R}^{n+k}. \tag{6.1.8}$$

We will use these components to check that the two functions given above are smooth maps:

First, we construct the set

$$V = (\varpi \times_B \varpi)^{-1}(U) = \{(a_1, a_2) \in A \times A, \varpi(a_1) = \varpi_(a_2) \in U\}, \tag{6.1.9}$$

on which we define an induced chart (V, χ) by

$$\begin{array}{cccc} \chi & : & V & \to & \mathbb{R}^{n+2k} \\ & & (a_1, a_2) & \mapsto & (\upsilon(\varpi(a_1)), \mathrm{pr}_2(\phi(a_1)), \mathrm{pr}_2(\phi(a_2))) \end{array}. \tag{6.1.10}$$

On this chart, the subtraction takes the form

$$\begin{array}{ccccc} \vec{\psi} \circ - \circ \chi^{-1} & : & \chi(V) & \to & \psi(X) \\ & & (u, v, w) & \mapsto & (u, v - w) \end{array}, \tag{6.1.11}$$

where we used the fact that ϕ restricts to an affine space isomorphism on each fiber, while $\vec{\phi}$ is defined by 6.1.6, and hence

$$\mathrm{pr}_2(\vec{\phi}(a_1 - a_2)) = \mathrm{pr}_2(\phi(a_1)) - \mathrm{pr}_2(\phi(a_2)), \tag{6.1.12}$$

while the base point remains unchanged,

$$\vec{\varpi}(a_1 - a_2) = \varpi(a_1) = \varpi(a_2). \tag{6.1.13}$$

The function $(u, v, w) \mapsto (u, v - w)$ is smooth, and so it follows that subtraction is smooth.

We proceed analogously with the addition. Let

$$W = (\varpi \times_B \vec{\varpi})^{-1}(U) = \{(a, e) \in A \times \vec{A}, \varpi(a) = \vec{\varpi}_(e) \in U\}, \tag{6.1.14}$$

on which we define an induced chart (W, ω) by

$$\begin{array}{cccc} \omega & : & W & \to & \mathbb{R}^{n+2k} \\ & & (a, e) & \mapsto & (\upsilon(\varpi(a)), \mathrm{pr}_2(\phi(a)), \mathrm{pr}_2(\vec{\phi}(e))) \end{array}. \tag{6.1.15}$$

On this chart, the addition takes the form

$$\begin{array}{ccccc} \psi \circ + \circ \omega^{-1} & : & \omega(W) & \to & \psi(X) \\ & & (u, v, w) & \mapsto & (u, v + w) \end{array}, \tag{6.1.16}$$

where we used the fact that ϕ restricts to an affine space isomorphism on each fiber, while $\vec{\phi}$ restricts to its linear derivative, and hence 6.1.6 holds, while the base point remains unchanged,

$$\varpi(a + e) = \varpi(a) = \vec{\varpi}(e). \tag{6.1.17}$$

The function $(u, v, w) \mapsto (u, v + w)$ is smooth, and so it follows that addition is smooth.

Finally, we have seen that both for addition and subtraction the base point is unchanged. Hence, both operations define bundle morphisms covering the identity on B. ■

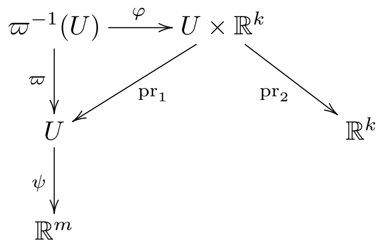
Note that also in the proof of the previous theorem we have relied on the fact that the local trivializations (U, ϕ) and $(U, \vec{\phi})$ of A and $\vec{A}$ are related. If this were not the case, the construction given in the proof would not have been possible. As another application, we will show that every vector bundle is an affine bundle by itself.

Theorem 6.1.3. *Every vector bundle $(E, B, \pi, \mathbb{R}^k)$ is an affine bundle modeled over itself, $\vec{E} = E$.*

Proof. By definition, the fibers E_p for $p \in B$ of a vector bundle are vector spaces, and thus also affine spaces modeled over themselves. Further, the local trivializations (U, ϕ) restrict to vector space isomorphisms on each fiber, and so $\varphi = \phi$ satisfies the requirements given in definition 6.1.2. ■

6.2 Induced charts and fiber coordinates

Similarly to vector bundles as discussed in section 3.2, one may also in the case of affine bundles construct a particularly useful set of coordinates by using induced charts. Given an affine bundle $(A, B, \varpi, \mathbb{R}^k)$ with a chart (U, ψ) of B and local trivialization (U, φ), one has the diagram



(6.2.1)

The induced chart (W, ω) of A is thus given by $W = \varphi^{-1}(U \times \mathbb{R}^k) = \varpi^{-1}(U)$ and

$$\omega(a) = (\psi(\varpi(a)), \mathrm{pr}_2(\varphi(a))) \tag{6.2.2}$$

for $a \in W$. Since φ restricts to a isomorphisms of affine spaces on every fiber A_p, the same holds for the combined function

$$A_p \xrightarrow{\varphi} \{p\} \times \mathbb{R}^k \xrightarrow{\mathrm{pr}_2} \mathbb{R}^k \tag{6.2.3}$$

The virtue of using these induced charts becomes more apparent if we also use the corresponding induced charts on the underlying vector bundle $(\vec{A}, B, \vec{\varpi}, \mathbb{R}^k)$, which are given from a local trivialization (U, ϕ) of $\vec{A}$ related to (U, φ) as in definition 6.1.2. In this case we find

$$(\mathrm{pr}_2 \circ \varphi)(a + e) = (\mathrm{pr}_2 \circ \varphi)(a) + (\mathrm{pr}_2 \circ \phi)(e) \tag{6.2.4}$$

for $a \in A_p$, $e \in \vec{A}_p$ and $p \in U$.

Finally, denoting the coordinates on B by (x^i), the coordinates on A by (x^i, y^μ) and the coordinates on $\vec{A}$ by $(x^i, \vec{y}^\mu)$, where $i = 1, \ldots, m$ and $\mu = 1, \ldots, k$, we find that the bundle morphisms we encountered in theorem 6.1.2 are expressed in coordinates as

$$\begin{array}{cccc} + & : & A \times_B \vec{A} & \to & A \\ & & (x^i, y^\mu, \vec{y}^\mu) & \mapsto & (x^i, y^\mu + \vec{y}^\mu) \end{array} \tag{6.2.5}$$

and

$$\begin{array}{cccc} - & : & A \times_B A & \to & \vec{A} \\ & & (x^i, y^\mu, y'^\mu) & \mapsto & (x^i, y^\mu - y'^\mu) \end{array} . \tag{6.2.6}$$

Similarly to the case of vector bundles, as shown in theorem 3.2.1 we find:

Theorem 6.2.1. *For the induced affine bundle charts the transition functions are affine transformations of the fiber coordinates.*

Proof. The proof is analogous to the proof of theorem 3.2.1, with the only difference being that for an affine bundle A the respective restrictions $\phi|_{A_p}$ and $\tilde{\phi}|_{A_p}$ are affine transformations instead of vector space isomorphisms. Hence, also $\tilde{\phi}_{A_p} \circ (\phi|_{A_p})^{-1}$ and its inverse are affine transformations of the fiber coordinates at the point p. ■

6.3 Affine bundle morphisms

As for vector bundles, also affine bundles allow us to define a particular class of bundle morphisms, which preserves not only the fiber bundle structure, but also the affine structure on each fiber. We define these morphisms as follows.

Definition 6.3.1 (Affine bundle morphism)**.** Let $(A_1, B_1, \varpi_1, F_1)$ and $(A_2, B_2, \varpi_2, F_2)$ be affine bundles. A *affine bundle morphism* (or *affine bundle homomorphism*) is a bundle morphism $\theta : A_1 \to A_2$ covering a map $\vartheta : B_1 \to B_2$ such that for each $p \in B_1$ the restriction of θ to the fiber $\varpi_i^{-1}(p)$ is an affine function between the affine spaces $\varpi_1^{-1}(p)$ and $\varpi_2^{-1}(\vartheta(p))$.

Also in this case we can define an isomorphism as an invertible morphism.

Definition 6.3.2 (Affine bundle isomorphism)**.** An *affine bundle isomorphism* is a bijective affine bundle morphism whose inverse is also an affine bundle morphism. If an affine bundle morphism between two affine bundles exists, these bundles are called *isomorphic*.

Finally, as seen before, also for affine bundles the following statement holds.

Theorem 6.3.1. *Let $(A_i, B_i, \varpi_i, F_i)$ with $i = 1, \ldots, 3$ be affine bundles and $\theta : A_1 \to A_2$ and $\theta' : A_2 \to A_3$ be affine bundle homomorphisms (isomorphisms) covering $\vartheta : B_1 \to B_2$ and $\vartheta' : B_2 \to B_3$. Then $\theta' \circ \theta : A_1 \to A_3$ is an affine bundle homomorphism (isomorphism) covering $\vartheta' \circ \vartheta : B_1 \to B_3$.*

Proof. ▶. . . ◀ ■

We have already seen in the definition of affine bundles that each affine bundle comes with a unique vector bundle, and that their geometries are closely related to each other. It follows that also their morphisms are closely related. A particular example for such a relation is shown in the following statement.

Theorem 6.3.2. *Let $(A_1, B_1, \varpi_1, F_1)$ and $(A_2, B_2, \varpi_2, F_2)$ be affine bundles and $\theta : A_1 \to A_2$ an affine bundle morphism covering a map $\vartheta : B_1 \to B_2$. Then there exists a unique vector bundle morphism $\vec{\theta} : \vec{A}_1 \to \vec{A}_2$ covering ϑ such that*

$$\theta(a + e) = \theta(a) + \vec{\theta}(e) \tag{6.3.1}$$

for all $(a, e) \in A_1 \times_{B_1} \vec{A}_1$.

Proof. Let $e \in \vec{A}_1$ with $\vec{\varpi}_1(e) = p \in B_1$. Then we can find $a, a' \in (A_1)_p$ such that $a - a' = e$. We use these to define

$$\vec{\theta}(e) = \theta(a) - \theta(a') \in \vec{A}_2 \,. \tag{6.3.2}$$

This is defined independently of the choice of the representatives a, a', since θ is an affine bundle morphism, and so it restricts to an affine function on every fiber. The latter also implies that $\vec{\theta}$ restricts to a linear function on every fiber, and that

$$\vec{\varpi}_2(\vec{\theta}(e)) = \varpi_2(\theta(a)) = \vartheta(\varpi_1(a)) = \vartheta(\vec{\varpi}_1(e)) \,. \tag{6.3.3}$$

We finally need to show that $\vec{\theta} : A_1 \to A_2$ is a smooth map. For $e \in \vec{A}_1$ and $p = \vec{\varpi}_1(e) \in B_1$, we can find an open set $U \subset B_1$ such that $p \in U$, as well as a local section $\alpha : U \to A_1$. On $\vec{\varpi}_1^{-1}(U)$, we can then define

$$\begin{array}{rcccl} \hat{\alpha} & : & \vec{\varpi}_1(U) & \to & A_1 \times_{B_1} \vec{A}_1 \\ & & e & \mapsto & (\alpha(\vec{\varpi}_1(e)), e \end{array} \,, \tag{6.3.4}$$

which is smooth, since it is constructed from smooth maps. This can be composed with the smooth map $+ : A_1 \times_{B_1} \vec{A}_1 \to A_1$, and then further with $\theta : A_1 \to A_2$. With these we have the smooth map

$$\begin{array}{rcccl} \theta_\alpha & : & \vec{\varpi}_1^{-1}(U) & \to & A_2 \times_{B_2} A_2 \\ & & e & \mapsto & (\theta(\alpha(\vec{\varpi}_1(e)) + e), \theta(\alpha(\vec{\varpi}_1(e)))) \end{array} \,. \tag{6.3.5}$$

Composing with $- : A_2 \times_{B_2} A_2 \to \vec{A}_2$, we have a smooth map which satisfies

$$\theta(\alpha(\vec{\varpi}_1(e)) + e) - \theta(\alpha(\vec{\varpi}_1(e))) = \vec{\theta}(\alpha(\vec{\varpi}_1(e)) + e - \alpha(\vec{\varpi}_1(e))) = \vec{\theta}(e) \,, \tag{6.3.6}$$

and so $\vec{\theta}$ is smooth. ∎

The unique vector bundle morphism $\vec{\theta}$ constructed in the previous theorem has its own name, and we define as follows.

Definition 6.3.3 (Linear derivative)**.** Given an affine bundle morphism $\theta : A_1 \to A_2$, the unique vector bundle morphism $\vec{\theta} : \vec{A}_1 \to \vec{A}_2$ from theorem 6.3.2 is called the *linear derivative* of θ.

The construction is illustrated in figure 6.1.

6.4 Sections of affine bundles

The fact that every fiber $A_p = \varpi^{-1}(p)$ of an affine bundle is an affine space modeled over the vector space $E_p = \pi^{-1}(p)$ allows for two operations. Given elements $a \in A_p$ and $e \in E_p$, one has their sum $a + e \in A_p$. Similarly, given two elements $a_1, a_2 \in A_p$ their difference $a_1 - a_2 \in E_p$ is defined. As it is also the case with linear operations on vector bundles, one may extend these operations to sections of the respective bundles, due to the fact that the affine and differentiable structures are compatible by the definition of an affine bundle. Hence, the following holds:

Theorem 6.4.1. *Let $(E, B, \pi, \mathbb{R}^k)$ be a vector bundle and $(A, B, \varpi, \mathbb{R}^k)$ an affine bundle modeled over E. Then the following objects,*

1. *for smooth sections $\varsigma \in \Gamma(A)$ and $\sigma \in \Gamma(E)$ the sum $\varsigma + \sigma \in \Gamma(A)$,*
2. *for smooth sections $\varsigma_1, \varsigma_2 \in \Gamma(A)$ the difference $\varsigma_1 - \varsigma_2 \in \Gamma(E)$,*

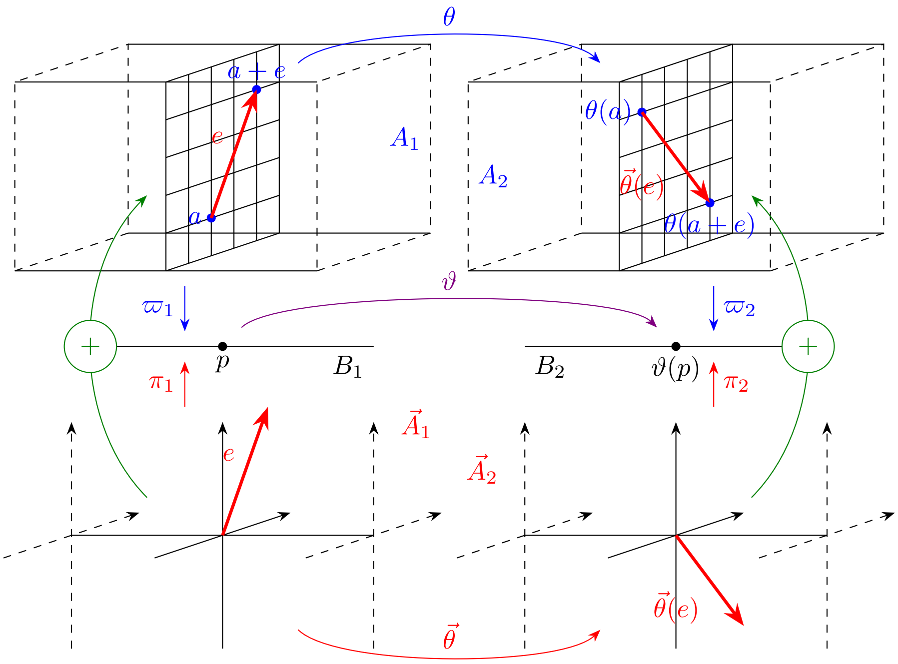


Figure 6.1: Construction of the linear derivative of an affine bundle morphism.

which are defined pointwise, are smooth sections.

Proof. The pair $(\varsigma, \sigma) : B \to A \times_B E$ is a smooth section of a fibered product. Since $+ : A \times_B E \to A$ is a smooth fiber bundle morphism covering the identity, their composition is a smooth section of A. The same argument holds for $\varsigma_1 - \varsigma_2$. ■

Note that in the previous statement we could also have used local sections, provided that they are defined on the same domain. For affine bundles, however, we have the following statement.

Theorem 6.4.2. *Every affine bundle has a global section.*

Proof. ►. . . ◄ ■

Note that there is an important difference with vector bundles. For the latter, there always exists a *canonical*, a priori uniquely defined global section, namely the zero section 3.3.1. For affine bundles, no such canonical section exists. To gain more insight into the space $\Gamma(A)$ of sections of an affine bundle, we show the following.

Theorem 6.4.3. *Let $(E, B, \pi, \mathbb{R}^k)$ be a vector bundle and $(A, B, \varpi, \mathbb{R}^k)$ an affine bundle modeled over E. Then there exists a one-to-one correspondence between global sections $\varsigma : B \to A$ of A and affine bundle isomorphisms $\theta : E \to A$ covering the identity whose linear derivative is $\vec{\theta} = \mathrm{id}_E$.*

Proof. In this construction we follow theorem 6.1.3 and understand E as an affine bundle modeled over itself. Given a section $\varsigma : B \to A$, one can define a map $\theta : E \to A$ as

$$\theta(e) = \varsigma(\pi(e)) + e\,. \tag{6.4.1}$$

The sum is well-defined, since

$$\varpi(\varsigma(\pi(e))) = \pi(e)\,, \tag{6.4.2}$$

and so both lie in fibers over the same base point $\pi(e)$. Now again

$$\varpi(\theta(e)) = \pi(e)\,, \tag{6.4.3}$$

and so we find $\varpi \circ \theta = \pi$, as necessary for a bundle morphism covering the identity. To show that it is an affine bundle morphism, we must check that it restricts to an affine function on every fiber. Let $\tilde{e} \in E_{\pi(e)}$. Then we have

$$\theta(e + \tilde{e}) = \varsigma(\pi(e)) + e + \tilde{e} = \theta(e) + \tilde{e}\,, \tag{6.4.4}$$

and so a displacement $\tilde{e}$ on the left hand side results in a displacement $\tilde{e}$ on the right hand side; since the latter depends linearly on the former, θ restricts to an affine function on every fiber. In fact, this linear relation is the identity, and so $\vec{\theta} = \mathrm{id}_E$. Further, from the smoothness of ς, π and the addition follows that also θ is smooth, since it is a combination of smooth maps. Finally, it is bijective, and we can explicitly give its inverse as

$$\theta^{-1}(a) = a - \varsigma(\varpi(a)) \tag{6.4.5}$$

for all $a \in A$. One checks analogously that also this is an affine bundle morphism. Hence, θ is an affine bundle isomorphism.

Conversely, given an affine bundle morphism $\theta : E \to A$ covering the identity on B whose linear derivative is $\vec{\theta} = \mathrm{id}_E$, one can define

$$\varsigma = \theta \circ 0\,, \tag{6.4.6}$$

where $0 : B \to E$ denotes the distinguished zero section of E. This is a smooth section, since both θ and 0 are smooth, and

$$\varpi \circ \theta \circ 0 = \pi \circ 0 = \mathrm{id}_B\,, \tag{6.4.7}$$

since 0 is a section and θ covers the identity. Finally, one easily checks that this construction reverses the construction of θ from ς given in the first part of the proof, establishing the one-to-one correspondence between these two objects. ■

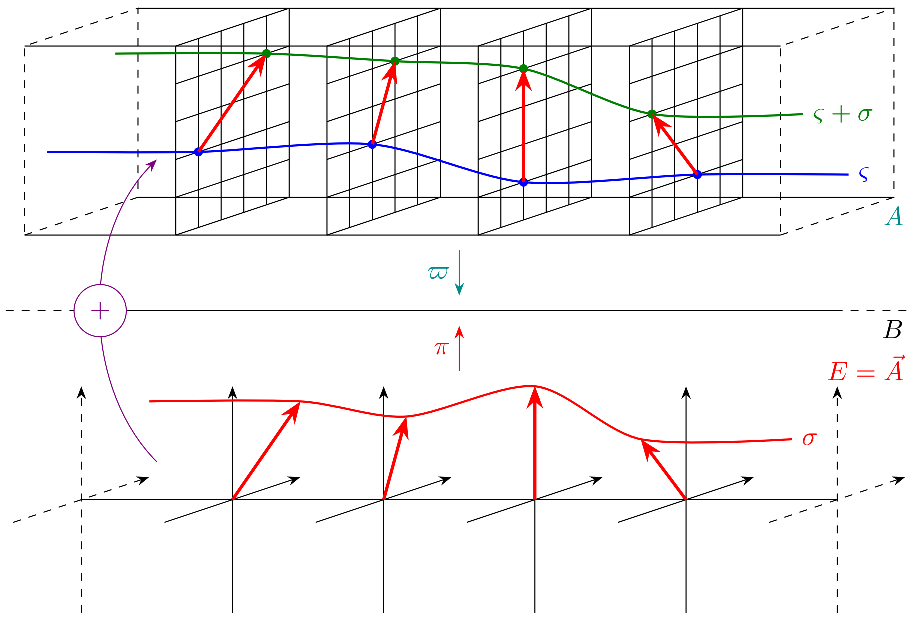


Figure 6.2: Relation between sections $\varsigma : B \to A$ and $\varsigma + \sigma : B \to A$ of an affine bundle $\varpi : A \to B$ and $\sigma : B \to E$ of the underlying vector bundle $\pi : E \to B$.

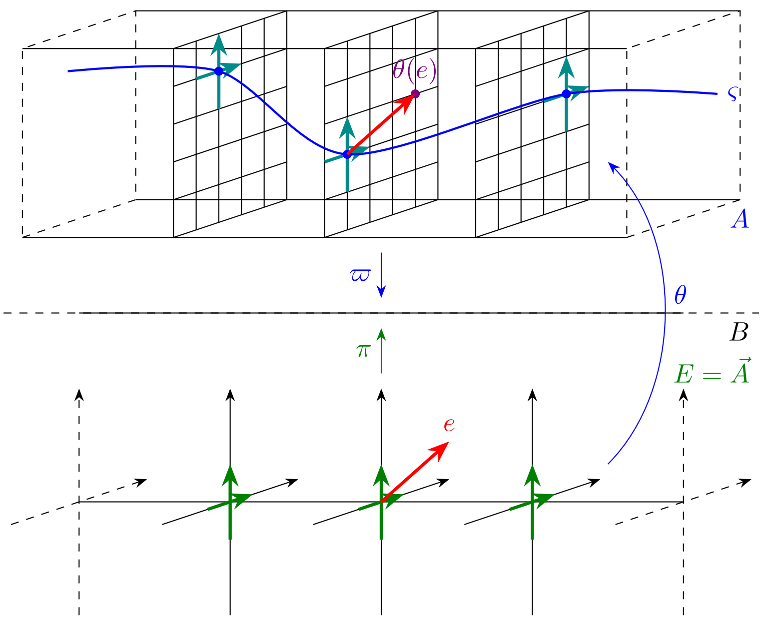


Figure 6.3: Relation between a section $\varsigma : B \to A$ of an affine bundle $\varpi : A \to B$ and an affine bundle morphism $\theta : E \to A$ from the underlying vector bundle $\pi : E \to B$ to A which satisfies $\vec{\theta} = \mathrm{id}_E$.

Chapter 7

Tangent bundle and vector fields

7.1 Derivations and tangent spaces

Every manifold is naturally equipped with a number of structures. One of the most basic and important structures is the tangent bundle. Geometrically it can be seen as the space of all vectors tangent to a manifold. In physics it appears most naturally in the context of mechanics: if the space of all possible positions of a point mass is modeled as a manifold, then its velocity is an element of the tangent bundle. The space of all tangent vectors at a given point is called the tangent space, and it can be defined in a number of different, but equivalent ways. Here we use a particularly simple definition in terms of derivations, and provide its geometric interpretation a bit later.

Definition 7.1.1 (Derivation)**.** Let M be a smooth manifold and $p \in M$. A *derivation* at p is a linear function $D : C^\infty(M, \mathbb{R}) \to \mathbb{R}$ such that it satisfies the Leibniz rule

$$D(fg) = D(f)g(p) + f(p)D(g) \tag{7.1.1}$$

for all $f, g \in C^\infty(M, \mathbb{R})$.

Recall that $C^\infty(M, \mathbb{R})$ denotes the space of smooth maps $f : M \to \mathbb{R}$. We remark that we could have chosen to work with manifolds of class C^k with finite $k > 0$ instead, and considered the larger space of functions $C^k(M, \mathbb{R})$. However, this poses some technical difficulties, which would require us to work with germs instead of maps. We avoid this here by working in the smooth category.

Further, note that we demand that a derivation D is linear, which means that for any $f, g \in C^\infty(M, \mathbb{R})$ and $\lambda, \mu \in \mathbb{R}$ we have

$$D(\lambda f + \mu g) = \lambda D(f) + \mu D(g)\,. \tag{7.1.2}$$

Together with the Leibniz rule, this has an important consequence.

Theorem 7.1.1. *For any $c \in \mathbb{R}$ and derivation D at $p \in M$ holds*

$$D(f_c) = 0\,, \tag{7.1.3}$$

where $f_c : M \to \mathbb{R}, p \mapsto c$ is the constant function.

Proof. For any $g \in C^\infty(M, \mathbb{R})$ holds

$$cD(g) = D(cg) = D(f_c g) = D(f_c)g(p) + f_c(p)D(g) = D(f_c)g(p) + cD(g)\,. \tag{7.1.4}$$

In particular, we can choose $g(p) \neq 0$, and solve this equation for $D(f_c) = 0$. ■

One easily checks the following:

Theorem 7.1.2. *The derivations at a point $p \in M$ of a manifold M, equipped with addition and scalar multiplication defined pointwise as*

$$(D_1 + D_2)(f) = D_1(f) + D_2(f) \quad \textit{and} \quad (\lambda D)(f) = \lambda D(f)\,, \tag{7.1.5}$$

form a vector space.

Proof. It is known from linear algebra that the pointwise sum and scalar multiple of linear functions is again a linear function. To check the Leibniz rule, we calculate

$$\begin{aligned}(D_1 + D_2)(fg) &= D_1(fg) + D_2(fg) \\ &= D_1(f)g(p) + f(p)D_1(g) + D_2(f)g(p) + f(p)D_2(g) \\ &= (D_1(f) + D_2(f))g(p) + f(p)(D_1(g) + D_2(g)) \\ &= (D_1 + D_2)(f)g(p) + f(p)(D_1 + D_2)(g)\end{aligned} \tag{7.1.6}$$

and

$$\begin{aligned}(\lambda D)(fg) &= \lambda D(fg) \\ &= \lambda(D(f)g(p) + f(p)D(g)) \\ &= (\lambda D)(f)g(p) + f(p)(\lambda D)(g)\,,\end{aligned} \tag{7.1.7}$$

showing that also the Leibniz rule holds. ■

This now leads us to the following definition.

Definition 7.1.2 (Tangent space). Let M be a smooth manifold and $p \in M$. The set of all derivations at p, equipped with addition and scalar multiplication defined as in theorem 7.1.2, is called the *tangent space* at p and denoted T_pM.

A question which arises immediately is whether this vector space has finite dimension, and how this depends on the manifold M. This is answered by the following statement:

Theorem 7.1.3. *Let M be a smooth manifold of dimension n. For each $p \in M$, the tangent space T_pM is a real vector space of dimension $n = \dim M$.*

Proof. Let (U, ϕ) be a chart such that $p \in U$. Since $\phi(U)$ is open, we can find $\epsilon > 0$ such that

$$V = \{x \in \mathbb{R}^n, \|x - \phi(p)\| < \epsilon\} \subset \phi(U)\,. \tag{7.1.8}$$

Now consider a function $f \in C^\infty(M, \mathbb{R})$, and write

$$\begin{array}{cccc} F & : & V & \to & \mathbb{R} \\ & & x & \mapsto & (f \circ \phi^{-1})(x) \end{array}\,. \tag{7.1.9}$$

Note that $F \in C^\infty(V, \mathbb{R})$. Using Hadamard's lemma we can write any smooth function F on V in the form

$$F(x) = F(x_0) + (x^a - x_0^a)\tilde{F}_a(x)\,, \tag{7.1.10}$$

where $x_0 = \phi(p)$ and $\tilde{F}_a$ are smooth functions on V. We can now define

$$\begin{array}{cccc} \phi_0^a & : & \phi^{-1}(V) & \to & \mathbb{R} \\ & & q & \mapsto & \phi^a(q) - \phi^a(p) = \phi^a(q) - x_0^a \end{array}\,, \tag{7.1.11}$$

as well as

$$\tilde f_a = \tilde F_a \circ \phi \quad : \quad \begin{array}{ccc} \phi^{-1}(V) & \to & \mathbb{R} \\ q & \mapsto & \tilde F_a(\phi(q)) \end{array} . \tag{7.1.12}$$

From these definitions follows that

$$f(q) = F(\phi(q)) = F(x_0) + \phi_0^a(q)\tilde F_a(\phi(q)) = f(p) + \phi_0^a(q)\tilde f_a(q) . \tag{7.1.13}$$

Now let D be a derivation at p. From the Leibniz rule and theorem 7.1.1 follows

$$D(f) = D(\phi_0^a \tilde f_a) = D(\phi_0^a)\tilde f_a(p) + \phi_0^a(p) D(\tilde f_a) = D(\phi_0^a)\tilde f_a(p) , \tag{7.1.14}$$

using $\phi_0^a(p) = 0$ in the last step. Note that the second factor is independent of D, and so the action of D on any function is fully determined by the n real numbers $u^a = D(\phi_0^a)$ which determine its action on ϕ_0^a. Now using the fact that

$$\tilde f_a(p) = \tilde F_a(x_0) = \left.\frac{\partial}{\partial x^a} F(x)\right|_{x=x_0} , \tag{7.1.15}$$

we can thus uniquely express every derivation as

$$D : f \mapsto D(f) = \sum_{a=1}^n u^a \left.\frac{\partial}{\partial x^a}(f \circ \phi^{-1})(x)\right|_{x=\phi(p)} \tag{7.1.16}$$

with $u \in \mathbb{R}^n$. Using theorem 7.1.2, this is obviously linear, and so it follows that T_pM is a vector space of dimension n. ■

The definition of the tangent space via derivations is probably the most intrinsic, as it does not refer to charts (even though we used them for the proof above). There are other, equivalent definitions. The following one, given in [Lan85, ch. II, § 2] shows more intuitively the vector space structure, and how the notion of tangent vectors is related to the atlas of a manifold which defines its geometry.

Definition 7.1.3 (Tangent vector)**.** Let M be a manifold of class C^k with $k \geq 1$ of dimension n and $\mathcal{A}$ its atlas, as well as $p \in M$. A *tangent vector* is an equivalence class $[U, \phi, u]$ of triples (U, ϕ, u), where $(U, \phi) \in \mathcal{A}$ is a chart with $p \in U$ and $u \in \mathbb{R}^n$, where two triples (U, ϕ, u) and (V, ψ, v) are regarded equivalent if and only if

$$\mathrm{D}(\psi \circ \phi^{-1})_{\phi(p)}(u) = v , \tag{7.1.17}$$

where D denotes the Jacobian.

In this definition it is less obvious that the tangent vector is an object which is intrinsic to the manifold M and "attached" to p. We can relate this definition to the previous one as follows:

Theorem 7.1.4. *Let M be a manifold of class C^k and dimension $\dim M = n$ with $k \geq 1$ and $p \in M$. There exists a one-to-one correspondence between tangent vectors at p and derivations at p.*

Proof. In theorem 7.1.3 we have seen that for each $p \in M$ a chart (U, ϕ) with $p \in U$ induces a linear bijection between T_pM and $\mathbb{R}^n$, which assigns to $D \in T_pM$ the element $u \in \mathbb{R}^n$ defined by the relation (7.1.16). Let $[U, \phi, u]$ be the tangent vector defined by the representative triple (U, ϕ, u). Another triple (V, ψ, v) defines the same tangent vector $[V, \psi, v] = [U, \phi, u]$ if and only if

$$v^a = u^b \left.\frac{\partial}{\partial x^b}\psi^a(\phi^{-1}(x))\right|_{x=\phi(p)} . \tag{7.1.18}$$

Note that D acts on a function $f \in C^\infty(M, \mathbb{R})$ as

$$\begin{aligned} D(f) &= u^a \frac{\partial}{\partial x^a}(f \circ \phi^{-1})(x)\bigg|_{x=\phi(p)} \\ &= u^a \frac{\partial}{\partial x^a}(f \circ \psi^{-1} \circ \psi \circ \phi^{-1})(x)\bigg|_{x=\phi(p)} \\ &= u^a \frac{\partial}{\partial x^a}\psi^b(\phi^{-1}(x))\bigg|_{x=\phi(p)} \frac{\partial}{\partial y^b}(f \circ \psi^{-1})(y)\bigg|_{y=\psi(p)} \\ &= v^b \frac{\partial}{\partial y^b}(f \circ \psi^{-1})(y)\bigg|_{y=\psi(p)}, \end{aligned} \tag{7.1.19}$$

where the last equality holds if and only if the equality (7.1.18) holds. Hence, $[V, \psi, v] = [U, \phi, u]$ if and only if (U, ϕ, u) and (V, ψ, v) define the same derivation $D \in T_pM$. ■

Making use of this one-to-one correspondence, we will therefore use the words derivation and tangent vector interchangeably, since they denote equivalent objects. Using their representation as vectors $u = u^a \mathbb{e}_a \in \mathbb{R}^n$ defined by a chart, we can also pictorially visualize them as arrows, whose length and direction represents its components, as shown in figure 7.1. Also the vector space operations are straightforward to visualize using the arrow representation. Figure 7.2 shows the sum of two tangent vectors, while figure 7.3 shows the multiplication by a scalar.

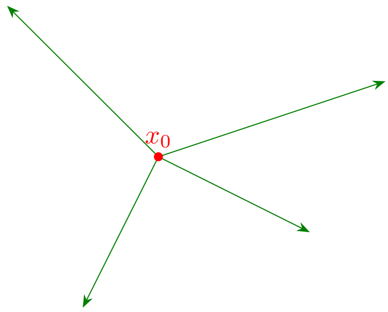


Figure 7.1: Visualization of tangent vectors as arrows.

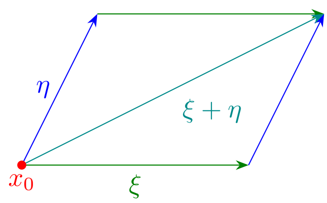


Figure 7.2: Visualization of the sum of two tangent vectors.

Figure 7.3: Visualization of scalar multiples of a tangent vector.

7.2 Tangent bundle

We have now defined an n-dimensional vector space at each point $p \in M$. It now becomes natural to consider these vector spaces to be fibers of a vector bundle over M. We define this vector bundle as follows.

Definition 7.2.1 (Tangent bundle). The *tangent bundle* of a manifold M of dimension n is the vector bundle $(TM, M, \tau, \mathbb{R}^n)$, whose total space is the disjoint union

$$TM = \biguplus_{p \in M} T_pM\,, \tag{7.2.1}$$

and the projection is the function $\tau : TM \to M$ such that $\tau(\xi) = p$ for $\xi \in T_pM$.

It is important to note that we take the *disjoint* union of all tangent spaces, i.e., we consider elements of TM to be different if they are taken from the tangent spaces T_pM and T_qM at different points $p \neq q$. For example, the function $D : C^\infty(M, \mathbb{R}) \to \mathbb{R}, f \mapsto 0$ is obviously a derivation both at p and q, and it would be contained only once in TM if we would naively take the union of all tangent spaces as defined at the beginning of this section. However, by taking the disjoint union, the elements of TM are actually pairs (p, D) of a point $p \in M$ and a derivation $D \in T_pM$, such that $(p, D) \neq (q, D)$. The projection τ is then simply the function $\tau : (p, D) \mapsto p$.

In order to constitute a vector bundle, the total space TM must, of course, carry the structure of a manifold. Also in order to be uniquely defined, this manifold structure must follow from the manifold structure on M. Given an atlas $\mathcal{A}$ on M, we can construct an atlas on TM as follows.

Let $p \in M$ and $(U, \phi) \in \mathcal{A}$ be a chart in an atlas $\mathcal{A}$ of M such that $p \in U$. Define $\tilde{U} = \tau^{-1}(U) \subset TM$ and let $\tilde{\phi} : \tilde{U} \to \mathbb{R}^{2n}$ the function that assigns to $(p, D) \in \tilde{U}$ the pair consisting of $\phi(p)$ and $u \in \mathbb{R}^n$ such that the relation (7.1.16) holds. Apply this procedure to every chart in $\mathcal{A}$. One easily checks that this yields an atlas $\tilde{\mathcal{A}}$ of TM.

Finally, we also need to show that $(TM, M, \tau, \mathbb{R}^n)$ is indeed a vector bundle of rank $n = \dim M$. This can be done as follows.

For $p \in M$, let (U, ϕ) be a chart of M such that $p \in U$. Define a function $\hat{\phi} : \tau^{-1}(U) \to U \times \mathbb{R}^n$ such that $\hat{\phi}(p, D) = (p, u)$ and $u \in \mathbb{R}^n$ as in (7.1.16). One can easily show that $\hat{\phi}$ has the properties listed in the definition 3.1.1 of a vector bundle.

The charts $(\tilde{U}, \tilde{\phi})$ on TM defined above have another nice property. Given two tangent vectors (p, D_1) and (p, D_2) at the same point $p \in M$ and constants λ_1 and λ_2, they satisfy

$$\tilde{\phi}(p, \lambda_1 D_1 + \lambda_2 D_2) = (\phi(p), \lambda_1 u_1 + \lambda_2 u_2)\,, \tag{7.2.2}$$

i.e., they are linear in the fiber coordinates. Given the canonical basis $(\mathbb{e}_a, a = 1, \ldots, n)$ they therefore define a basis

$$\partial_a = \tilde{\phi}^{-1}(\phi(p), \mathbb{e}_a) \in T_pM\,, \quad a = 1, \ldots, n \tag{7.2.3}$$

for each tangent space T_pM, which we call the *coordinate basis* induced by the chart (U, ϕ) on M. In most cases these are the most convenient coordinates on TM. If the point $p \in M$ is fixed, they allow to write a tangent vector $\xi \in T_pM$ in the form $\xi = \xi^a \partial_a$, where we also use the *Einstein summation convention* that the occurrence of an upper and a lower index implies a sum over all values that this index takes. However, note that this notation "hides" the chart (U, ϕ) - it is implicit in the (chart-dependent) basis ∂_a. This becomes clear if we recall that ξ

defines a derivation on functions, which is now expressed as

$$\xi(f) = \xi^a \partial_a f = \sum_{a=1}^{n} \xi^a \left. \frac{\partial(f \circ \phi^{-1})(x)}{\partial x^a} \right|_{x=\phi(p)} . \tag{7.2.4}$$

The right hand side of this equation shows that the basis vectors ∂_a, and therefore the components ξ^a, depend on the choice of the chart, while the total expression $\xi(f)$ is independent of this choice.

7.3 Tangent vectors of curves

We now come to more practical aspects of the tangent bundle, which are closer to physics. One of the most important aspects is the tangent vector of a curve, which can be interpreted as the velocity of a point mass along its trajectory and which is defined as follows.

Definition 7.3.1 (Tangent vector of a curve)**.** Let $\gamma \in C^\infty(\mathbb{R}, M)$ be a curve on a manifold M. Its *tangent vector* at $t \in \mathbb{R}$ is the derivation $\dot{\gamma}(t) \in T_{\gamma(t)}M$ defined by

$$\dot{\gamma}(t)(f) = (f \circ \gamma)'(t) \tag{7.3.1}$$

for $f \in C^\infty(M, \mathbb{R})$.

Frequently we need to consider a reparametrization of a curve, i.e., a change of the curve parameter. This can be seen as a special case of defining a curve $\tilde{\gamma} = \gamma \circ u$ with $u \in C^\infty(\mathbb{R}, \mathbb{R})$. We find that the following holds.

Theorem 7.3.1. *Let $\gamma \in C^\infty(\mathbb{R}, M)$ be a curve on a manifold M and $u \in C^\infty(\mathbb{R}, \mathbb{R})$. For the curve $\tilde{\gamma} = \gamma \circ u$ holds $\dot{\tilde{\gamma}}(t) = u'(t)\dot{\gamma}(u(t))$ for all $t \in \mathbb{R}$.*

Proof. Using the fact that for any $f \in C^\infty(M,)$ the composition $f \circ \gamma \in C^\infty(\mathbb{R}, \mathbb{R})$, we can use the well-known chain rule to calculate

$$\dot{\tilde{\gamma}}(t)(f) = (f \circ \gamma \circ u)'(t) = u'(t)(f \circ \gamma)'(u(t)) = u'(t)\dot{\gamma}(u(t))(f) . \tag{7.3.2}$$

■

Another possibility to obtain a new curve from a known one is the composition $\varphi \circ \gamma \in C^\infty(\mathbb{R}, N)$ for $\gamma \in C^\infty(\mathbb{R}, M)$ and $\varphi \in C^\infty(M, N)$. Its tangent vectors will be discussed in section 10.2.

We are also interested in a coordinate description, so we will work in local coordinates (x^a) defined by a chart (U, ϕ), such that $\gamma(t) \in U$. For a curve $\gamma \in C^\infty(\mathbb{R}, M)$ we then obtain the coordinate expression $t \mapsto (\phi \circ \gamma)(t)$, which assigns to each $t \in \mathbb{R}$ a point $x = (x^a, a = 1, \ldots, n) \in \mathbb{R}^n$. Using the same coordinate chart, a function $f \in C^\infty(M, \mathbb{R})$ is expressed as $x \mapsto (f \circ \phi^{-1})(x)$. The composition $f \circ \gamma : \mathbb{R} \to \mathbb{R}$ is thus expressed as $f \circ \phi^{-1} \circ \phi \circ \gamma$. By the chain rule we then have

$$\begin{aligned} \dot{\gamma}(t_0)(f) &= (f \circ \gamma)'(t_0) \\ &= (f \circ \phi^{-1} \circ \phi \circ \gamma)'(t_0) \\ &= \sum_{a=1}^{n} \left. \frac{\partial(f \circ \phi^{-1})(x)}{\partial x^a} \right|_{x=\gamma(t_0)} \cdot \left. \frac{\partial(\phi \circ \gamma)^a(t)}{\partial t} \right|_{t=t_0} . \end{aligned} \tag{7.3.3}$$

Of course this is a rather lengthy and cumbersome notation, and so one usually uses a shorter notation, in particular in the physics literature. Recall that we made use of the coordinate basis (7.2.3) of the tangent space in order to express the application (7.2.4) of a tangent vector to a function. Using the coordinate basis ∂_a of $T_{\gamma(t)}M$, one conveniently expresses the tangent vector as $\dot{\gamma}(t) = \dot{\gamma}^a(t)\partial_a$ in components $\dot{\gamma}^a(t)$. Further, one usually omits the map ϕ in the coordinate expression of the curve, and writes $\gamma^a(t)$ instead of $(\phi \circ \gamma)^a(t)$. This notation is not only shorter, but also has the convenient property that

$$\dot{\gamma}^a(t) = \frac{\mathrm{d}}{\mathrm{d}t}\gamma^a(t) \,. \tag{7.3.4}$$

Hence, written in components the tangent vector is nothing but the ordinary derivative with respect to the curve parameter t. However, one must still keep in mind that both sides of this equation depend on the choice of the chart (although the equality holds in any chart).

We conclude this section with another helpful construction which we will use later. Note that for each $t \in \mathbb{R}$ the tangent vector $\dot{\gamma}(t)$ of a curve $\gamma \in C^\infty(\mathbb{R}, M)$ on M is an element of $T_{\gamma(t)}M$. Hence, the curve γ defines a function $t \mapsto \dot{\gamma}(t)$, which deserves its own name.

Definition 7.3.2 (Canonical lift of a curve)**.** Let $\gamma \in C^\infty(\mathbb{R}, M)$ be a curve on a manifold M. Its *canonical lift* is the curve $\dot{\gamma} : t \mapsto \dot{\gamma}(t)$ on TM.

Using the induced charts on TM it is not difficult to check that the canonical lift is indeed an element of $C^\infty(\mathbb{R}, TM)$. Also it follows immediately that $\tau \circ \dot{\gamma} = \gamma$.

7.4 Vector fields

If we consider a fluid instead of a point mass, we have a velocity at each point of the fluid, so we need to assign a tangent space element to every point. We already encountered this type of assignment and called it a section. Sections of the tangent bundle are so important that they deserve their own name.

Definition 7.4.1 (Vector field)**.** A *vector field* on a manifold M is a section of the tangent bundle TM. The space of all vector fields on M is denoted $\Gamma(TM)$ or $\mathrm{Vect}(M)$.

Let $X \in \mathrm{Vect}(M)$. If we use local coordinates (x^a) defined by a chart (U, ϕ) on M, which further induces a chart $(\tilde{U}, \tilde{\phi})$ on TM with coordinates (x^a, v^a), then the coordinate expression

$$\tilde{\phi} \circ X \circ \phi^{-1} : \phi(U) \to \mathbb{R}^{2n} \tag{7.4.1}$$

assigns to each $x \in \phi(U)$ a pair (x, v). Writing $v = v^a \mathbb{e}_a$ in the canonical basis of $\mathbb{R}^n$, and using the coordinate basis ∂_a of T_pM with $\phi(p) = x$, we can write $X(p) = X^a(p)\partial_a$. Extending this notation, we can write the whole vector field as $X = X^a\partial_a$. Note that when evaluating the vector field at a point p to obtain $X(p)$, one has to take both the coordinate expression X^a and the basis vector ∂_a at this point p.

Since a vector field assigns to any point of a manifold a derivation at that point, we can define the following construction.

Definition 7.4.2 (Action of a vector field on a function)**.** Let M be a manifold, $X \in \mathrm{Vect}(M)$ a vector field on M and $f \in C^\infty(M, \mathbb{R})$ a real function on M. For each $p \in M$, the vector field X defines a derivation $X(p) \in T_pM$. Via these derivations X *acts* on f, i.e., it defines a real function $Xf \in C^\infty(M, \mathbb{R})$ given by

$$(Xf)(p) = X(p)(f) \tag{7.4.2}$$

for all $p \in M$.

We illustrate this definition using local coordinates (x^a). In these coordinates the vector field X takes the form $X^a\partial_a$, and $Xf = X^a\partial_a f$, which should be read in the obvious way:

$$(Xf)(p) = X^a(p)(\partial_a f)(p)\,. \tag{7.4.3}$$

Some more properties follow from the definition:

- Xf is $\mathbb{R}$-linear in the first argument:

$$(\lambda X + \mu Y)f = \lambda(Xf) + \mu(Yf) \quad \text{for} \quad \lambda, \mu \in \mathbb{R}\,. \tag{7.4.4}$$

- Xf is $\mathbb{R}$-linear in the second argument:

$$X(\lambda f + \mu g) = \lambda(Xf) + \mu(Xg) \quad \text{for} \quad \lambda, \mu \in \mathbb{R}\,. \tag{7.4.5}$$

- Xf satisfies the Leibniz rule for the second argument:

$$X(fg) = (Xf)g + f(Xg)\,. \tag{7.4.6}$$

7.5 Commutator of vector fields

The action of vector fields on functions, which form a vector space, gives us a hint that the set $\mathrm{Vect}(M)$ can be equipped with more structure, turning it into an algebra. We define this structure as follows.

Definition 7.5.1 (Commutator of vector fields)**.** Let M be a manifold and $X, Y \in \mathrm{Vect}(M)$ vector fields. Their *commutator* is the unique vector field $[X, Y] \in \mathrm{Vect}(M)$ such that for all $f \in C^\infty(M, \mathbb{R})$,

$$[X, Y]f = X(Yf) - Y(Xf)\,. \tag{7.5.1}$$

Of course one must check that such a unique vector field $[X, Y]$ really exists, i.e., that the definition above assigns to each point $p \in M$ an element in T_pM, i.e., a derivation at p. It is clear from the definition above that

$$\begin{array}{rcccc} [X, Y](p) & : & C^\infty(M, \mathbb{R}) & \to & \mathbb{R} \\ & & f & \mapsto & ([X, Y]f)(p) \end{array} \tag{7.5.2}$$

is a linear function. To see that it is a derivation, one calculates

$$\begin{aligned} [X, Y](fg) &= X(Y(fg)) - Y(X(fg)) \\ &= X((Yf)g + f(Yg)) - Y((Xf)g - f(Xg)) \\ &= (X(Yf))g + (Yf)(Xg) + (Xf)(Yg) + f(X(Yg)) \\ &\quad - (Y(Xf))g - (Xf)(Yg) - (Yf)(Xg) - f(Y(Xg)) \\ &= ([X, Y]f)g + f([X, Y]g)\,. \end{aligned} \tag{7.5.3}$$

Evaluating this at p we find

$$([X,Y](fg))(p) = ([X,Y]f)(p)g(p) + f(p)([X,Y]g)(p)\,. \tag{7.5.4}$$

This is exactly the Leibniz rule (7.1.1) for a derivation $D = [X,Y](p)$.

If we express the vector fields $X = X^a\partial_a$ and $Y = Y^a\partial_a$ in a coordinate basis, we can use the product rule for derivatives on $\mathbb{R}^n$ to see that

$$\begin{aligned} [X,Y]f &= X^a\partial_a(Y^b\partial_b f) - Y^a\partial_a(X^b\partial_b f) \\ &= (X^a\partial_a Y^b - Y^a\partial_a X^b)\partial_b f \\ &= [X,Y]^b\partial_b f\,. \end{aligned} \tag{7.5.5}$$

This gives us an explicit formula for the components $[X,Y]^a$ in these coordinates. This formula has the same form in all coordinate systems, since we have made no reference to particular coordinates in the definition of $[X,Y]$. Another property of the commutator is now easy to show.

Theorem 7.5.1. *The set* $\text{Vect}(M)$ *of vector fields on a manifold* M *carries the structure of a real Lie algebra, with the Lie bracket given by the commutator.*

Proof. It follows from the linearity of Xf in the first argument that $\text{Vect}(M)$ is a real vector space. The same property implies that the commutator $[X,Y]$ is linear in both arguments. Further, one can see immediately from the definition that it is antisymmetric. Finally, we check the Jacobi identity

$$\begin{aligned} 0 &\overset{?}{=} [X,[Y,Z]]f + [Y,[Z,X]]f + [Z,[X,Y]]f \\ &= X(Y(Zf)) - X(Z(Yf)) - Y(Z(Xf)) + Z(Y(Xf)) + Y(Z(Xf)) - Y(X(Zf)) \\ &\quad - Z(X(Yf)) + X(Z(Yf)) + Z(X(Yf)) - Z(Y(Xf)) - X(Y(Zf)) + Y(X(Zf))\,. \end{aligned} \tag{7.5.6}$$

One easily checks that this indeed vanishes. ∎

Here we have used the term Lie algebra in the usual sense as an algebra whose product (the Lie bracket) is antisymmetric and satisfies the Jacobi identity. We will see later that there is a much deeper connection between vector fields and Lie algebras when we come to the discussion of Lie groups in chapter 15. Another useful formula is the following.

Theorem 7.5.2. *The commutator of vector fields satisfies the Leibniz rule*

$$[X, fY] = (Xf)Y + f[X,Y] \tag{7.5.7}$$

for all $X, Y \in \text{Vect}(M)$ *and* $f \in C^\infty(M,\mathbb{R})$.

Proof. Acting with $[X, fY]$ on another function $g \in C^\infty(M,\mathbb{R})$ we find

$$\begin{aligned} [X,fY]g &= X(f(Yg)) - fY(Xg) \\ &= (Xf)(Yg) + fX(Yg) - fY(Xg) \\ &= ((Xf)Y + f[X,Y])g\,. \quad ∎ \end{aligned} \tag{7.5.8}$$

7.6 Distributions

Chapter 8

Cotangent bundle and covector fields

8.1 Ideals of functions and cotangent spaces

We now come to a concept which is somehow dual to the tangent bundle. While we have defined elements of the tangent bundle as derivations, which act on functions $f \in C^\infty(M, \mathbb{R})$, elements of the cotangent bundle can be defined as equivalence classes of functions, where two functions are regarded equivalent if and only if they yield the same value if we act on them with a derivation. This will be formalized in our definition of the *cotangent space*.

Definition 8.1.1 (Cotangent space)**.** Let M be a manifold and $p \in M$. Let $I_p \subset C^\infty(M, \mathbb{R})$ be the ideal of real functions f on M for which $f(p) = 0$ and $I_p^2 \subset I_p$ the ideal generated by functions fg with $f, g \in I_p$. Both I_p and I_p^2 are vector spaces, and I_p^2 is a subspace of I_p. The *cotangent space* T_p^*M at p is the quotient vector space I_p/I_p^2.

Recall that an element of the quotient vector space I_p/I_p^2 is defined as the equivalence class

$$[f]_p = \{g \in I_p, f - g \in I_p^2\}, \tag{8.1.1}$$

i.e., two functions $f, g \in I_p$ belong to the same equivalence class, $[f]_p = [g]_p$, if and only if $f - g \in I_p^2$. We call f a *representative* of the class $[f]_p$. As it was also the case for the tangent space, it is not obvious whether T_p^*M is of finite dimension. In fact, the following holds, and this time we will prove it using Hadamard's lemma.

Theorem 8.1.1. *Let M be a manifold of dimension $\dim M = n$ and $p \in M$. The cotangent space T_p^*M is a vector space of dimension n.*

Proof. Let (U, ϕ) be a chart such that $p \in U$. Since $\phi(U)$ is open, we can find $\epsilon > 0$ such that

$$V = \{x \in \mathbb{R}^n, \|x - \phi(p)\| < \epsilon\} \subset \phi(U). \tag{8.1.2}$$

Now consider a function $f \in I_p$, and write

$$\begin{array}{rccc} F & : & V & \to & \mathbb{R} \\ & & x & \mapsto & (f \circ \phi^{-1})(x) \end{array}. \tag{8.1.3}$$

Note that $F \in C^\infty(V, \mathbb{R})$. Using Hadamard's lemma we can write any smooth function F on V in the form

$$F(x) = F(x_0) + (x^a - x_0^a)\tilde{F}_a(x)\,, \tag{8.1.4}$$

where $x_0 = \phi(p)$ and $\tilde{F}_a$ are smooth functions on V. Since we have chosen $f \in I_p$, it follows that

$$F(x_0) = (f \circ \phi^{-1})(\phi(p)) = f(p) = 0\,, \tag{8.1.5}$$

and therefore

$$F(x) = (x^a - x_0^a)\tilde{F}_a(x)\,. \tag{8.1.6}$$

We now want to determine which functions $f \in I_p$ belong to I_p^2. For this purpose, let us first recall that every component ϕ^a of a chart $\phi : U \to \mathbb{R}^n$ defines a smooth function $\phi^a : U \to \mathbb{R}$ with

$$\begin{array}{rcccl} \phi^a \circ \phi^{-1} & : & V & \to & \mathbb{R} \\ & & x & \mapsto & x^a \end{array}\,, \tag{8.1.7}$$

and let us define

$$\begin{array}{rcccl} \phi_0^a & : & \phi^{-1}(V) & \to & \mathbb{R} \\ & & q & \mapsto & \phi^a(q) - \phi^a(p) = \phi^a(q) - x_0^a \end{array}\,. \tag{8.1.8}$$

Clearly, we have $\phi_0^a(p) = 0$ and thus $\phi_0^a \in I_p$. Let us now first assume that $\tilde{F}_a(x_0) = 0$, and define

$$\begin{array}{rcccl} \tilde{f}_a = \tilde{F}_a \circ \phi & : & \phi^{-1}(V) & \to & \mathbb{R} \\ & & q & \mapsto & \tilde{F}_a(\phi(q)) \end{array}\,. \tag{8.1.9}$$

Assuming $\tilde{F}_a(x_0) = 0$, we have $\tilde{f}_a(p) = 0$, and thus $\tilde{f}_a \in I_p$. It then follows that

$$f(q) = F(\phi(q)) = \phi_0^a(q)\tilde{F}_a(\phi(q)) = \phi_0^a(q)\tilde{f}_a(q)\,, \tag{8.1.10}$$

and so f is a sum of products of functions $\phi_0^a, \tilde{f}_a \in I_p$, and thus $f \in I_p^2$. Conversely, let us assume that $f = gh$ with functions $g, h \in I_p$, and define $G = g \circ \phi^{-1}$ and $H = h \circ \phi^{-1}$ as above. Then we have

$$F(x) = G(x)H(x) = (x^a - x_0^a)(x^b - x_0^b)\tilde{G}_a(x)\tilde{H}_b(x) = (x^a - x_0^a)\tilde{F}_a(x)\,, \tag{8.1.11}$$

and thus $\tilde{F}_a(x_0) = 0$. Hence, we have shown that $f \in I_p^2$ if and only if $\tilde{F}_a(x_0) = 0$. Since the expansion (8.1.6) is linear in $\tilde{F}_a$, it further follows that for any two functions $f, f' \in I_p$ we have $f - f' \in I_p^2$, and thus $[f]_p = [f']_p$, if and only if $\tilde{F}_a(x_0) = \tilde{F}'_a(x_0)$. Hence, the equivalence class $[f]_p$ is uniquely determined by $\tilde{F}_a(x_0) \in \mathbb{R}^n$. In particular, choosing

$$\begin{array}{rcccl} f' & : & U & \to & \mathbb{R} \\ & & q & \mapsto & \tilde{f}_a(p)\phi_0^a(q) = \tilde{F}_a(x_0)\phi_0^a(q) \end{array}\,, \tag{8.1.12}$$

we have $[f]_p = [f']_p$. We can thus write every equivalence class $[f]_p$ uniquely as

$$[f]_p = [f']_p = [\tilde{F}_a(x_0)\phi_0^a]_p = \tilde{F}_a(x_0)[\phi_0^a]_p = \tilde{f}_a(p)[\phi_0^a]_p\,, \tag{8.1.13}$$

in terms of constant coefficients $\tilde{f}_a(p)$ with respect to a basis $([\phi_0^a]_p)$. Hence, $T_p^*M = I_p/I_p^2$ is a vector space of dimension $\dim M = n$. ■

It is conventional to denote the coordinate basis elements $[\phi_0^a]_p$ of T_p^*M introduced above by $\mathrm{d}x^a$. Note that both the components $\tilde{F}_a(x_0)$ and the basis vectors $\mathrm{d}x^a$ depend on the choice of the chart (U, ϕ). This leaves us with the question how to calculate the components $\tilde{F}_a(x_0)$ for a given function f. Using the expansion (8.1.6) one easily checks that

$$\begin{aligned} \left.\frac{\partial}{\partial x^a}(f \circ \phi^{-1})(x)\right|_{x=\phi(p)} &= \left.\frac{\partial}{\partial x^a}F(x)\right|_{x=x_0} \\ &= \left.\frac{\partial}{\partial x^a}\left[(x^b - x_0^b)\tilde{F}_b(x)\right]\right|_{x=x_0} \\ &= \tilde{F}_a(x_0)\,. \end{aligned} \tag{8.1.14}$$

Hence, the components $\tilde{F}_a(x_0)$ are simply the directional derivatives $\partial_a f$ we introduced when we discussed the tangent bundle, and so we can write

$$[f]_p = \partial_a f \, \mathrm{d}x^a \, , \tag{8.1.15}$$

where it is understood that ∂_a is the coordinate basis of T_pM. We will see later that this coincidence is a consequence of the relation between the tangent and cotangent bundles.

The aforementioned construction also offers a possibility to pictorially visualize cotangent vectors. Note that within a chart a cotangent vector defines a linear function

$$\hat{F}(x) = (x^a - x_0^a)\tilde{F}_a(x_0) = (x^a - x_0^a) \left. \frac{\partial}{\partial x^a} F(x) \right|_{x=x_0} , \tag{8.1.16}$$

which is the unique linear function with $\hat{F}(x_0) = 0$, whose partial derivatives at x_0 agree with that of F, and vice versa. The level sets of this function are parallel hyperplanes (i.e., spaces of codimension 1), with the hyperplane passing through x_0 representing the value 0, and hyperplanes are more dense if the magnitude of the cotangent vector is larger. One may therefore visualize a cotangent vector as a stack of hyperplanes, with their orientation and density (inverse distance) representing orientation and magnitude of the covector. This is shown in figure 8.1. In fact, one may also draw only the hyperplanes $\hat{F}(x) = 0$ and $\hat{F}(x) = 1$ and omit all others, since their location follows by linearity; this is used in [Bur85, HO01, Jan21]. Another possible way to depict the magnitude is to encode it in the size of the drawn hyperplane element; however, this is less intuitive in terms of its geometric interpretation.

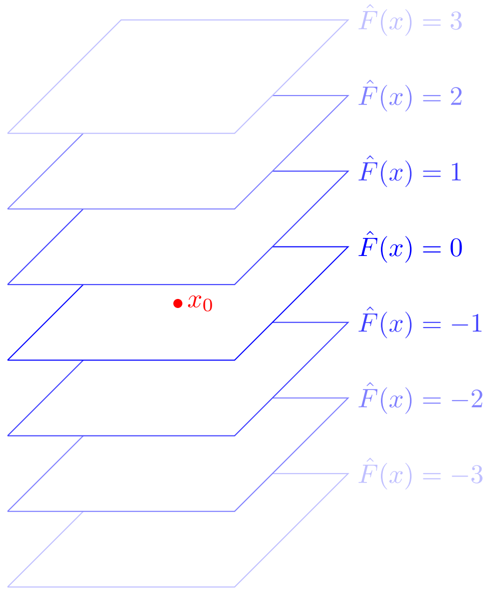


Figure 8.1: Visualization of a cotangent vector as a stack of hyperplanes. In this picture the chart (and hence the manifold) has dimension 3, and so the hyperplanes have are actual planes of dimension 2.

Using the visualization of cotangent vectors as hyperplanes defined by the level sets of a linear function $\hat{F}$, we can also easily visualize the vector space structure of the cotangent space. If two covectors are represented by linear functions $\hat{F}$ and $\hat{G}$, their sum is represented by $\hat{F}+\hat{G}$; its level sets are shown in figure 8.2. Similarly, we can visualize the scalar multiplication. Figure 8.3 shows the multiplication by 2. Note that this moves the hyperplanes twice as close to each other, since the magnitude of a covector is encoded in the *density*, i.e., the *inverse distance* of the hyperplanes. Finally, figure 8.4 shows the negative of a covector, which simply reverses the stack of hyperplanes.

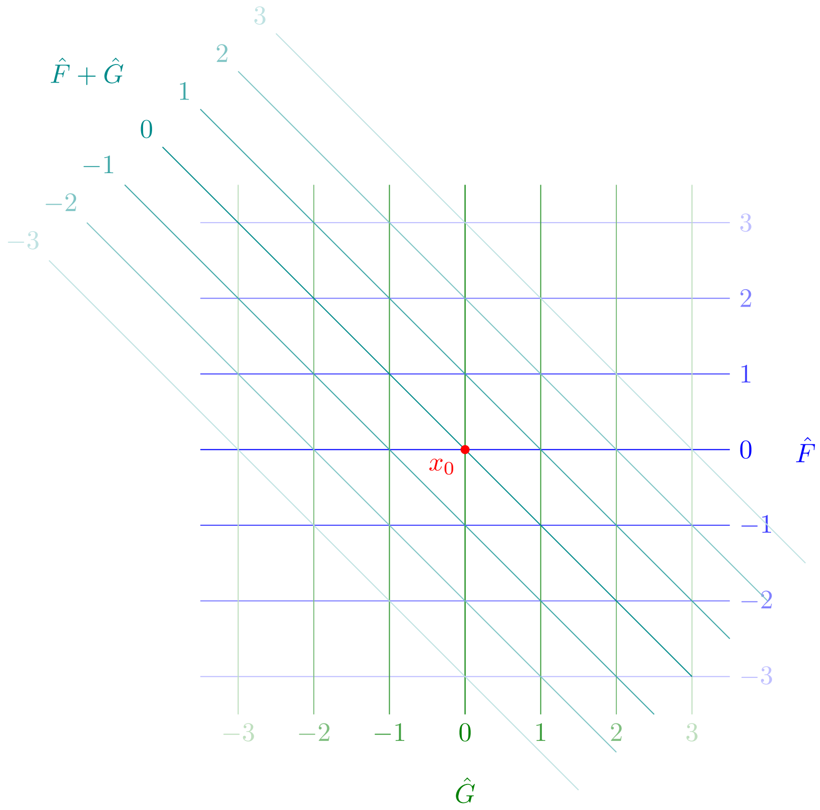


Figure 8.2: Visualization of the sum of two covectors, represented by the level sets (hyperplanes) of the linear functions $\hat{F}$ and $\hat{G}$. Each level set of $\hat{F} + \hat{G}$ passes through the intersections of level sets of $\hat{F}$ and $\hat{G}$ whose sum equals the value of $\hat{F} + \hat{G}$.

8.2 Cotangent bundle

We can proceed similarly to the construction of the tangent bundle and assemble the cotangent spaces to form the cotangent bundle.

Definition 8.2.1 (Cotangent bundle)**.** The *cotangent bundle* of a manifold M of dimension n is the vector bundle $(T^*M, M, \bar{\tau}, \mathbb{R}^n)$, whose total space is the disjoint union

$$T^*M = \biguplus_{p \in M} T_p^*M \,, \tag{8.2.1}$$

and the projection is the function $\bar{\tau} : T^*M \to M$ such that $\bar{\tau}(\alpha) = p$ for $\alpha \in T_p^*M$.

Also here we take the disjoint union, in full analogy to the construction of the tangent bundle, even though this is not really necessary here: two arbitrary covectors $[f]_p$ and $[g]_q$ are always distinct for distinct points $p \neq q$. Of course one still has to provide an atlas on T^*M in order to turn it into a manifold, and construct the local trivializations in order to show that it is indeed a vector bundle of rank n over M. We will not prove this here, since the construction proceeds in full analogy to the tangent bundle, but now using the coordinate basis $\mathrm{d}x^a$ instead of ∂_a.

We finally remark that in contrast to the tangent bundle TM, where the coordinate basis (∂_a)

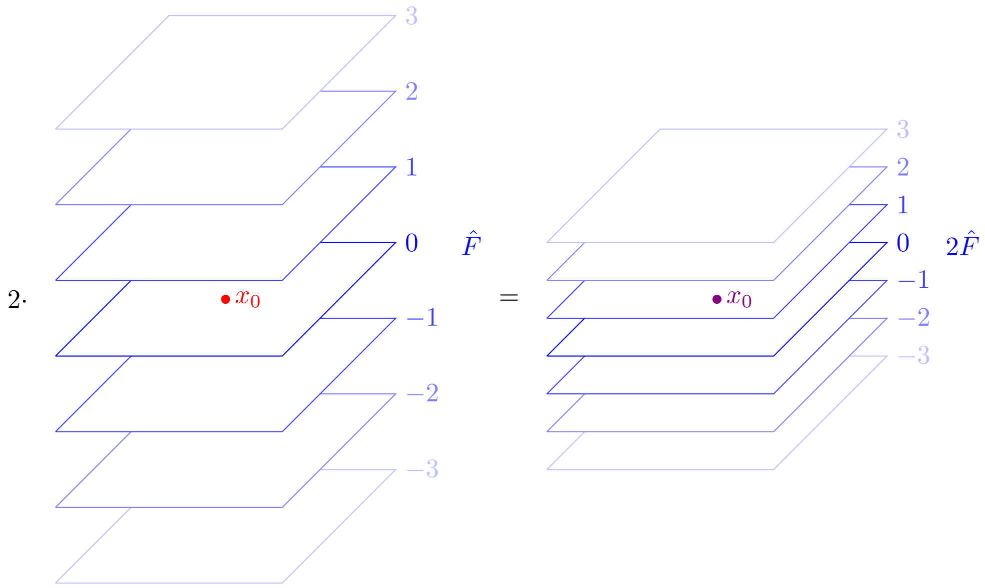


Figure 8.3: Visualization of the multiplication of a cotangent vector by a scalar. In place of level set to the value $\hat{F}(x) = 1$ we find that $2\hat{F}$ has the value $2\hat{F}(x) = 2$. We thus find that the level sets (hyperplanes) of $2\hat{F}$ are twice as dense as those of $\hat{F}$.

carries a lower index and we introduced coordinates (x^a, v^a), the coordinate basis $(\mathrm{d}x^a)$on T^*M carries an upper index, and so one would introduce coordinates (x^a, p_a) on T^*M and write a cotangent vector as $p_a \, \mathrm{d}x^a$, making use of the Einstein summation convention. Again, this is not by accident, or a matter of inconvenience, but simply a consequence of the fact that tangent and cotangent bundles are duals.

8.3 Covector fields

Sections of the cotangent bundle are of similar importance as sections of the tangent bundle, and also deserve their own name. Hence, we introduce the following notion.

Definition 8.3.1 (Covector field)**.** A *covector field* (or *1-form*) on a manifold M is a section of the cotangent bundle T^*M. The space of all covector fields on M is denoted $\Gamma(T^*M)$ or $\Omega^1(M)$.

The term one-form and the notation $\Omega^1(M)$ will become clear in the next chapter, when we discuss general k-forms, with $0 \le k \le \dim M$. As it was also the case with vector fields, we can use coordinates (x^a) to write a covector field in the form $\omega = \omega_a \, \mathrm{d}x^a$, where the component functions ω_a are smooth. Note that $\mathrm{d}x^a$ now denotes a *covector field*, while in the previous section it was used to denote a single *covector* at a fixed point $p \in M$. We also encountered this ambiguity of notation with the symbol ∂_a, and so one must be careful whenever it is being used.

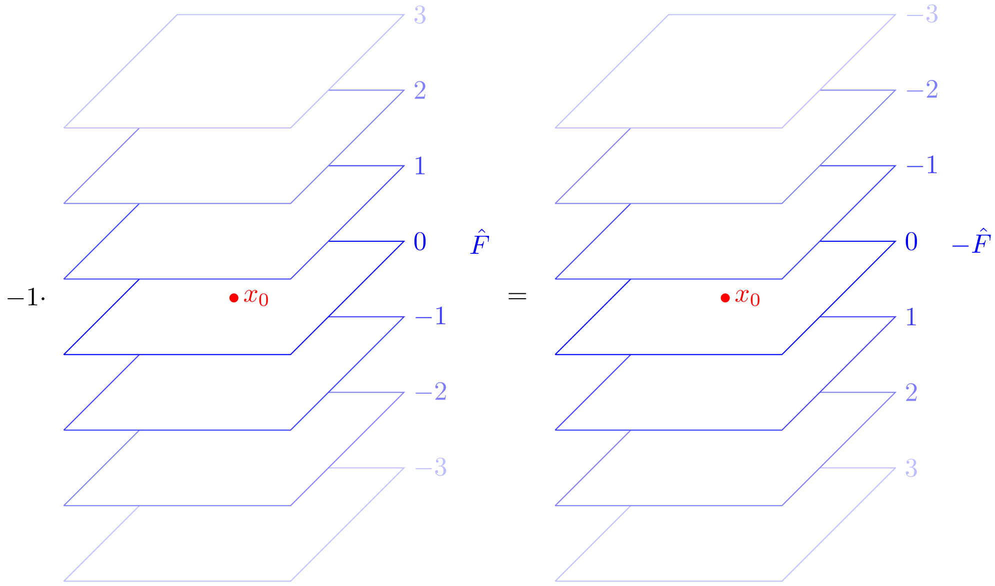


Figure 8.4: Visualization of the negative of a cotangent vector by a scalar, which reverses the stack of hyperplanes.

8.4 Total differential

There is a particular class of covector fields which are obtained from real functions by an operation which we define as follows.

Definition 8.4.1 (Total differential)**.** Let M be a manifold and $f \in C^\infty(M, \mathbb{R})$ a function on M. Its *total differential* is the covector field

$$\begin{array}{rcccc} \mathrm{d}f & : & M & \to & T^*M \\ & & p & \mapsto & [f - f(p)]_p \end{array} \tag{8.4.1}$$

which assigns to each point $p \in M$ the equivalence class $[f - f(p)]_p \in T_p^*M$ of $f - f(p) \in I_p$ modulo I_p^2.

To clarify this definition, recall that we defined the tangent space T_p^*M at $p \in M$ as the quotient space I_p/I_p^2, where I_p is the ideal of functions which vanish at p. We can therefore specify an element of T_p^*M by providing an element of I_p, i.e., by a representative. Note that in general a function $f \in C^\infty(M, \mathbb{R})$ will not be in I_p, but the function $\tilde{f} : q \mapsto f(q) - f(p)$ satisfies $\tilde{f} = f - f(p) \in I_p$. This is the function we choose as the representative. Of course one still has to prove that $\mathrm{d}f$ is indeed a (smooth) section, but we will omit the formal proof here, and instead simply state the following.

Theorem 8.4.1. *The total differential is a linear function* $\mathrm{d} : C^\infty(M, \mathbb{R}) \to \Omega^1(M)$ *which satisfies the Leibniz rule,*

$$\mathrm{d}(fg) = \mathrm{d}f\, g + f\, \mathrm{d}g \tag{8.4.2}$$

for $f, g \in C^\infty(M, \mathbb{R})$.

Proof. We restrict ourselves to proving linearity and the Leibniz rule here, and omit the proof of smoothness. Let $f, g \in C^\infty(M, \mathbb{R})$ and $\mu, \nu \in \mathbb{R}$, and define $h = \mu f + \nu g$. For $p \in M$ we then have

$$
\begin{aligned}
\mathrm{d}h(p) &= [h - h(p)]_p \\
&= [\mu f + \nu g - \mu f(p) - \nu g(p)]_p \\
&= \mu[f - f(p)]_p + \nu[g - g(p)]_p \\
&= \mu\,\mathrm{d}f(p) + \nu\,\mathrm{d}g(p)\,,
\end{aligned} \tag{8.4.3}
$$

which shows that d is linear. Now let $h = fg$ instead, which yields

$$
\begin{aligned}
\mathrm{d}h(p) &= [h - h(p)]_p \\
&= [fg - f(p)g(p)]_p \\
&= [fg - f(p)g(p) - (f - f(p))(g - g(p))]_p \\
&= [fg(p) + f(p)g - 2f(p)g(p)]_p \\
&= [f - f(p)]_p g(p) + f(p)[g - g(p)]_p\,,
\end{aligned} \tag{8.4.4}
$$

where we used the fact that $(f - f(p))(g - g(p)) \in I_p^2$, so that we may add it to obtain another representative of the same equivalence class. This shows that also the Leibniz rule is satisfied. ■

Another important relation besides the linearity and the Leibniz rule is the following, which can be understood as a chain rule.

Theorem 8.4.2. *Let M be a manifold, $f \in C^\infty(M, \mathbb{R})$ a function on M and $u \in C^\infty(\mathbb{R}, \mathbb{R})$. Then for all $p \in M$ holds*

$$
\mathrm{d}(u \circ f)(p) = u'(f(p))\mathrm{d}f(p)\,. \tag{8.4.5}
$$

Proof. Let $p \in M$. Following Hadamard's lemma, we can write u in the form

$$
u(x) = u(f(p)) + (x - f(p))\tilde{u}(x)\,, \tag{8.4.6}
$$

and we use the abbreviations $U = u \circ f$ and $\tilde{U} = \tilde{u} \circ f$. Then we have

$$
\begin{aligned}
\mathrm{d}U(p) &= [U - U(p)]_p \\
&= [(f - f(p))\tilde{U}]_p \\
&= [(f - f(p))\tilde{U}(p) - (f - f(p))(\tilde{U} - \tilde{U}(p))]_p \\
&= \tilde{U}(p)[(f - f(p))]_p \\
&= u'(f(p))\mathrm{d}f(p)\,,
\end{aligned} \tag{8.4.7}
$$

where we used the fact that $\tilde{U} - \tilde{U}(p) \in I_p$ and

$$
u'(f(p)) = (f(p) - f(p))\tilde{u}'(f(p)) + \tilde{u}(f(p)) = \tilde{u}(f(p)) = \tilde{U}(p)\,. \tag{8.4.8}
$$

■

Another possibility to obtain a new function from a known one is the composition $f \circ \varphi \in C^\infty(M, \mathbb{R})$ for $\varphi \in C^\infty(M, N)$ and $f \in C^\infty(N, \mathbb{R})$. Its differential will be discussed in section 11.3.

We finally express $\mathrm{d}f$ in the coordinate basis $(\mathrm{d}x^a)$. Using the coordinate expression (8.1.15) we see that this is simply

$$
\mathrm{d}f = \partial_a f\,\mathrm{d}x^a\,, \tag{8.4.9}
$$

which looks identical to (8.1.15), but where the (ambiguous) notation is now supposed to denote the action of the coordinate vector field ∂_a on f.

8.5 Duality of tangent and cotangent bundles

When we defined the cotangent space and cotangent bundle we already had some hints that there is a duality between the tangent and cotangent bundles. We now make this precise and discuss this topic in detail. We start by proving the following theorem.

Theorem 8.5.1. *The tangent space T_pM and the cotangent space T_p^*M at any point p on a manifold M are dual vector spaces.*

Proof. To show that T_pM is the dual vector space of T_p^*M, we need to show that there is an isomorphism $\theta : T_pM \to (T_p^*M)^*$, which we construct as follows. Recall that the elements of $T_p^*M = I_p/I_p^2$ are equivalence classes $[f]_p = f + I_p^2$ of functions $f \in I_p$. For such an equivalence class $[f]_p \in T_p^*M$ and a derivation $D \in T_pM$ we define

$$\begin{array}{cccc} \theta(D) & : & T_p^*M & \to & \mathbb{R} \\ & & [f]_p & \mapsto & D(f) \end{array} . \tag{8.5.1}$$

We still need to check that this is well-defined and does not depend on the choice of the representative f. Since derivations are linear functions, this is equivalent to showing that D vanishes on I_p^2. Since the elements of I_p^2 are products of functions $f, g \in I_p$, we have

$$D(fg) = D(f)g(p) + f(p)D(g) = 0\,, \tag{8.5.2}$$

since $f(p) = g(p) = 0$. Further, we see that $\theta(D)$ is linear, since

$$\begin{aligned} \theta(D)(\lambda[f]_p + \mu[g]_p) &= \theta(D)([\lambda f + \mu g]_p) \\ &= D(\lambda f + \mu g) \\ &= \lambda D(f) + \mu D(g) \\ &= \lambda\theta(D)([f]_p) + \mu\theta(D)([g]_p)\,. \end{aligned} \tag{8.5.3}$$

To see that θ is an isomorphism of the vector spaces T_pM and $(T_p^*M)^*$, we need to show that it is linear and possesses an inverse. Linearity follows from

$$\theta(\lambda D_1 + \mu D_2)([f]_p) = \lambda D_1(f) + \mu D_2(f) = \lambda\theta(D_1)([f]_p) + \mu\theta(D_2)([f]_p)\,. \tag{8.5.4}$$

We finally show the existence of an inverse $\vartheta : (T_p^*M)^* \to T_pM$ by explicit construction. Let $\alpha \in (T_p^*M)^*$ and define

$$\vartheta(\alpha) : C^\infty(M, \mathbb{R}) \to \mathbb{R}, f \mapsto \alpha([f - f(p)]_p) \tag{8.5.5}$$

To see that $\vartheta(\alpha)$ is a derivation, we check its linearity

$$\begin{aligned} \vartheta(\alpha)(\lambda f + \mu g) &= \alpha([\lambda(f - f(p)) + \mu(g - g(p))]_p) \\ &= \alpha(\lambda[f - f(p)]_p + \mu[g - g(p)]_p) \\ &= \lambda\alpha([f - f(p)]_p) + \mu\alpha([g - g(p)]_p) \\ &= \lambda\vartheta(\alpha)(f) + \mu\vartheta(\alpha)(g) \end{aligned} \tag{8.5.6}$$

and product rule

$$\begin{aligned} \vartheta(\alpha)(fg) &= \alpha([fg - f(p)g(p)]_p) \\ &= \alpha([(f - f(p))(g - g(p)) + f(p)(g - g(p)) + (f - f(p))g(p)]_p) \\ &= f(p)\alpha([g - g(p)]_p) + \alpha([f - f(p)]_p)g(p) \\ &= f(p)\vartheta(\alpha)(g) + \vartheta(\alpha)(f)g(p)\,. \end{aligned} \tag{8.5.7}$$

We finally need to check that the functions θ and ϑ defined above are inverses of each other. We first check that

$$\theta(\vartheta(\alpha))([f]_p) = \vartheta(\alpha)(f) = \alpha([f - \underbrace{f(p)}_{=0}]_p) = \alpha([f]_p) \tag{8.5.8}$$

for $\alpha \in (T_p^*M)^*$ and $f \in I_p$. Conversely, for $D \in T_pM$ and $f \in C^\infty(M, \mathbb{R})$ we have

$$\vartheta(\theta(D))(f) = \theta(D)([f - f(p)]_p) = D(f - f(p))\,. \tag{8.5.9}$$

To see that the latter equals $D(f)$, we need to show that a derivation D vanishes on a constant function c. This follows from the linearity of D together with the product rule, since

$$D(c)f = D(cf) - cD(f) = cD(f) - cD(f) = 0 \tag{8.5.10}$$

for all $f \in C^\infty(M, \mathbb{R})$. We have thus shown that θ and ϑ are indeed inverses of each other, so that $T_pM \cong (T_p^*M)^*$. Since T_p^*M is a finite-dimensional real vector space of dimension $\dim M$, which we have shown using Hadamard's lemma, it follows that also $(T_p^*M)^*$ and thus T_pM are real vector spaces of dimension $\dim M$. Finally, since the double dual V^{**} of a finite-dimensional vector space V is again isomorphic to V, it follows that also $T_p^*M \cong (T_pM)^*$. ■

This rather lengthy proof was necessary since we provided an own definition for both tangent and cotangent spaces. The advantage of the approach we used here is that it gave us a deeper understanding of the structure of these spaces and an interpretation for their elements in terms of functions on the manifold, which will be useful during the remainder of the lecture course. Instead of explicitly writing the isomorphisms θ and ϑ constructed above, we will simply write

$$\langle D, u \rangle = \theta(D)(u) \tag{8.5.11}$$

for the canonical pairing between $D \in T_pM$ and $u \in T_p^*M$. To visualize this pairing, we can make use of the visualizations of tangent vectors and cotangent vectors shown in figure 7.1 and figure 8.1, respectively. Drawing both visualizations in one diagram, as shown in figure 8.5, with a common origin for the tangent vector arrow and the hyperplane stack, the endpoint of the arrow marks a hyperplane, corresponding to a level set of $\hat{F}$. The corresponding value of $\hat{F}$ denotes the value of $D(f) = \langle D, [f]_p \rangle$. Using the same method, one can visualize the linearity of the canonical pairing in each argument, as shown in figures 8.6, 8.7 and 8.8.

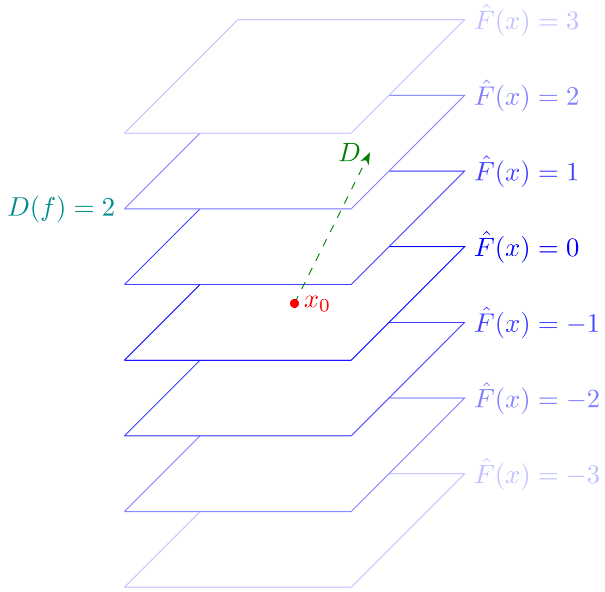


Figure 8.5: Visualization of the canonical pairing between a vector and a covector.

In the literature one often finds another approach, which simply defines the cotangent bundle as the dual of the tangent bundle. So far, we have shown only that the tangent and cotangent spaces at each point $p \in M$ are dual vector spaces. We will now go one step further and show that also the geometry of the tangent and cotangent bundles is related, by proving the following statement:

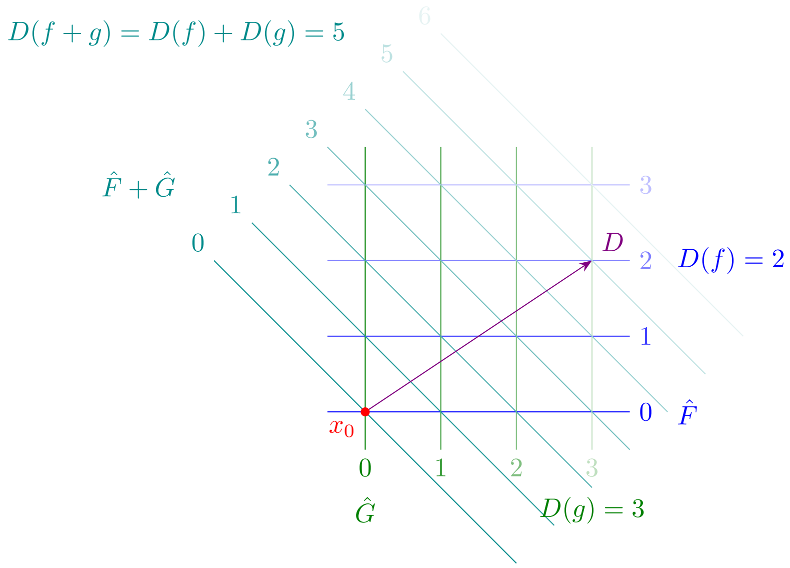


Figure 8.6: Visualization of the addition rule $D(f+g) = D(f) + D(g)$.

Theorem 8.5.2. *The cotangent bundle T^*M of a manifold M is the dual bundle of the tangent bundle TM (and vice versa).*

Proof. Recall that the local trivializations of TM and T^*M are defined by the coordinate bases (∂_a) and $(\mathrm{d}x^a)$ constructed from charts (U, ϕ). The former is constructed such that for a tangent vector $v \in T_pM$ at a point $p \in M$ and a function $f \in C^\infty(M, \mathbb{R})$ holds $v(f) = v^a \partial_a f$, while we expressed a cotangent vector $[f]_p \in T_p^*M$ as $\partial_a f \mathrm{d}x^a$. It thus follows directly that

$$\begin{aligned} v^a \partial_b f \langle \partial_a, \mathrm{d}x^b \rangle &= \langle v^a \partial_a, \partial_b f \mathrm{d}x^b \rangle \\ &= \langle v, [f]_p \rangle \\ &= v(f) \\ &= v^a \partial_a f \\ &= v^a \partial_b f \delta_a^b \,, \end{aligned} \tag{8.5.12}$$

so that $\langle \partial_a, \mathrm{d}x^b \rangle = \delta_a^b$. Hence, the coordinate basis (∂_a) of T_pM and $(\mathrm{d}x^a)$ of T_p^*M at any point $p \in M$ are dual bases. Since these bases define the local trivializations, following section 4.1 we find that TM and T^*M are dual vector bundles. ■

Given this result, we can now also make use of the canonical pairing 4.1.2 and write

$$\langle X, \omega \rangle(p) = \langle X(p), \omega(p) \rangle \tag{8.5.13}$$

for $X \in \Gamma(TM)$ and $\omega \in \Gamma(T^*M)$. A useful example of this operation is obtained if the covector field is given as the total differential of a function, which leads to the following relation:

Theorem 8.5.3. *Let M be a manifold, $f \in C^\infty(M, \mathbb{R})$ a function on M and $X \in \mathrm{Vect}(M)$ a vector field on M. Then $Xf = \mathrm{tr}_1^1(X \otimes \mathrm{d}f) = \langle X, \mathrm{d}f \rangle$.*

Proof. It is clear from the definition of a tensor contraction that $\mathrm{tr}_1^1(X \otimes \mathrm{d}f) = \langle X, \mathrm{d}f \rangle$. To see that this also equals Xf, recall from definition 7.4.2 that for every $p \in M$ we obtain $(Xf)(p)$ by applying the derivation $v = X(p) \in T_pM$ to f. Further, $\mathrm{d}f$ is defined for all p as the equivalence class $\mathrm{d}f(p) = [f - f(p)]_p \in T_p^*M = I_p/I_p^2$. Finally, the pairing $\langle v, [f - f(p)]_p \rangle$ is given by $v(f)$, which completes the proof. ■

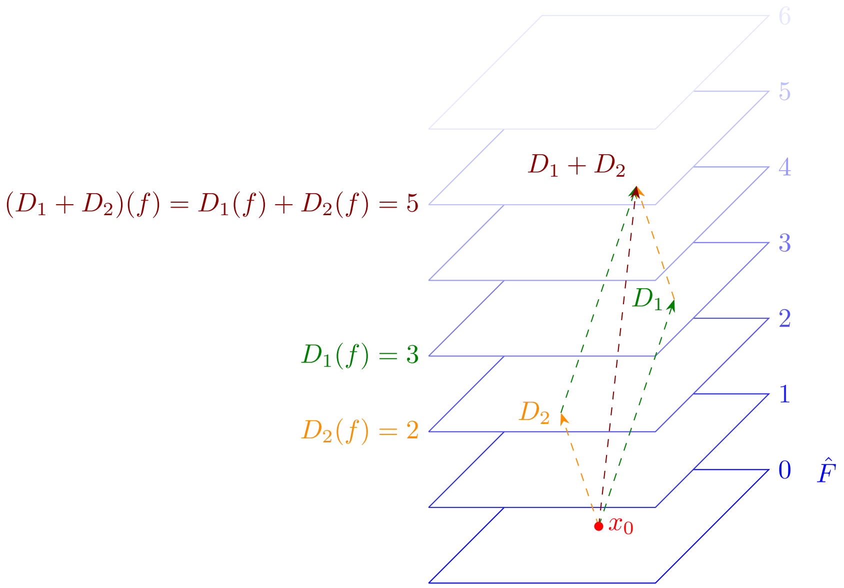


Figure 8.7: Visualization of the addition rule $(D_1 + D_2)(f) = D_1(f) + D_2(f)$.

There is an even faster way to see this using coordinates, where one easily reads off

$$\begin{aligned}\langle X, \mathrm{d}f\rangle &= \langle X^a\partial_a, \partial_b f\, \mathrm{d}x^b\rangle \\ &= X^a\partial_b f\langle\partial_a, \mathrm{d}x^b\rangle \\ &= X^a\partial_b f\delta^b_a \\ &= X^a\partial_a f \\ &= Xf\end{aligned} \tag{8.5.14}$$

8.6 Tensors over the tangent bundle

The tangent and cotangent bundles we introduced so far are the building blocks of another structure, called *tensor bundles*, which we will frequently encounter during the remainder of the course and extensively use in physics. In fact, physical quantities are usually modeled by *tensor fields* on a spacetime manifold, i.e., sections of a tensor bundle. In this section we will explain this notion. We begin with the following definition:

Definition 8.6.1 (Tensor bundle of the tangent bundle)**.** Let M be a manifold. The *tensor bundle* of type (r, s) for $r, s \in \mathbb{N}$ is the tensor product bundle

$$T^r_s M = \underbrace{TM \otimes \ldots \otimes TM}_{r \text{ times}} \otimes \underbrace{T^*M \otimes \ldots \otimes T^*M}_{s \text{ times}} . \tag{8.6.1}$$

This is of course simply the definition of the tensor bundle 4.3.3, where we use the vector bundle $E = TM$. Note that it is conventional to denote this bundle as $T^r_s M$ (and not TM^r_s, as one might naively expect), which is the reason for pointing out its definition separately in this section. Also the following fact about the tensor bundle now directly follows from our more general knowledge on the dimensions of the tangent and cotangent bundles, as well as tensor product bundles; see definitions 4.3.1, 7.2.1 and 8.2.1.

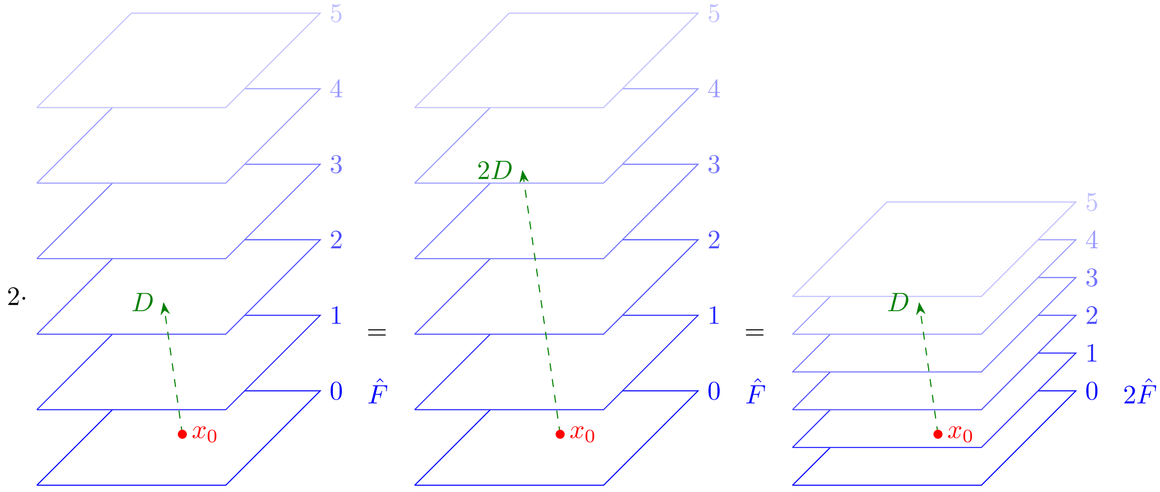


Figure 8.8: Visualization of the multiplication rule $(2D)(f) = D(2f) = 2D(f)$.

Theorem 8.6.1. *The tensor bundle $T^r_s M$ as defined above is a vector bundle of rank $(\dim M)^{r+s}$ over M.*

Proof. By construction, the tangent bundle is a vector bundle of rank $\dim M$. The same holds for the cotangent bundle. By repeatedly applying theorem 4.3.1 one finds that the rank of $T^r_s M$ is $(\dim M)^{r+s}$. ■

The rank of the tensor bundle can also be seen by introducing coordinates (x^a) on M. For any point $p \in M$ we then have the coordinate bases (∂_a) of $T_p M$ and $(\mathrm{d}x^a)$ of $T^*_p M$. The corresponding coordinate basis of $T^r_{s\,p} M$ is then given by the elements

$$\partial_{a_1} \otimes \ldots \otimes \partial_{a_r} \otimes \mathrm{d}x^{b_1} \otimes \ldots \otimes \mathrm{d}x^{b_s} \,, \tag{8.6.2}$$

where each index runs from 1 to $\dim M$, so that the basis has $(\dim M)^{r+s}$ elements. Any element $V \in T^r_{s\,p} M$ can be expanded using this basis in the form

$$V = V^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} \partial_{a_1} \otimes \ldots \otimes \partial_{a_r} \otimes \mathrm{d}x^{b_1} \otimes \ldots \otimes \mathrm{d}x^{b_s} \,, \tag{8.6.3}$$

with r upper and s lower indices.

Now it is also clear that vector fields are tensor fields of type $(1, 0)$, while covector fields are tensor fields of type $(0, 1)$. In the same spirit as with general tensor fields discussed in section 5.1, we can also regard real functions as tensor fields of rank $(0, 0)$ in the tangent bundle. In physics, a tensor field of type $(0, 0)$ is also called a *scalar field*.

Finally, we point out that tensor fields, which are sections of $T^r_s M$, are of course a special case of the more general case of tensor fields we discussed intensively in chapter 5. In particular, all operations we have defined on general tensor fields apply to those for the tangent bundle, and one finds the same canonical tensor fields.

Chapter 9

Differential forms

9.1 Antisymmetric tensors and differential forms

In section 4.4 we already discussed the exterior power bundle, whose sections can be seen as totally antisymmetric tensors. We now come to a particularly useful class of such bundles, which are constructed from the cotangent bundle. Its sections are called *differential forms* and play a role for calculating derivatives and integrals. We start with their formal definition.

Definition 9.1.1 (Differential form)**.** A *differential form of rank* k (or k*-form*) on a manifold M is a section of the exterior power bundle $\Lambda^k T^*M$ for $k \in \mathbb{N}$. The space of all k-forms on M is denoted $\Omega^k(M)$, while the space of all differential forms is denoted

$$\Omega^\bullet(M) = \bigoplus_{k=0}^{\dim M} \Omega^k(M)\,. \tag{9.1.1}$$

The space $\Omega^\bullet(M)$, whose elements are formal sums of differential forms of any degree k, is useful to describe operators which act on all differential forms, irrespective of their degree. However, in practice one rarely considers linear combinations of differential forms with different degree, and only works with "homogeneous" elements, which belong to a particular subspace $\Omega^k(M) \subset \Omega^\bullet(M)$.

Given coordinates (x^a) on M, we can use the coordinate basis covector fields $(\mathrm{d}x^a)$ to construct a basis of $\Lambda^k T^*M$ with basis elements of the form $\mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k}$, as shown in section 4.4. A differential form $\omega \in \Omega^k(M)$ can thus be expressed in the form

$$\omega = \frac{1}{k!}\omega_{a_1\cdots a_k}\mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k}\,, \tag{9.1.2}$$

where the components are totally antisymmetric, $\omega_{a_1\cdots a_k} = \omega_{[a_1\cdots a_k]}$. It thus becomes clear that a k-form is simply a totally antisymmetric tensor field of type $(0,k)$. Here we used the bracket notation (4.4.7) introduced in section 4.4.

There are some special cases. For $k = 0$ we have $\Lambda^0 T^*M \cong M \times \mathbb{R}$, so that a 0-form is simply a real function and $\Omega^0(M) \cong C^\infty(M,\mathbb{R})$. We also encountered $\Lambda^1 T^*M \cong T^*M$, so that a 1-form is the same as a covector field. This justifies the notation $\Omega^1(M)$ for the space of covector fields introduced in the last lecture. Finally, for $k = \dim M$, we get again a vector bundle of rank 1, as for $k = 0$. However, these vector bundles are in general not isomorphic!

In the following we will study a few operations on differential forms and their properties.

9.2 Exterior product

Recall from linear algebra that given a vector space V, the exterior algebra defines a wedge product

$$\begin{array}{cccc} \wedge : & \Lambda^k V \times \Lambda^l V & \to & \Lambda^{k+l} V \\ & (u, v) & \mapsto & u \wedge v \end{array}, \tag{9.2.1}$$

which acts on basis vectors in the obvious way,

$$(e_{i_1} \wedge \ldots \wedge e_{i_k}) \wedge (e_{j_1} \wedge \ldots \wedge e_{j_l}) = e_{i_1} \wedge \ldots \wedge e_{i_k} \wedge e_{j_1} \wedge \ldots \wedge e_{j_l}\,, \tag{9.2.2}$$

and is linear in both u and v. Pointwise application of the wedge product to differential forms allows us to define the following:

Definition 9.2.1 (Exterior product)**.** Let M be a manifold and $\alpha \in \Omega^k(M)$ and $\beta \in \Omega^l(M)$. Their *exterior product* (or *wedge product*) is the differential form $\alpha \wedge \beta \in \Omega^{k+l}(M)$ such that for all $p \in M$ holds

$$(\alpha \wedge \beta)(p) = \alpha(p) \wedge \beta(p)\,. \tag{9.2.3}$$

Using coordinates (x^a), we have

$$\begin{aligned} \alpha \wedge \beta &= \frac{1}{k!l!}(\alpha_{a_1 \cdots a_k} \mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k}) \wedge (\beta_{b_1 \cdots b_l} \mathrm{d}x^{b_1} \wedge \ldots \wedge \mathrm{d}x^{b_l}) \\ &= \frac{1}{k!l!}\alpha_{[a_1 \cdots a_k}\beta_{b_1 \cdots b_l]} \mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k} \wedge \mathrm{d}x^{b_1} \wedge \ldots \wedge \mathrm{d}x^{b_l}\,. \end{aligned} \tag{9.2.4}$$

The antisymmetrization comes from the fact that the wedge product of the basis elements $(\mathrm{d}x^a)$ is totally antisymmetric.

The following properties of the exterior product follow directly from the properties of the wedge product.

Theorem 9.2.1. *For $\alpha \in \Omega^k(M)$, $\beta \in \Omega^l(M)$ and $\gamma \in \Omega^r(M)$, the exterior product satisfies:*

1. *Graded anticommutativity:*
$$\alpha \wedge \beta = (-1)^{kl} \beta \wedge \alpha\,. \tag{9.2.5}$$
2. *Associativity:*
$$\alpha \wedge (\beta \wedge \gamma) = (\alpha \wedge \beta) \wedge \gamma = \alpha \wedge \beta \wedge \gamma\,. \tag{9.2.6}$$
3. *$\mathbb{R}$-linearity in each factor.*

Proof. ▶. . . ◀ ■

A special case is given if $k = 0$ or $l = 0$. In this case one of the terms in the exterior product is a real function $f \in C^\infty(M, \mathbb{R})$, and the exterior product reduces to the ordinary product

$$f \wedge \alpha = \alpha \wedge f = f\alpha. \tag{9.2.7}$$

9.3 Exterior derivative

We have seen in section 8.4 that the total differential $\mathrm{d}f$ of a function $f \in \Omega^0(M) \cong C^\infty(M, \mathbb{R})$ is a covector field, and hence a 1-form. The total differential can thus be viewed as a function $\mathrm{d} : \Omega^0(M) \to \Omega^1(M)$, which is a special case of the following construction.

Definition 9.3.1 (Exterior derivative)**.** For a manifold M, the *exterior derivative* $\mathrm{d} : \Omega^k(M) \to \Omega^{k+1}(M)$ for all $k \in \mathbb{N}$ is the unique linear function such that:

- $\mathrm{d}f$ is the total differential for any $f \in \Omega^0(M) \cong C^\infty(M, \mathbb{R})$.
- $\mathrm{d}(\mathrm{d}\omega) = 0$ for any $\omega \in \Omega^k(M)$.
- $\mathrm{d}(\alpha \wedge \beta) = \mathrm{d}\alpha \wedge \beta + (-1)^p \alpha \wedge \mathrm{d}\beta$ for $\alpha \in \Omega^p(M)$ and $\beta \in \Omega^q(M)$, where $p, q \in \mathbb{N}$.

In coordinates (x^a) we can write a k-form as $\omega = \frac{1}{k!}\omega_{a_1\cdots a_k}\mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k}$ and use the definition above to derive the formula

$$\begin{aligned}
\mathrm{d}\omega &= \frac{1}{k!}\mathrm{d}(\omega_{a_1\cdots a_k}\mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k}) \\
&= \frac{1}{k!}\mathrm{d}(\omega_{a_1\cdots a_k}) \wedge \mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k} \\
&\quad + \frac{1}{k!}\omega_{a_1\cdots a_k}\sum_{i=1}^{k}(-1)^{i-1}\mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_{i-1}} \wedge \mathrm{d}(\mathrm{d}x^{a_i}) \wedge \mathrm{d}x^{a_{i+1}} \wedge \ldots \wedge \mathrm{d}x^{a_k} \\
&= \frac{1}{k!}\partial_{[b}\omega_{a_1\cdots a_k]}\mathrm{d}x^b \wedge \mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k}\,,
\end{aligned} \tag{9.3.1}$$

where the antisymmetrization in the last line again comes from the total antisymmetry of the wedge product.

The following notions are closely related to the exterior derivative.

Definition 9.3.2 (Closed form)**.** A k-form $\omega \in \Omega^k(M)$ on a manifold M is called *closed* if $\mathrm{d}\omega = 0$.

Definition 9.3.3 (Exact form)**.** A k-form $\omega \in \Omega^k(M)$ on a manifold M is called *exact* if there exists a $(k-1)$-form $\sigma \in \Omega^{k-1}(M)$ such that $\mathrm{d}\sigma = \omega$.

Now the following statement is straightforward.

Theorem 9.3.1. *Every exact form is closed.*

Proof. This follows immediately from the fact that $\mathrm{d}^2 = 0$. ■

The converse, however, is not true: not every closed form is exact. In fact, there is a deep connection between the question which closed forms are exact and the topology of a manifold, known under the keyword of *de Rham cohomology*. We will not pursue this topic here. However, it is helpful to remark that the differential forms form a *cochain complex*

$$0 \to \Omega^0(M) \xrightarrow{\mathrm{d}} \Omega^1(M) \xrightarrow{\mathrm{d}} \ldots \xrightarrow{\mathrm{d}} \Omega^{\dim M}(M) \xrightarrow{\mathrm{d}} 0\,, \tag{9.3.2}$$

which means that for any of the spaces $\Omega^k(M)$ in this diagram the image of the homomorphism on the left (which is given by the exact k-forms) is contained in the kernel of the homomorphism on the right (which is given by the closed k-forms). This is a consequence of theorem 9.3.1 and hence of $\mathrm{d}^2 = 0$.

9.4 Interior product

Also the pairing $\langle X, \omega\rangle$ between a vector field $X \in \mathrm{Vect}(M)$ and a covector field $\omega \in \Omega^1(M)$ introduced in the previous lecture is a special case of a more general construction, which we discuss in this section and which is defined as follows.

Definition 9.4.1 (Interior product)**.** For a manifold M, the *interior product* $\iota : \mathrm{Vect}(M) \times \Omega^{k+1}(M) \to \Omega^k(M)$ is the unique function such that for any $X \in \mathrm{Vect}(M)$:

- $\iota_X \alpha = \langle X, \alpha\rangle$ for $\alpha \in \Omega^1(M)$.
- $\iota_X(\lambda\alpha + \mu\beta) = \lambda\iota_X\alpha + \mu\iota_X\beta$ for $\lambda, \mu \in \mathbb{R}$ and $\alpha, \beta \in \Omega^{k+1}(M)$.
- $\iota_X(\alpha \wedge \beta) = (\iota_X\alpha) \wedge \beta + (-1)^p\alpha \wedge (\iota_X\beta)$ for $\alpha \in \Omega^p(M)$ and $\beta \in \Omega^q(M)$, where $p, q \in \mathbb{N}$.

Instead of writing $\iota_X\omega$ for the interior product, it is also common to write $X \lrcorner\, \omega$. We will use both notations interchangeably, depending on which one is more convenient.

For a vector field $X = X^a\partial_a$ and a differential form $\omega = \frac{1}{k!}\omega_{a_1\cdots a_k}\mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k}$ expressed in coordinates (x^a) we can directly use the properties given in the definition above to read off the coordinate formula

$$\begin{aligned}
\iota_X\omega &= \frac{1}{k!}\iota_{X^b\partial_b}(\omega_{a_1\cdots a_k}\mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k})\\
&= \frac{1}{k!}X^b\omega_{a_1\cdots a_k}\sum_{i=1}^{k}(-1)^{i-1}\langle\partial_b, \mathrm{d}x^{a_i}\rangle\mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_{i-1}} \wedge \mathrm{d}x^{a_{i+1}} \wedge \ldots \wedge \mathrm{d}x^{a_k}\\
&= \frac{1}{k!}\omega_{a_1\cdots a_k}\sum_{i=1}^{k}(-1)^{i-1}X^{a_i}\mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_{i-1}} \wedge \mathrm{d}x^{a_{i+1}} \wedge \ldots \wedge \mathrm{d}x^{a_k}\\
&= \frac{1}{(k-1)!}X^{a_1}\omega_{a_1\cdots a_k}\mathrm{d}x^{a_2} \wedge \ldots \wedge \mathrm{d}x^{a_k}\,,
\end{aligned} \tag{9.4.1}$$

where the last line follows from the fact that we took the components $\omega_{a_1\cdots a_k}$ to be totally antisymmetric. Repeating this process k times, we find the helpful formula

$$\iota_{X_k}\cdots\iota_{X_1}\omega = X_1^{a_1}\cdots X_k^{a_k}\omega_{a_1\cdots a_k}\,. \tag{9.4.2}$$

In particular, if one chooses for the vector fields the coordinate basis vector fields, one finds

$$\iota_{\partial_{a_k}}\cdots\iota_{\partial_{a_1}}\omega = \omega_{a_1\cdots a_k}\,. \tag{9.4.3}$$

This antisymmetry also plays a role in the following statement.

Theorem 9.4.1. *For $X, Y \in \mathrm{Vect}(M)$ and $\omega \in \Omega^k(M)$ the interior product satisfies $\iota_X(\iota_Y\omega) = -\iota_Y(\iota_X\omega)$.*

Proof. ▶. . . ◀ ■

We will not prove this here, and instead present another theorem, which can be helpful in practical calculations.

Theorem 9.4.2. *Given a k-form $\omega \in \Omega^k(M)$ and $k+1$ vector fields $X_0, \ldots X_k \in \mathrm{Vect}(M)$, the exterior derivative, interior product and Lie bracket are related by*

$$\begin{aligned}\iota_{X_k} \cdots \iota_{X_0} \mathrm{d}\omega = &\sum_{i=0}^{k} (-1)^i X_i \left(\iota_{X_k} \cdots \iota_{X_{i+1}} \iota_{X_{i-1}} \cdots \iota_{X_0} \omega\right) \\ &+ \sum_{i=0}^{k-1} \sum_{j=i+1}^{k} (-1)^{i+j} \iota_{X_k} \cdots \iota_{X_{j+1}} \iota_{X_{j-1}} \cdots \iota_{X_{i+1}} \iota_{X_{i-1}} \cdots \iota_{X_0} \iota_{[X_i,X_j]} \omega \,.\end{aligned} \tag{9.4.4}$$

Proof. ►...◄ ■

For a 1-form $\omega \in \Omega^1(M)$ this formula reduces to

$$\begin{aligned}\iota_Y \iota_X \mathrm{d}\omega &= X(\iota_Y \omega) - Y(\iota_X \omega) - \iota_{[X,Y]} \omega \\ &= X(\langle Y, \omega \rangle) - Y(\langle X, \omega \rangle) - \langle [X, Y], \omega \rangle \,.\end{aligned} \tag{9.4.5}$$

Finally, we mention another helpful formula regarding the inner product with the commutator of two vector fields:

Theorem 9.4.3. *For $X, Y \in \mathrm{Vect}(M)$ and $\omega \in \Omega^k(M)$ the interior product with $[X, Y]$ is given by*

$$\iota_{[X,Y]} \omega = \mathrm{d}\iota_X \iota_Y \omega + \iota_X \mathrm{d}\iota_Y \omega - \iota_Y \mathrm{d}\iota_X \omega - \iota_Y \iota_X \mathrm{d}\omega \,. \tag{9.4.6}$$

Proof. ►...◄ ■

As we will see in section 16.5, this is closely related to the commutator of the interior product and the Lie derivative.

9.5 Vector-valued differential forms

Definition 9.5.1 (Vector-valued differential form). Let $(E, B, \pi, \mathbb{R}^n)$ be a vector bundle. A *differential k-form with values in E* is a section of the bundle $E \otimes \Lambda^k T^* B$. The space of all E-valued k-forms is denoted $\Omega^k(B, E)$. If $E = B \times V$ is a trivial bundle whose fiber is the vector space V, also the notation $\Omega^k(B, V)$ is used.

Chapter 10

Differential and pushforward

10.1 Differential and pushforward

The tensor bundles we have discussed in the previous sections, including the special cases of tangent and cotangent bundles, belong to a particular class called *natural bundles*. These bundles allow relating elements of their total spaces if only a map between their base spaces is given. One of these operations, which applies to the tangent bundle, can be defined as follows.

Definition 10.1.1 (Differential and pushforward)**.** Let M and N be manifolds and $\varphi : M \to N$ a smooth map. The *differential* of φ is the smooth map $\varphi_* : TM \to TN$ which assigns to a tangent vector $v \in T_pM$ at $p \in M$ (which is a derivation at p acting on functions on M) its *pushforward* $\varphi_*(v) \in T_{\varphi(p)}N$ (derivation at $\varphi(p) \in N$ acting on functions on N) along φ defined by

$$\varphi_*(v)(f) = v(f \circ \varphi) \tag{10.1.1}$$

for $f \in C^\infty(N, \mathbb{R})$.

To see that this definition makes sense and indeed yields a map $\varphi_* : TM \to TN$ one of course needs to check, following definition 7.1.1, that $\varphi_*(v)$ as defined above is a derivation and that φ_* is smooth. It is not very difficult to check this, as we will see by proving an even stronger statement.

Theorem 10.1.1. *The differential $\varphi_* : TM \to TN$ of a smooth map $\varphi : M \to N$ is a vector bundle homomorphism covering φ.*

Proof. We first need to check that for every $p \in M$ and $v \in T_pM$ we indeed have $\varphi_*(v) \in T_{\varphi(p)}N$, so that φ_* is a bundle morphism covering φ. In other words, we must check that $\varphi_*(v)$ is a derivation at $\varphi(p)$. First, $\varphi_*(v)$ is clearly linear, since precomposition with φ is linear, i.e.,

$$(\mu f + \nu g) \circ \varphi = \mu(f \circ \varphi) + \nu(g \circ \varphi) \tag{10.1.2}$$

for $\mu, \nu \in \mathbb{R}$ and $f, g \in C^\infty(N, \mathbb{R})$ and v is linear. Further, the Leibniz rule follows from

$$\begin{aligned}
\varphi_*(v)(fg) &= v((fg) \circ \varphi) \\
&= v((f \circ \varphi)(g \circ \varphi)) \\
&= v(f \circ \varphi)g(\varphi(p)) + f(\varphi(p))v(g \circ \varphi) \\
&= \varphi_*(v)(f)g(\varphi(p)) + f(\varphi(p))\varphi_*(v)(g)\,.
\end{aligned} \tag{10.1.3}$$

Hence, $\varphi_*(v) \in T_{\varphi(p)}N$. We then check that this map is linear on every fiber. Given another vector $w \in T_pM$ at the same point $p \in M$ we have

$$\begin{aligned}\varphi_*(\mu v + \nu w)(f) &= (\mu v + \nu w)(f \circ \varphi) \\ &= \mu v(f \circ \varphi) + \nu w(f \circ \varphi) \\ &= \mu \varphi_*(v)(f) + \nu \varphi_*(w)(f)\,,\end{aligned} \tag{10.1.4}$$

and so it is indeed linear. Finally, we also need to check its smoothness. However, since this proof makes use of charts, we defer it to section 10.3, where we derive its coordinate representation. ■

This of course raises the question, whether and under which circumstances the differential becomes a vector bundle isomorphism. To answer it, we first prove the following statement which further relates the differential of a map to the properties of derivatives, by showing that it satisfies a generalization of the chain rule.

Theorem 10.1.2. *Let M, N, O be manifolds and $\varphi_1 : M \to N$ and $\varphi_2 : N \to O$ smooth maps. Then their differentials satisfy*

$$(\varphi_2 \circ \varphi_1)_* = \varphi_{2*} \circ \varphi_{1*}\,. \tag{10.1.5}$$

Proof. Let $f \in C^\infty(O, \mathbb{R})$ be a function on O and $v \in TM$. It follows that

$$\begin{aligned}\varphi_{2*}(\varphi_{1*}(v))(f) &= \varphi_{1*}(v)(f \circ \varphi_2) \\ &= v((f \circ \varphi_2) \circ \varphi_1) \\ &= v(f \circ (\varphi_2 \circ \varphi_1)) \\ &= (\varphi_2 \circ \varphi_1)_*(v)(f)\,,\end{aligned} \tag{10.1.6}$$

using the fact that map composition $\circ$ is associative. ■

Since the identity map on every manifold acts as a neutral element in map composition, one may of course assume that the differential maps this element to the corresponding neutral element in the composition of vector bundle morphisms between tangent bundles. This is indeed the case, as one can easily show also explicitly.

Theorem 10.1.3. *Let M be a manifold and $\mathrm{id}_M : M \to M$ the identity on M. Its differential is given by $(\mathrm{id}_M)_* = \mathrm{id}_{TM}$.*

Proof. Given any vector $v \in TM$ and a function $f \in C^\infty(M, \mathbb{R})$ one has

$$(\mathrm{id}_M)_*(v)(f) = v(f \circ \mathrm{id}_M) = v(f)\,. \tag{10.1.7}$$

Since this holds for any function f, and a tangent vector is uniquely characterized by its action on functions, it follows that $(\mathrm{id}_M)_*(v) = v$, and thus $(\mathrm{id}_M)_* = \mathrm{id}_{TM}$. ■

With these two statements it is now straightforward to prove the following.

Theorem 10.1.4. *The differential $\varphi_* : TM \to TN$ of a smooth map $\varphi : M \to N$ is a vector bundle isomorphism if and only if φ is a diffeomorphism.*

Proof. If φ_* is a vector bundle isomorphism, then it is invertible and its inverse is also a vector bundle isomorphism. But then also the covered map φ must be invertible and have a smooth inverse, and thus be a diffeomorphism. Conversely, given a diffeomorphism φ, both its differential φ_* and the differential $(\varphi^{-1})_*$ are vector bundle homomorphisms. One finds that the latter is the inverse of the former, since

$$(\varphi^{-1})_* \circ \varphi_* = (\varphi^{-1} \circ \varphi)_* = (\mathrm{id}_M)_* = \mathrm{id}_{TM}\,, \tag{10.1.8}$$

using the previous two statements. Hence, φ_* is a vector bundle isomorphism. ■

Another class of maps which is worth studying if given by constant maps. In this case we find the following.

Theorem 10.1.5. *Let M, N be manifolds, $q \in N$ and $\varphi_q : M \to N, p \mapsto q$ the constant map which maps every point $p \in M$ to the same point q. Then its differential is given by*

$$\begin{array}{rccc} \varphi_{q*} & : & TM & \to & TN \\ & & v & \mapsto & 0 \in T_q N \end{array} . \tag{10.1.9}$$

Proof. Let $f \in C^\infty(N, \mathbb{R})$, $p \in M$ and $v \in T_p M$. Then we have

$$\varphi_{q*}(v)(f) = v(f \circ \varphi_q) = 0 \,, \tag{10.1.10}$$

since

$$\begin{array}{rccc} f \circ \varphi_q & : & M & \to & \mathbb{R} \\ & & p & \mapsto & f(q) \end{array} \tag{10.1.11}$$

is a constant function, and so any derivation acts trivially as of theorem 7.1.1. ■

The previous examples of diffeomorphisms and the constant function are in a certain sense extremes: while the differential of a diffeomorphism is bijective and thus preserves the information given by a tangent vector, the differential of a constant function sends all tangent vectors to the same zero element and thus discards all information. The following notion gives a measure for the amount of information, i.e., the number of vector components in a given basis, is preserved.

Definition 10.1.2 (Rank)**.** Let M, N be manifolds and $\varphi : M \to N$ a smooth map. For $p \in M$, the *rank* of φ in p is the rank of its differential φ_* in p. A map is of *constant rank* if it has the same rank in all points $p \in M$.

Maps of constant rank play an important role when it comes to defining certain subbundles of vector bundles, as we will see later.

10.2 Pushforward and curves

Although the definition 10.1.1 is probably the most clear and practical from an algebraic point of view, it is not very intuitive from a geometric perspective. We will therefore also discuss how the pushforward relates to the other definitions of tangent vectors we have given. The most geometric picture is given by that of the tangent vector of a curve discussed in section 7.3, and which leads us to the following statement.

Theorem 10.2.1. *Let M and N be manifolds, $\gamma \in C^\infty(\mathbb{R}, M)$ a curve on M and $\varphi : M \to N$ a smooth map. Then the pushforward $\varphi_*(\dot\gamma(t))$ is given by*

$$\varphi_*(\dot\gamma(t)) = \dot\Gamma(t) \,, \tag{10.2.1}$$

where $\Gamma = \varphi \circ \gamma \in C^\infty(\mathbb{R}, N)$.

Proof. First, note that by definition 7.3.1 we have $\dot\gamma(t) \in T_{\gamma(t)}M$ and $\dot\Gamma(t) \in T_{\Gamma(t)}N = T_{\varphi(\gamma(t))}N$, and so the base points $\gamma(t)$ and $\Gamma(t)$ are indeed related by φ. Following definition 10.1.1 we then find

$$\begin{aligned} \varphi_*(\dot\gamma(t))(f) &= \dot\gamma(t)(f \circ \varphi) \\ &= (f \circ \varphi \circ \gamma)'(t) \\ &= (f \circ \Gamma)'(t) \\ &= \dot\Gamma(t)(f) \end{aligned} \tag{10.2.2}$$

for any $f \in C^\infty(N, \mathbb{R})$. Since this holds for any smooth function f, and a tangent vector is uniquely defined by its base point and action on functions, we conclude that $\varphi_*(\dot{\gamma}(t)) = \dot{\Gamma}(t)$. ■

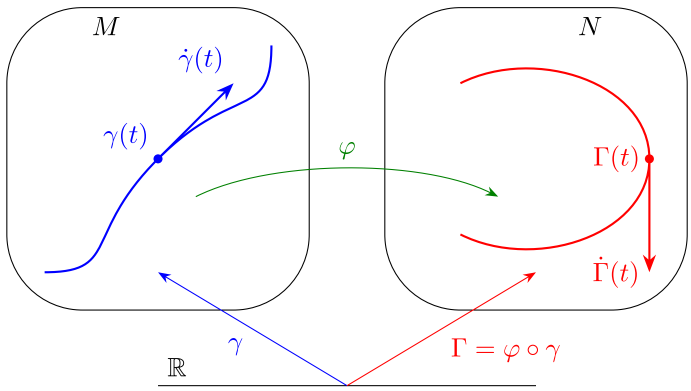


Figure 10.1: Illustration of the pushforward of the tangent vector of a curve. At each point $\gamma(t)$ the tangent vector $\dot{\gamma}(t)$ is mapped to the corresponding tangent vector of the composed curve $\Gamma = \varphi \circ \gamma$.

This is illustrated in figure 10.1. With this knowledge, the next statement follows immediately.

Theorem 10.2.2. *Let M and N be manifolds, $\gamma \in C^\infty(\mathbb{R}, M)$ a curve on M and $\varphi : M \to N$ a smooth map. The canonical lift $\dot{\Gamma} \in C^\infty(\mathbb{R}, TN)$ of $\Gamma = \varphi \circ \gamma$ is given by $\dot{\Gamma} = \varphi_* \circ \dot{\gamma}$, where $\dot{\gamma} \in C^\infty(\mathbb{R}, TM)$ is the canonical lift of γ.*

Proof. This follows immediately from definition 7.3.2 and theorem 10.2.1. ■

With the pushforward, we can also find this helpful statement on canonical lifts of curves.

Theorem 10.2.3. *A curve $\Gamma : \mathbb{R} \to TM$ is the canonical lift, $\Gamma = \dot{\gamma}$, of a curve $\gamma : \mathbb{R} \to M$ if and only if $\tau_* \circ \dot{\Gamma} = \Gamma$.*

Proof. For a curve γ, we have that the projection of its canonical lift along $\tau : TM \to M$ recovers the original curve, $\tau \circ \dot{\gamma}$, and so Γ can only be the canonical lift of $\tau \circ \Gamma$. The canonical lift of $\tau \circ \Gamma$ is $\tau_* \circ \dot{\Gamma}$, according to theorem 10.2.2, and this agrees with Γ if and only if $\tau_* \circ \dot{\Gamma} = \Gamma$. ■

10.3 Pushforward and charts

To get a better picture of the differential and the pushforward, we can write them in coordinates. In order to clearly distinguish between coordinate dependent and independent quantities, we explicitly write out all involved charts. Consider $p \in M$ and $v \in T_pM$. On N we pick a chart (V, χ) such that $\varphi(p) \in V$. We then pick a chart (U, ψ) on M such that $p \in U$ and $U \subset \varphi^{-1}(V)$. (Note that we can always obtain such a chart by choosing an arbitrary chart (U', ψ') around p and defining $U = U' \cap \varphi^{-1}(V)$ and $\psi = \psi'|_U$.) Denote the induced charts on TM and TN, which are obtained from the coordinate bases corresponding to (U, ψ) and (V, χ), by $(\tilde{U}, \tilde{\psi})$ and $(\tilde{V}, \tilde{\chi})$. We further denote by (x^a) the corresponding coordinates on M and (y^μ) the coordinates on N, as well as by $(x^a, \bar{x}^a)$ and $(y^\mu, \bar{y}^\mu)$ the induced coordinates on TM and TN. We use different indices (Latin for M and Greek for N) here to distinguish between objects living on different manifolds, and to make clear that Latin indices run from 1 to $m = \dim M$, while Greek indices run from 1 to $n = \dim N$. In these coordinates a map $\varphi : M \to N$ is expressed by

$$\hat{\varphi} = \chi \circ \varphi \circ \psi^{-1} \quad : \quad \begin{array}{ccc} \psi(U) & \to & \mathbb{R}^n \\ x & \mapsto & y = \hat{\varphi}(x) \end{array}, \tag{10.3.1}$$

where we introduced the abbreviation $\hat{\varphi}$.

Now let $v \in T_pM$ be a tangent vector and $g \in C^\infty(M, \mathbb{R})$ be a smooth function on M. The induced chart $(\tilde{U}, \tilde{\psi})$ is defined such that $\tilde{U} = \tau^{-1}(U)$, where $\tau : TM \to M$ is the tangent bundle projection, and

$$\begin{array}{rccc} \tilde{\psi} & : & \tilde{U} & \to & \mathbb{R}^{2m} \\ & & v & \mapsto & (\psi(\tau(v)), \bar{\psi}(v)) \end{array} , \tag{10.3.2}$$

where $\bar{\psi}(v) \in \mathbb{R}^m$ follows from

$$v(g) = \bar{\psi}^a(v) \, \frac{\partial}{\partial x^a}(g \circ \psi^{-1})(x)\Big|_{x=\psi(\tau(v))} . \tag{10.3.3}$$

In other words, in the coordinate basis ∂_a of T_pM, v is written as $v = v^a \partial_a = \bar{\psi}^a(v)\partial_a$. In the same fashion one constructs the induced chart $(\tilde{V}, \tilde{\chi})$ on TN.

Now let $f \in C^\infty(N, \mathbb{R})$ and $g = f \circ \varphi$. Our aim is to derive the coordinate expression for $\varphi_*(v)$, i.e., the expression

$$\tilde{\chi}(\varphi_*(v)) = (\chi(\tau'(\varphi_*(v))), \bar{\chi}(\varphi_*(v))) = (\chi(\varphi(\tau(v))), \bar{\chi}(\varphi_*(v))) , \tag{10.3.4}$$

where $\tau' : TN \to N$ denotes the tangent bundle projection and we used the fact that φ_* is a vector bundle homomorphism in the first component of this tuple. For the second component $\bar{\chi}(\varphi_*(v))$ we make use of the definition

$$\begin{aligned} (\varphi_*(v))(f) &= v(f \circ \varphi) \\ &= \bar{\psi}^a(v) \, \frac{\partial}{\partial x^a}(f \circ \varphi \circ \psi^{-1})(x)\Big|_{x=\psi(\tau(v))} \\ &= \bar{\psi}^a(v) \, \frac{\partial}{\partial x^a}(f \circ \chi^{-1} \circ \chi \circ \varphi \circ \psi^{-1})(x)\Big|_{x=\psi(\tau(v))} \\ &= \bar{\psi}^a(v) \, \frac{\partial}{\partial x^a}(f \circ \chi^{-1} \circ \hat{\varphi})(x)\Big|_{x=\psi(\tau(v))} \\ &= \underbrace{\bar{\psi}^a(v) \, \frac{\partial}{\partial x^a}\hat{\varphi}^\mu(x)\Big|_{x=\psi(\tau(v))}}_{=\bar{\chi}^\mu(\varphi_*(v))} \, \frac{\partial}{\partial y^\mu}(f \circ \chi^{-1})(y)\Big|_{y=\chi(\varphi(\tau(v)))} \\ &= \bar{\chi}^\mu(\varphi_*(v)) \, \frac{\partial}{\partial y^\mu}(f \circ \chi^{-1})(y)\Big|_{y=\chi(\varphi(\tau(v)))} , \end{aligned} \tag{10.3.5}$$

using the chain rule for functions $\hat{\varphi}$ and $f \circ \chi^{-1}$ defined on subsets of Euclidean spaces. This allows us to simply read off

$$\bar{\chi}(\varphi_*(v)) = \bar{\psi}^a(v) \, \frac{\partial}{\partial x^a}\hat{\varphi}(x)\Big|_{x=\psi(\tau(v))} = \mathrm{D}\hat{\varphi}(\psi(\tau(v))) \cdot \bar{\psi}(v) . \tag{10.3.6}$$

Here $\mathrm{D}\hat{\varphi}(x)$ denotes the Jacobian of $\hat{\varphi}$ at $x \in \psi(U)$. If the charts are clear from the context, one often omits them and identifies the points on the manifold with their coordinates, so that the map φ is simply expressed as the assignment $x \mapsto y$. Following this convention, the pushforward is also written as

$$\varphi_*(v) = \varphi_*(v^a \partial_a) = v^a \frac{\partial y^\mu}{\partial x^a} \partial_\mu , \tag{10.3.7}$$

where ∂_a and ∂_μ are the coordinate bases of T_pM and $T_{\varphi(p)}N$ induced by their respective charts. Its application to a function f is then written as

$$\varphi_*(v)(f) = v(f \circ \varphi) = v^a \partial_a f(y(x)) = v^a \frac{\partial y^\mu}{\partial x^a} \partial_\mu f(y) . \tag{10.3.8}$$

The appearance of the Jacobian also explains the name differential for the map φ_*, as it is basically some kind of derivative of φ. This relates to our previous finding 10.1.2 that the differential itself satisfies a chain rule.

The coordinate description found above finally also allows us to relate the pushforward also to the third interpretation of a tangent vector in terms of equivalence classes related to charts, which we gave in definition 7.1.3. ▶. . . ◀

10.4 Immersions

The property of the differential to assign to each map a vector bundle homomorphism, which restricts to a linear function on each fiber, allows us to classify maps by the properties of these linear functions, and thus make use of the tools of linear algebra. We start this study by defining the following class of maps.

Definition 10.4.1 (Immersion)**.** A map $\varphi : M \to N$ between two manifolds M and N is called an *immersion* if and only if for each $p \in M$ the restriction $\varphi_{*p} : T_pM \to T_{\varphi(p)}N$ of the differential $\varphi_* : TM \to TN$ to T_pM is injective.

Before we come to examples for immersions which we have actually already encountered, we discuss them more generally. Recall that both for maps in general, as well as for diffeomorphisms in particular, we have found that their composition yields again a map, or a diffeomorphism, respectively. The same statement holds also for immersions, as we shall see next.

Theorem 10.4.1. *Let M, N, O be manifolds and $\varphi : M \to N$ and $\psi : N \to O$ immersions. Then also $\psi \circ \varphi : M \to O$ is a immersion.*

Proof. Let $p \in M$. Since φ is an immersion, any distinct $u_1, u_2 \in T_pM$ will have distinct images $v_1 = \varphi_*(u_1), v_2 = \varphi_*(u_2) \in T_{\varphi(p)}N$. Since also ψ is an immersion, these have distinct images $w_1 = \psi_*(v_1), w_2 = \psi_*(v_2) \in T_{\psi(\varphi(p))}O$. Hence,

$$w_1 = \psi_*(\varphi_*(u_1)) = (\psi \circ \varphi)_*(u_1) \quad \text{and} \quad w_2 = \psi_*(\varphi_*(u_2)) = (\psi \circ \varphi)_*(u_2) \tag{10.4.1}$$

are distinct, and so $\psi \circ \varphi$ is an immersion. ■

One may of course ask whether also the converse is true, and one can draw any conclusions about the individual maps φ and ψ if $\psi \circ \varphi$ is an immersion. It turns out that this is indeed the case.

Theorem 10.4.2. *Let M, N, O be manifolds and $\varphi : M \to N$ and $\psi : N \to O$ maps such that $\psi \circ \varphi : M \to O$ is an immersion. Then also φ is an immersion.*

Proof. Let $p \in M$ and $u_1, u_2 \in T_pM$ two distinct tangent vectors. Since we assumed $\psi \circ \varphi$ to be an immersion, it follows that

$$w_1 = \psi_*(\varphi_*(u_1)) = (\psi \circ \varphi)_*(u_1) \neq (\psi \circ \varphi)_*(u_2) = \psi_*(\varphi_*(u_2)) = w_2 \tag{10.4.2}$$

are distinct. Hence, also $v_1 = \varphi_*(u_1)$ and $v_2 = \varphi_*(u_2)$ must be distinct. Since this holds for all $p \in M$ and distinct $u_1, u_2 \in T_pM$, it follows that φ is an immersion. ■

We will make use of this result when we show for a few classes of maps that they are immersions. However, we start with a simpler example; see also theorem 10.5.3 for a related statement.

Theorem 10.4.3. *Every diffeomorphism is an immersion.*

Proof. Recall from theorem 10.1.4 that the differential φ_* of a map φ is a vector bundle isomorphism if and only if φ is a diffeomorphism. In this case, φ_* restricts to a vector space isomorphism on every fiber, and is therefore in particular injective, so that φ is a immersion. ∎

For a diffeomorphism, we know that both the properties of surjectivity and injectivity on every fiber are satisfied. This is not surprising, since a diffeomorphism is bijective by definition, and also its inverse is a diffeomorphism. As one may expect, this changes as soon as one considers maps which do not possess an inverse. Nevertheless, one can still find situations in which immersions occur. A simple example are product manifolds. We start with the following basic observation.

Theorem 10.4.4. *Let M, N be manifolds and $M \times N$ their product. Then for every $q \in N$ the map*

$$\begin{array}{rcccc} \varphi_q & : & M & \to & M \times N \\ & & x & \mapsto & (x, q) \end{array} \tag{10.4.3}$$

is an immersion.

Proof. Consider the projection $\mathrm{pr}_1 : M \times N \to M$, and note that $\mathrm{pr}_1 \circ \varphi_q = \mathrm{id}_M$. The identity id_M is a diffeomorphism, and hence an immersion. Following theorem 10.4.2, also φ_q is an immersion. ∎

In the previous example, instead of fixing a point $q \in N$, we could also have considered an arbitrary smooth map $f : M \to N$, and then shown that $\varphi_f : M \to M \times N, p \mapsto (p, f(p))$ is an immersion, as we will do next.

Theorem 10.4.5. *Let M, N be manifolds and $M \times N$ their product. Then for map $f : M \to N$ the map*

$$\begin{array}{rcccc} \varphi_f & : & M & \to & M \times N \\ & & x & \mapsto & (x, f(x)) \end{array} \tag{10.4.4}$$

is an immersion.

Proof. The proof proceeds in full analogy to the preceding one. Consider the projection $\mathrm{pr}_1 : M \times N \to M$, and note that $\mathrm{pr}_1 \circ \varphi_f = \mathrm{id}_M$. The identity id_M is a diffeomorphism, and hence an immersion. Following theorem 10.4.2, also φ_f is an immersion. ∎

According to theorem 2.3.2, the map φ_f from the previous example is simply a section of the trivial fiber bundle $(M \times N, M, \mathrm{pr}_1, N)$. Instead of restricting ourselves to sections of a trivial bundle, we can also go one step further and consider sections of general fiber bundles. Here we work with local sections, since global sections may not exist for a given bundle.

Theorem 10.4.6. *Let (E, B, π, F) be a fiber bundle. Every local section $\sigma : U \to E$ with $U \subseteq B$ is an immersion.*

Proof. The proof essentially follows again the same line of argument as before, using the fact that the projection $\pi : E \to B$ satisfies $\pi \circ \sigma = \mathrm{id}_U$. Now $\mathrm{id}_U = \mathrm{id}_M|_U$ is a diffeomorphism, and hence an immersion. It thus follows from theorem 10.4.2 that also σ is an immersion. ∎

The previous examples show the full virtue of theorem 10.4.2. In order to show that a map φ is an immersion, one finds a map ψ of which it is known that $\psi \circ \varphi$ is an immersion.

10.5 Submersions

We now come to a closely related concept, which is in some sense dual to the previously introduced notion of an immersion. It can be defined as follows.

Definition 10.5.1 (Submersion)**.** A map $\varphi : M \to N$ between two manifolds M and N is called a *submersion* if and only of for each $p \in M$ the restriction $\varphi_{*p} : T_pM \to T_{\varphi(p)}N$ of the differential $\varphi_* : TM \to TN$ to T_pM is surjective.

While for an immersion we demanded that it restricts to an injective linear function on every tangent space, we thus now consider the case of surjective functions. It turns out that many properties of submersions are very similar to those of immersions, as we will show in this section. We first study the composition of submersions.

Theorem 10.5.1. *Let M, N, O be manifolds and $\varphi : M \to N$ and $\psi : N \to O$ submersions. Then also $\psi \circ \varphi : M \to O$ is a submersion.*

Proof. Let $p \in M$. Since ψ is a submersion, for every $u \in T_{\psi(\varphi(p))}O$ there exists some $v \in T_{\varphi(p)}N$ such that $\psi_*(v) = u$. Since also φ is a submersion, there further exists $w \in T_pM$ such that $\varphi_*(w) = v$, and therefore

$$(\psi \circ \varphi)_*(w) = \psi_*(\varphi_*(w)) = \psi_*(v) = u\,, \tag{10.5.1}$$

and so $\psi \circ \varphi$ is a submersion. ■

Again as in the case of immersions, one may ask for a converse statement, which allows to draw conclusions on the individual maps φ and ψ from their composition, in analogy to theorem 10.4.2. It turns out that this is indeed the case, but only under certain circumstances. This will be formulated as follows.

Theorem 10.5.2. *Let M, N, O be manifolds and $\varphi : M \to N$ and $\psi : N \to O$ maps such that $\psi \circ \varphi : M \to O$ is a submersion and φ is surjective. Then also ψ is a submersion.*

Proof. Let $q \in N$. Since φ is surjective, there exists a $p \in M$ such that $\varphi(p) = q$. Since $\psi \circ \varphi$ is a submersion, it follows that $(\psi \circ \varphi)_*|_{T_pM}$ is surjective, i.e., for every $u \in T_{\psi(\varphi(p))}O = T_{\psi(q)}O$ there exists $w \in T_pM$ such that $\psi_*(\varphi_*(w)) = u$. Defining $v = \varphi_*(w)$, we have thus found an element $v \in T_qN$ such that $\psi_*(v) = u$. Since this holds for all $q \in N$ and $u \in T_{\psi(q)}O$, it follows that ψ is a submersion. ■

Note that the assumption that φ is surjective is crucial in the proof above. The reason is that in order to show that ψ is a submersion, we must show that the restriction of ψ_* to the tangent space T_qN is surjective for all $q \in N$. But in order to find a preimage $v \in T_qN$, we made use of φ_*, and this is possible only if q belongs to the image $\varphi(M) \subseteq N$ of φ. Hence, we must demand that this image $\varphi(M)$ contains all of N, i.e., that φ is surjective. Compare this with the related statement 10.4.2, where no such restriction is needed. The reason is that here we must check that the restriction of φ_* to T_pM for all $p \in M$ is injective, and in order to show that, we require that p lies in the domain of $\psi \circ \varphi$. However, this is the case by definition, and so no restriction is needed. Another notable difference with theorem 10.4.2 is that for submersions we conclude on the properties of the *second* map ψ, while for immersions we conclude on the *first* map φ.

By definition, a map is a submersion if and only if the restriction of its differential to any fiber is surjective. The latter is in particular true if it is a vector bundle isomorphism, as in the case of the diffeomorphism, as we have also discussed in the case of immersions in theorem 10.4.3. The latter is therefore easy to prove.

Theorem 10.5.3. *Every diffeomorphism is a submersion.*

Proof. The proof proceeds in full analogy to the proof of theorem 10.4.3, by again using the fact that for a diffeomorphism φ, the differential φ_* restricts to a vector space isomorphism on every fiber, and is therefore in particular surjective on every fiber. ■

Another important class of submersions, which we discuss now, is given by projections. We first start with the most simple case, given by the projection of a product manifold. We find that the following holds.

Theorem 10.5.4. *Let M, N be manifolds and $M \times N$ their product. Then the projections $\mathrm{pr}_1 : M \times N \to M$ and $\mathrm{pr}_2 : M \times N \to N$ are submersions.*

Proof. Here we will restrict ourselves to the map $\mathrm{pr}_1 : M \times N \to M$, since the proof for pr_2 proceeds in full analogy. Let $(p, q) \in M \times N$. To show that $\mathrm{pr}_{1*} : T(M \times N) \to TM$ restricts to a surjective linear map on $T_{(p,q)}(M \times N)$, consider the map

$$\begin{array}{cccc} \varphi_q & : & M & \to & M \times N \\ & & x & \mapsto & (x, q) \end{array} . \tag{10.5.2}$$

This map satisfies, of course, $\varphi_q(p) = (p, q)$ and $\mathrm{pr}_1 \circ \varphi_q = \mathrm{id}_M$. From the latter now follows

$$\mathrm{pr}_{1*} \circ \varphi_{q*} = (\mathrm{pr}_1 \circ \varphi_q)_* = \mathrm{id}_{M*} = \mathrm{id}_{TM} . \tag{10.5.3}$$

Given any vector $v \in T_pM$, we can thus define $w = \varphi_{q*}(v) \in T_{(p,q)}(M \times N)$, and it satisfies $\mathrm{pr}_{1*}(w) = v$. Hence, the restriction of pr_{1*} to $T_{(p,q)}(M \times N)$ is surjective. Since this holds for all $(p, q) \in M \times N$, it follows that $\mathrm{pr}_1 : M \times N \to M$ is a submersion. ■

Recall that locally any fiber bundle takes the form of a product, through its local trivializations. Since the differential of a map depends only on its local properties, one may therefore expect that also the projection of a fiber bundle is a submersion. This can be shown as follows.

Theorem 10.5.5. *Let (E, B, π, F) be a fiber bundle. The projection $\pi : E \to B$ is a submersion.*

Proof. There are different possibilities to prove this statement, either by constructing a similar argument as in the proof of the preceding theorem 10.5.4, or by using its result together with the properties of the local trivializations. Here we start with the former. Let $e \in E$ and $\pi(e) = p \in B$. Using theorem 2.3.1, we can construct a local section as follows. Let (U, ϕ) with $p \in U \subset B$ and $\phi : \pi^{-1}(U) \to U \times F$ a local trivialization around p, and construct a local section σ by defining

$$\begin{array}{cccc} \sigma & : & U & \to & E \\ & & x & \mapsto & \phi^{-1}(x, \mathrm{pr}_2(\phi(e))) \end{array} . \tag{10.5.4}$$

This is a local section, since $\phi(e) \in U \times F$, and so $\mathrm{pr}_2(\phi(e)) \in F$, and

$$\pi \circ \sigma = \mathrm{pr}_1 \circ \phi \circ \sigma = \mathrm{id}_U . \tag{10.5.5}$$

Further, it satisfies

$$\sigma(p) = \phi^{-1}(p, \mathrm{pr}_2(\phi(e))) = \phi^{-1}(\pi(e), \mathrm{pr}_2(\phi(e))) = \phi^{-1}(\mathrm{pr}_1(\phi(e)), \mathrm{pr}_2(\phi(e))) = \phi^{-1}(\phi(e)) = e . \tag{10.5.6}$$

Now we can further proceed in full analogy to the proof of theorem 10.5.4. Given $v \in T_pB$, we define $w = \sigma_*(v) \in T_eE$. Since $\pi \circ \sigma = \mathrm{id}_U = \mathrm{id}_M|_U$, it follows that $\pi_*(w) = v$. Since this holds for all $v \in T_pB$, it follows that $\pi_*|_{T_eE}$ is surjective. Further, this can be done for all $e \in E$, and so π is a submersion.

Alternatively, one can also use theorem 10.5.4 directly for the proof. ►. . . ◄ ■

The previous examples exhibit some sense of duality between immersions and submersions. While sections of fiber bundles constitute immersions, projections are submersions. Also note that in order to prove the statements for submersions, we needed significantly more work than for the corresponding statements for immersions. The reason is that while for immersions we could make use of theorem 10.4.2, while the related theorem 10.5.2 for submersions is not helpful here, since we have no surjective map to compose with. Instead, we had to choose a different, non-surjective map for any point we considered, and used the existence of such maps, or constructed them explicitly.

Finally, we come to the coordinate representation of submersions. Here we find that the following theorem holds.

Theorem 10.5.6. *Let M, N be manifolds and $\varphi : M \to N$ a submersion. For every $p \in M$, there exist charts (U, ψ) of M with $p \in U$ and (V, χ) of N with $\varphi(U) \subset V$, such that $\chi \circ \varphi \circ \psi^{-1} : \psi(U) \to \chi(V)$ is given by the ordinary orthogonal projection*

$$(x^1, \ldots, x^n, x^{n+1}, \ldots x^m) \mapsto (x^1, \ldots, x^n)\,, \tag{10.5.7}$$

where $m = \dim M$ and $n = \dim N$.

Proof. ►. . . ◄ ■

Chapter 11

Pullback

11.1 Pullback of functions

While the pushforward transfers objects (vectors) along a map in the same direction as the map points, the pullback works in the opposite direction and transfers objects (sections of bundles) from the target manifold to the source manifold. We have already encountered this in section 2.9 in the discussion of pullback bundles, where sections of arbitrary bundles can be pulled back. For particular cases, it is also possible to consider the pullback to bundles which are not pullback bundles. In fact, there are different notions of such pullbacks, depending on the type of object to which it is applied. The simplest possible case is the pullback of a function.

Definition 11.1.1 (Pullback of a function)**.** Let M and N be manifolds and $\varphi : M \to N$ a smooth map. The *pullback* of a function $f \in C^\infty(N, \mathbb{R})$ to M along φ is the function $\varphi^*(f) = f \circ \varphi \in C^\infty(M, \mathbb{R})$.

It is clear that $\varphi^*(f)$ is a smooth function on M, since the composition of smooth maps is smooth.

11.2 Pullback of covector fields

A slightly more sophisticated type of pullback is defined as follows.

Definition 11.2.1 (Pullback of a covector field)**.** Let M and N be manifolds and $\varphi : M \to N$ a smooth map. The *pullback* of a covector field $\omega \in \Omega^1(N)$ to M along φ is the covector field $\varphi^*(\omega) \in \Omega^1(M)$ such that for all $p \in M$ and $v \in T_pM$ holds

$$\langle v, \varphi^*(\omega)(p) \rangle = \langle \varphi_*(v), \omega(\varphi(p)) \rangle . \tag{11.2.1}$$

Note that there is a fundamental difference between the pullback and the pushforward, besides the fact that they transfer objects in different directions: while the pushforward takes *single* tangent vectors from TM to TN, the pullback takes whole *sections* of T^*N to sections of T^*N. This can be understood as follows.

A map $\varphi : M \to N$ assigns to each point $p \in M$ a point $\varphi(p) \in N$, but this map is in general not surjective or injective. Given a single vector $v \in T_pM$, the pushforward yields a single vector $\varphi_*(v) \in T_{\varphi(p)}N$. However, we cannot use the pushforward and apply it to a vector field $X \in \mathrm{Vect}(M)$ to obtain a vector field on $Y \in \mathrm{Vect}(N)$, because the latter would be a map $Y : N \to TN$ which assigns a unique vector to each $q \in N$. But the pushforward does not yield any vector at points $q \in N$ which lie outside the image of φ. Further, if φ is not injective, it maps different vectors $X(p)$ and $X(p')$ with $\varphi(p) = \varphi(p') = q$ into T_qN, so that $Y(q)$ would not be uniquely defined.

The converse holds for the pullback. We cannot pull a single covector $\alpha \in T_q^*N$ back to M, because q may lie outside the image of φ and thus have no preimage at all, or may have multiple preimages. But if we have a covector field $\omega \in \Omega^1(N)$, which assigns a covector $\omega(q)$ to each point $q \in N$, we can obtain a covector field $\varphi^*(\omega) \in \Omega^1(M)$ as follows. We need to construct a section of T^*M, which assigns to each $p \in M$ a covector $\varphi^*(\omega)(p) \in T_p^*M$. Here we make use of the fact that T_pM and T_p^*M are dual vector spaces, so that we can identify such a covector with a linear function on T_pM. To construct such a function, we take a vector $v \in T_pM$ and push it (linearly) to a vector $\varphi_*(v) \in T_{\varphi(p)}N$. Now we use the covector $\omega(\varphi(p)) \in T^*_{\varphi(p)}N$, which is a linear function on $T_{\varphi(p)}N$. This is exactly the construction given in the definition of the pullback.

To illustrate this definition we write the pullback in coordinates. Let (x^a) be coordinates on M and (y^α) coordinates on N, as in the previous section. Using these coordinates a covector field $\omega \in \Omega^1(N)$ takes the form $\omega_\alpha \mathrm{d}y^\alpha$, while a vector $v \in T_xM$ can be written as $v = v^a\partial_a$. The definition of the pullback then reads

$$\langle v, \varphi^*(\omega)(x)\rangle = \langle \varphi_*(v), \omega(\varphi(x))\rangle = \varphi_*(v)^\alpha \omega_\alpha(y(x)) = v^a \frac{\partial y^\alpha}{\partial x^a}\omega_\alpha(y(x)) \,, \tag{11.2.2}$$

so that $\varphi^*(\omega)$ can be written in coordinates in the form

$$\varphi^*(\omega)(x) = \omega_\alpha(y(x))\frac{\partial y^\alpha}{\partial x^a}\mathrm{d}x^a \,. \tag{11.2.3}$$

11.3 Pullback of differential forms

We now have pullbacks of 0-forms (real functions) and 1-forms (covector fields) on N. One may already guess that this procedure can be extended to arbitrary k-forms on N. For this purpose, recall that an element of $\Lambda^k T_q^*N$ can be viewed as an alternating multilinear form on T_qN, i.e., a function from $T_qN \times \ldots \times T_qN$ to $\mathbb{R}$ which is linear in each argument and totally antisymmetric with respect to permutations of its arguments. With this in mind we can define the pullback of a differential form as follows.

Definition 11.3.1 (Pullback of a differential form)**.** Let M and N be manifolds and $\varphi : M \to N$ a smooth map. The *pullback* of a k-form $\omega \in \Omega^k(N)$ to M along φ is the k-form $\varphi^*(\omega) \in \Omega^k(M)$ such that for all $p \in M$ and $v_1, \ldots, v_k \in T_pM$ holds

$$\varphi^*(\omega)(p)(v_1, \ldots, v_k) = \omega(\varphi(p))(\varphi_*(v_1), \ldots, \varphi_*(v_k)) \,. \tag{11.3.1}$$

Again one easily checks that this definition indeed yields a k-form on M. Also the coordinate expression can be easily derived. Following the same procedure as above one easily sees that

$$\varphi^*(\omega_{\alpha_1\cdots\alpha_k}\mathrm{d}y^{\alpha_1} \wedge \ldots \wedge \mathrm{d}y^{\alpha_k}) = \omega_{\alpha_1\cdots\alpha_k}\frac{\partial y^{\alpha_1}}{\partial x^{a_1}} \cdots \frac{\partial y^{\alpha_k}}{\partial x^{a_k}}\mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k} \,. \tag{11.3.2}$$

A bit less obvious are the following very useful properties of the pullback of differential forms.

Theorem 11.3.1. *Let M and N be manifolds and $\varphi : M \to N$ a smooth map. For any differential forms $\alpha \in \Omega^k(N), \beta \in \Omega^l(N)$ on N the pullback satisfies*

$$\varphi^*(\alpha) \wedge \varphi^*(\beta) = \varphi^*(\alpha \wedge \beta) \quad \text{and} \quad \mathrm{d}(\varphi^*(\alpha)) = \varphi^*(\mathrm{d}\alpha)\,. \tag{11.3.3}$$

Proof. ▶. . . ◀ ■

The proof is rather lengthy, but simple, so we will not discuss it here.

11.4 Pullback of covariant tensor fields

We finally generalize the pullback even further. In a similar way as an element of $\Lambda^k T_q^* N$ can be regarded as an alternating multilinear form on $T_q N$, an element of $\bigotimes^k T_q^* N$ corresponds to a (general) multilinear form on $T_q N$. This allows us to extend the pullback to covariant tensor fields, i.e., tensor fields of type $(0, k)$. In fact, the definition is identical to the case of a differential form.

Definition 11.4.1 (Pullback of a covariant tensor field)**.** Let M and N be manifolds and $\varphi : M \to N$ a smooth map. The *pullback* of a covariant tensor field $A \in \Gamma(T_k^0 N)$ to M along φ is the covariant tensor field $\varphi^*(A) \in \Gamma(T_k^0 M)$ such that for all $p \in M$ and $v_1, \ldots, v_k \in T_p M$ holds

$$\varphi^*(A)(p)(v_1, \ldots, v_k) = A(\varphi(p))(\varphi_*(v_1), \ldots, \varphi_*(v_k))\,. \tag{11.4.1}$$

It should be clear now that this is indeed a tensor field of type $(0, k)$ on M and that its coordinate expression is given by

$$\varphi^*(A_{\alpha_1 \cdots \alpha_k} \mathrm{d}y^{\alpha_1} \otimes \ldots \otimes \mathrm{d}y^{\alpha_k}) = A_{\alpha_1 \cdots \alpha_k} \frac{\partial y^{\alpha_1}}{\partial x^{a_1}} \cdots \frac{\partial y^{\alpha_k}}{\partial x^{a_k}} \mathrm{d}x^{a_1} \otimes \ldots \otimes \mathrm{d}x^{a_k}\,. \tag{11.4.2}$$

Chapter 12

Diffeomorphisms and coordinate transformations

12.1 Pullback along diffeomorphisms

We have seen in the previous sections that the ways we can transfer objects along an arbitrary smooth map $\varphi : M \to N$ are limited since φ is in general neither injective nor surjective. We can remove these limitations by taking φ to be a diffeomorphism, i.e., a bijective map whose inverse is again smooth. In this case the differential $\varphi_* : TM \to TN$ becomes a vector bundle isomorphism, and we can make use of various derived vector bundle isomorphisms to transfer single tensors and tensor fields freely between both manifolds. This will be done in this section. We start be defining the pullback of a vector field.

Definition 12.1.1 (Pullback of a vector field)**.** Let M and N be manifolds and $\varphi : M \to N$ a diffeomorphism. The *pullback* of a vector field $X \in \mathrm{Vect}(N)$ to M along φ is the vector field $\varphi^*(X) \in \mathrm{Vect}(M)$ such that $\varphi^*(X)(p) = \varphi_*^{-1}(X(\varphi(p)))$ for each $p \in M$.

In the definition we have explicitly used the inverse of φ_*, which should remind us that this construction is valid only if φ is a diffeomorphism. In coordinates (x^a) on M and (y^a) on N (where we now use the same type of letters for the indices, because diffeomorphic manifolds necessarily have the same dimension) we find that

$$\varphi^*(X) = X^a \frac{\partial x^b}{\partial y^a} \partial_b\,, \tag{12.1.1}$$

which follows from the rule for the derivative of inverse functions on $\mathbb{R}^n$. Since we can now pull back both vector and covector fields, we can also pull back arbitrary tensor fields. The definition is as follows.

Definition 12.1.2 (Pullback of a tensor field)**.** Let M and N be manifolds and $\varphi : M \to N$ a diffeomorphism. The *pullback* of tensor fields on N to tensor fields on M is defined as the linear function $\varphi^* : \Gamma(T^r_sN) \to \Gamma(T^r_sM)$ that for any r vector fields $X_1, \ldots, X_r \in \mathrm{Vect}(N)$ and s 1-forms $\omega_1, \ldots, \omega_s \in \Omega^1(N)$ holds

$$\varphi^*(X_1 \otimes \ldots \otimes X_r \otimes \omega_1 \otimes \ldots \otimes \omega_s) = \varphi^*(X_1) \otimes \ldots \otimes \varphi^*(X_r) \otimes \varphi^*(\omega_1) \otimes \ldots \otimes \varphi^*(\omega_s)\,. \tag{12.1.2}$$

Using the definitions of the pullbacks of various tensor fields, one can now derive the following property.

Theorem 12.1.1. *The pullback along a diffeomorphism commutes with the contraction of a tensor field, i.e.,* $\varphi^*(\operatorname{tr}^k_l T) = \operatorname{tr}^k_l \varphi^*(T)$.

Proof. We use the same notation as in definition 12.1.2 and start with the case $T = X \otimes \omega$, where $X \in \operatorname{Vect}(N)$ and $\omega \in \Omega^1(N)$. In this case the only possible contraction has $k = l = 1$ and reads $\operatorname{tr}^1_1(X \otimes \omega) = \langle X, \omega \rangle$. For every $p \in M$ we have $\varphi^*(X)(p) = \varphi_*^{-1}(X(\varphi(p)))$. By the definition of the pullback, we then have

$$\begin{aligned}
\operatorname{tr}^1_1 \varphi^*(X \otimes \omega)(p) &= \langle \varphi^*(X)(p), \varphi^*(\omega)(p) \rangle \\
&= \langle X(\varphi(p)), \omega(\varphi(p)) \rangle \\
&= (\operatorname{tr}^1_1(X \otimes \omega))(\varphi(p)) \\
&= \varphi^*(\operatorname{tr}^1_1(X \otimes \omega))(p) \,,
\end{aligned} \tag{12.1.3}$$

where the latter follows from the pullback of a scalar function. Since this holds for all $p \in M$, we conclude

$$\operatorname{tr}^1_1 \varphi^*(X \otimes \omega) = \varphi^*(\operatorname{tr}^1_1(X \otimes \omega)) \,. \tag{12.1.4}$$

This is now easily extended to the contraction of a general tensor product, since

$$\operatorname{tr}^k_l(T \otimes X \otimes U \otimes \omega \otimes V) = \langle X, \omega \rangle (T \otimes U \otimes V) \,, \tag{12.1.5}$$

where k and l are chosen such that the contraction is over X and ω, since

$$\begin{aligned}
\operatorname{tr}^k_l \varphi^*(T \otimes X \otimes U \otimes \omega \otimes V) &= \operatorname{tr}^k_l(T' \otimes X' \otimes U' \otimes \omega' \otimes V') \\
&= \langle X', \omega' \rangle (T' \otimes U' \otimes V') \\
&= \langle X, \omega \rangle' (T' \otimes U' \otimes V') \\
&= \varphi^*(\langle X, \omega \rangle (T \otimes U \otimes V)) \\
&= \varphi^*(\operatorname{tr}^k_l(T \otimes X \otimes U \otimes \omega \otimes V)) \,,
\end{aligned} \tag{12.1.6}$$

where we abbreviated $T' = \varphi^*(T)$, and analogously for the other fields. Finally, since the pullback is linear by definition, this result holds for any linear combination of tensor products, and hence for any tensor field. ■

This can also be seen from the coordinate expression of the pullback. In coordinates we find for a tensor field $A \in \Gamma(T^r_s N)$ the pullback

$$\begin{aligned}
&\varphi^*(A^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} \partial'_{a_1} \otimes \ldots \otimes \partial'_{a_r} \otimes \mathrm{d}y^{b_1} \otimes \ldots \otimes \mathrm{d}y^{b_s}) \\
&\quad = A^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} \frac{\partial x^{c_1}}{\partial y^{a_1}} \cdots \frac{\partial x^{c_r}}{\partial y^{a_r}} \frac{\partial y^{b_1}}{\partial x^{d_1}} \cdots \frac{\partial y^{b_s}}{\partial x^{d_s}} \partial_{c_1} \otimes \ldots \otimes \partial_{c_r} \otimes \mathrm{d}x^{d_1} \otimes \ldots \otimes \mathrm{d}x^{d_s} \,,
\end{aligned} \tag{12.1.7}$$

where we wrote (∂'_a) for the coordinate basis of $T_y N$. For the contraction, using

$$\frac{\partial x^c}{\partial y^a} \frac{\partial y^b}{\partial x^c} = \delta^b_a \,, \tag{12.1.8}$$

one sees that it commutes with the pullback.

12.2 Coordinate transformations

The transformations using diffeomorphisms shown in the previous section entail the following two special cases if the source and target manifolds are the same, $M = N$:

- If we can use a single chart (U, ψ) to describe the diffeomorphism (for example, if $U = M$ is the whole manifold or if φ maps U to itself), then we may use the same coordinates (x^a) to describe points and their images. A point p with coordinates $\psi^a(p) = x^a$ is mapped to a (in general different) point $p' = \varphi(p)$ with coordinates $\psi^a(p') = x'^a$. We can apply this diffeomorphism to a tensor field A and obtain its pullback $A' = \varphi^*(A)$. The components of this (in general different) tensor field are given by

$$A'^{c_1 \cdots c_r}{}_{d_1 \cdots d_s}(p) = A^{a_1 \cdots a_r}{}_{b_1 \cdots b_s}(p') \frac{\partial x^{c_1}}{\partial x'^{a_1}} \cdots \frac{\partial x^{c_r}}{\partial x'^{a_r}} \frac{\partial x'^{b_1}}{\partial x^{d_1}} \cdots \frac{\partial x'^{b_s}}{\partial x^{d_s}} \,. \tag{12.2.1}$$

Since points change, $p \neq p'$ in general, this transformation is usually called an *active diffeomorphism*.

- Another special case is obtained if we use two different charts (U, ψ) and (V, χ), and hence different coordinates (x^a) and (y^a), to describe tensor fields at the same point $p \in U \cap V$, while the diffeomorphism we consider is the identity $\varphi = \mathrm{id}_M$. In this case points and tensor fields stay the same, since $\varphi(p) = p$ and $\varphi^* A = A$, but their coordinate expressions change. This means that we express tensor fields in a different basis

$$\partial'_{a_1} \otimes \ldots \otimes \partial'_{a_r} \otimes \mathrm{d}y^{b_1} \otimes \ldots \otimes \mathrm{d}y^{b_s} = \frac{\partial x^{c_1}}{\partial y^{a_1}} \cdots \frac{\partial x^{c_r}}{\partial y^{a_r}} \frac{\partial y^{b_1}}{\partial x^{d_1}} \cdots \frac{\partial y^{b_s}}{\partial x^{d_s}} \partial_{c_1} \otimes \ldots \otimes \partial_{c_r} \otimes \mathrm{d}x^{d_1} \otimes \ldots \otimes \mathrm{d}x^{d_s} \tag{12.2.2}$$

at the same point p. Such a pure coordinate transformation is also called a *passive diffeomorphism*, since it does not change points or tensor fields, but only their coordinate description.

We see that both types of transformations are described by the same general formula (12.1.7). This is not very surprising. Given an active diffeomorphism $\varphi : M \to M$ and a chart (U, ψ), we may define a new chart (V, χ) by $\psi = \chi \circ \varphi$, such that $\psi(p) = \chi(p')$. Then we find that the coordinate expression of a tensor field A at a point p' in the chart (V, χ) is exactly the same as that of its pullback $A' = \varphi^* A$ at the original point p, but using the different chart (U, ψ), which follows from the fact that both are given by $(\psi^{-1})^* A' = (\psi^{-1})^* \varphi^* A = (\chi^{-1})^* A$.

12.3 Background independence

The notions of active and passive diffeomorphisms often lead to confusion, in particular regarding the question whether a given physical theory is invariant under active or passive coordinate transformations. With the notions introduced in the previous section, we may clarify this issue.

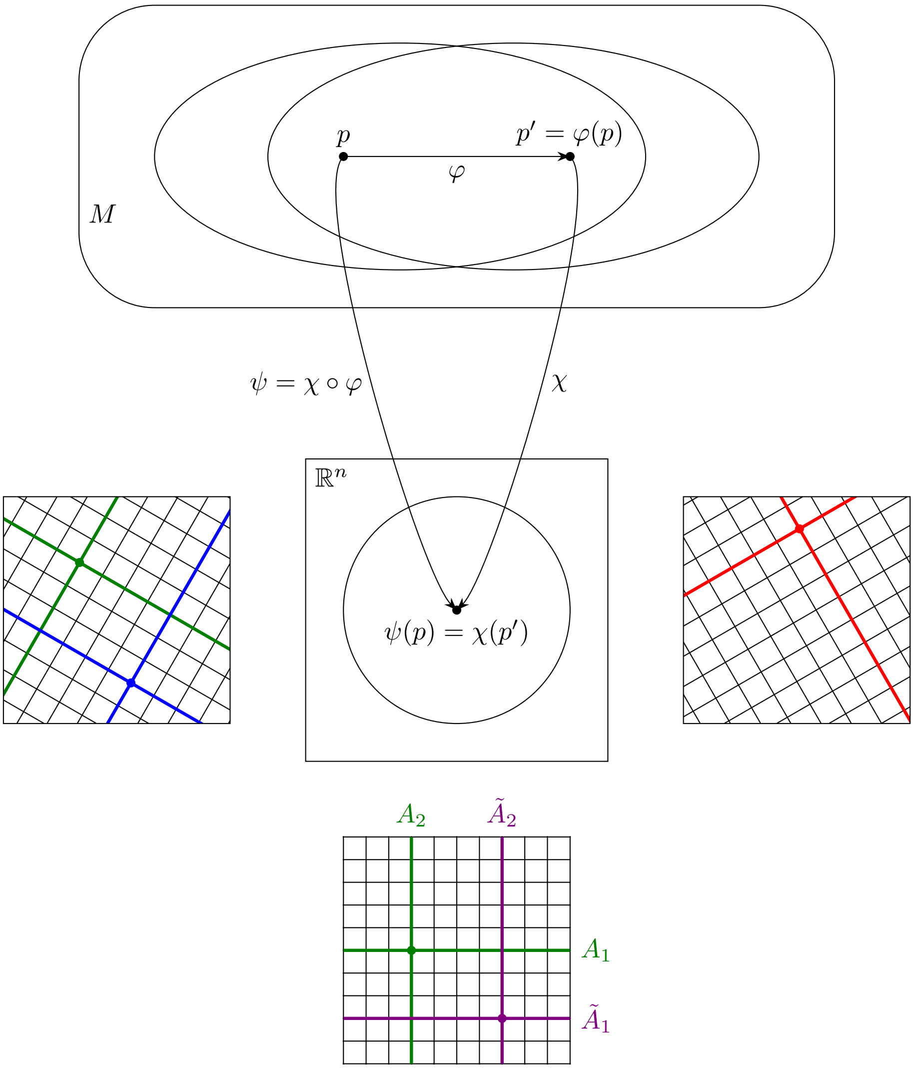


Figure 12.1: Relation between active and passive diffeomorphisms.

Chapter 13

Submanifolds

13.1 Immersed submanifolds

13.2 Embedded submanifolds

Definition 13.2.1 (Embedded submanifold)**.** Let M be a manifold of dimension n. A subset $S \subset M$ is called an *embedded submanifold* of dimension k if for each $p \in M$ there exists a chart (U, ϕ) of M with $p \in U$ such that $S \cap U = \phi^{-1}(\mathbb{R}^k)$, where we identify $\mathbb{R}^k = \{v^i \mathbb{e}_i \in \mathbb{R}^n \mid v^i = 0 \forall n > k\}$.

Theorem 13.2.1. *Let M, N be manifolds and $\varphi : M \to N$ a map of constant rank. Then for each $q \in N$, $\varphi^{-1}(q)$ is an embedded submanifold of M.*

Proof. ►...◄ ■

Theorem 13.2.2. *Let M, N be manifolds and $\varphi : M \to N$ a submersion. Then for each $q \in N$, $\varphi^{-1}(q)$ is an embedded submanifold of M.*

Proof. ►...◄ ■

Theorem 13.2.3. *Let $\pi : E \to B$ be a fiber bundle. For each $p \in B$, the fiber $E_p = \pi^{-1}(p)$ is an embedded submanifold of E.*

Proof. ►...◄ ■

13.3 Bundles over submanifolds

13.4 Foliations

Chapter 14

Manifolds with boundary and corners

14.1 Manifolds with boundary

Definition 14.1.1 (Chart with boundary)**.** Let M be a set. A *chart (with boundary)* of dimension $n \in \mathbb{N}$ on M is a pair (U, ϕ), where $U \subset M$ is a subset of M and $\phi : U \to \{\mathbb{R}^n | x^1 \geq 0\}$ is an injective function, such that the image $\phi(U) \subset \{\mathbb{R}^n | x^1 \geq 0\}$ is open.

Definition 14.1.2 (Interior and boundary)**.** Let M be a manifold with boundary. A point $p \in M$ is called:

1. *interior point*, if there exists a chart (U, ϕ) such that $\phi^1(u) \neq 0$,
2. *boundary point*, if there exists a chart (U, ϕ) such that $\phi^1(u) = 0$.

The set of all interior points of M is called its *interior* and denoted M°, while the set of all boundary points is called its *boundary* and denoted ∂M.

Theorem 14.1.1. *If a point $p \in M$ of a manifold with boundary M is a boundary (interior) point with respect to some chart, it is a boundary (interior) point with respect to all charts in which it is contained.*

Proof. ▶ . . . ◀ ■

Theorem 14.1.2. *A manifold with boundary M is the disjoint union of its interior and its boundary, $M = M^\circ \uplus \partial M$.*

Proof. ▶ . . . ◀ ■

14.2 Manifolds with corners

Definition 14.2.1 (Chart with corners)**.** Let M be a set. A *chart (with corners)* of dimension $n \in \mathbb{N}$ on M is a pair (U, ϕ), where $U \subset M$ is a subset of M and $\phi : U \to \{\mathbb{R}^n | x^1 \geq 0 \wedge \dots x^n \geq 0\}$ is an injective function, such that the image $\phi(U) \subset \{\mathbb{R}^n | x^1 \geq 0 \wedge \dots x^n \geq 0\}$ is open.

Chapter 15

Lie groups and actions

15.1 Lie groups

In this section we will introduce manifolds which carry an additional algebraic structure, namely that of a group. In order to work with this structure, it must be compatible with the manifold structure. We make this precise in the following definition.

Definition 15.1.1 (Lie group)**.** A *Lie group* is a manifold G which carries the structure of a group, such that the group multiplication $\cdot : G \times G \to G$ and the inverse $\bullet^{-1} : G \to G$ are smooth maps.

This compatibility condition is a bit similar to the compatibility condition for vector bundles, where we wanted the vector space operations (addition and scalar multiplication) to be smooth operations. There are many examples for Lie groups which frequently appear in physics:

Example 15.1.1. The group $(\mathbb{R}, +)$ of real numbers with the addition as group operation is a Lie group of dimension 1.

Example 15.1.2. The complex numbers $z \in \mathbb{C}$ with $|z| = 1$ and group operation the multiplication is a Lie group of dimension 1 which is diffeomorphic to the circle S^1. It is also denoted $\mathrm{U}(1)$.

Example 15.1.3. The following matrix groups for $n, p, q \in \mathbb{N}$ are Lie groups, where the group multiplication is given by matrix multiplication:

- The (real) *general linear group* $\mathrm{GL}(n, \mathbb{R})$ (or simply $\mathrm{GL}(n)$) of real invertible $n \times n$ matrices is a Lie group of dimension n^2.
- The complex *general linear group* $\mathrm{GL}(n, \mathbb{C})$ of complex invertible $n \times n$ matrices is a Lie group of dimension $2n^2$.
- The (real) *special linear group* $\mathrm{SL}(n, \mathbb{R})$ (or simply $\mathrm{SL}(n)$) of real invertible $n \times n$ matrices with determinant 1 is a Lie group of dimension $n^2 - 1$.
- The complex *special linear group* $\mathrm{SL}(n, \mathbb{C})$ of complex invertible $n \times n$ matrices with determinant 1 is a Lie group of dimension $2n^2 - 2$.

- The *orthogonal group* $\mathrm{O}(n)$ of real $n \times n$ matrices such that $AA^t = \mathbb{1}$ is a Lie group of dimension $n(n-1)/2$.
- The *indefinite orthogonal group* $\mathrm{O}(p,q)$ with $n = p+q$ of real $n \times n$ matrices such that $A\eta A^t = \eta$, where
$$\eta = \mathrm{diag}(\underbrace{-1,\ldots,-1}_{p\text{ times}},\underbrace{1,\ldots,-1}_{q\text{ times}})\,, \tag{15.1.1}$$
is a Lie group of dimension $n(n-1)/2$.
- The *special orthogonal group* $\mathrm{SO}(n)$ of real $n \times n$ matrices with determinant 1 such that $AA^t = \mathbb{1}$ is a Lie group of dimension $n(n-1)/2$.
- The *indefinite special orthogonal group* $\mathrm{SO}(p,q)$ with $n = p+q$ of real $n \times n$ matrices with determinant 1 such that $A\eta A^t = \eta$ is a Lie group of dimension $n(n-1)/2$.
- The (real) *symplectic group* $\mathrm{Sp}(2n,\mathbb{R})$ (or simply $\mathrm{Sp}(n)$) of real $2n \times 2n$ matrices such that $A\Omega A^t = \Omega$, where
$$\Omega = \begin{pmatrix} 0 & \mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{pmatrix}\,, \tag{15.1.2}$$
is a Lie group of dimension $n(2n+1)$.
- The complex *symplectic group* $\mathrm{Sp}(2n,\mathbb{C})$ of complex $2n \times 2n$ matrices such that $A\Omega A^t = \Omega$ is a Lie group of dimension $2n(2n+1)$.
- The *unitary group* $\mathrm{U}(n)$ of complex $n \times n$ matrices such that $AA^\dagger = \mathbb{1}$ is a Lie group of dimension n^2.
- The *special unitary group* $\mathrm{SU}(n)$ of complex $n \times n$ matrices with determinant 1 such that $AA^\dagger = \mathbb{1}$ is a Lie group of dimension $n^2 - 1$.

Note that the dimensions listed here are the *real* dimensions, i.e., the dimensions of G as real manifolds. The group $\mathrm{SO}(1,3)$ is also called the *Lorentz group*.

The aforementioned matrix groups play a particularly important role in physics. Of comparable importance are their inhomogeneous extensions.

Example 15.1.4. Let G be one of the matrix groups defined above, whose elements are $n \times n$ matrices with entries in $\mathbb{K}$ (either $\mathbb{R}$ or $\mathbb{C}$). Then G acts from the left on $\mathbb{K}^n$ via multiplication from the left. On the space $\mathbb{K}^n \times G$ define a product such that
$$\begin{array}{cccccc} \cdot & : & (\mathbb{K}^n \times G) \times (\mathbb{K}^n \times G) & \to & \mathbb{K}^n \times G \\ & & ((v,A),(w,B)) & \mapsto & (v + Aw, AB) \end{array}\,. \tag{15.1.3}$$
We may also express this multiplication law by writing the group elements as block matrices in the form
$$\begin{pmatrix} A & v \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} B & w \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} AB & v + Aw \\ 0 & 1 \end{pmatrix}\,. \tag{15.1.4}$$
The resulting group is called the *semidirect product*, denoted $\mathbb{K}^n \rtimes G$, and is also a Lie group. For the matrix groups mentioned above, it is conventionally denoted by prepending the letter "I" (for inhomogeneous) to the name of the group, e.g., the group $\mathrm{ISO}(1,3) = \mathbb{R}^4 \rtimes \mathrm{SO}(1,3)$ is called the *Poincaré group*.

15.2 Lie group homomorphisms

In order to relate different Lie groups to each other, we need the same compatibility condition for homomorphisms between Lie groups.

Definition 15.2.1 (Lie group homomorphism / isomorphism)**.** Let G, H be Lie groups. A *Lie group homomorphism* from G to H is a smooth map $\varphi : G \to H$ such that $\varphi(gg') = \varphi(g)\varphi(g')$ for all $g, g' \in G$. If it is also a diffeomorphism, it is called a *Lie group isomorphism*.

Note that it is sufficient to demand compatibility with the group multiplications and smooth structures; the following two properties then follow.

Theorem 15.2.1. *Let G, H be Lie groups and $\varphi : G \to H$ a Lie group homomorphism. Then $\varphi(e_G) = e_H$ and $\varphi(g^{-1}) = \varphi(g)^{-1}$ for all $g \in G$.*

Proof. First, from the property of the unit elements follows

$$e_H = \varphi(g)^{-1}\varphi(g) = \varphi(g)^{-1}\varphi(ge_G) = \varphi(g)^{-1}\varphi(g)\varphi(e_G) = \varphi(e_G)\,, \tag{15.2.1}$$

and thus further

$$\varphi(g)^{-1} = \varphi(g)^{-1}e_H = \varphi(g)^{-1}\varphi(e_G) = \varphi(g)^{-1}\varphi(gg^{-1}) = \varphi(g)^{-1}\varphi(g)\varphi(g^{-1}) = \varphi(g^{-1}) \tag{15.2.2}$$

for all $g \in G$. ∎

There are numerous homomorphisms and isomorphisms between the groups given in the examples above.

Example 15.2.1. The map $\varphi : \mathbb{R} \to \mathrm{U}(1), x \mapsto e^{ix}$ is a Lie group homomorphism.

Example 15.2.2. The map $\varphi : \mathrm{U}(1) \to \mathrm{SO}(2)$ defined by

$$\varphi(z) = \begin{pmatrix} \mathrm{Re}(z) & \mathrm{Im}(z) \\ -\mathrm{Im}(z) & \mathrm{Re}(z) \end{pmatrix} \tag{15.2.3}$$

is a Lie group isomorphism.

Lie group isomorphisms have various applications. A particularly interesting class of isomorphisms arise when both groups are identical, $G_1 = G_2 = G$. Such kind of isomorphisms deserve their own name.

Definition 15.2.2 (Lie group automorphism)**.** Let G be a Lie group. An *automorphism* of G is a Lie group isomorphism $\varphi : G \to G$. The group of all Lie group automorphisms of G is denoted $\mathrm{Aut}(G)$.

We remark that under certain circumstances $\mathrm{Aut}(G)$ also has the structure of a Lie group, but this is not always the case. The following class of automorphisms will be relevant for our later constructions:

Definition 15.2.3 (Inner automorphism)**.** Let G be a Lie group and $g \in G$. The *inner automorphism* $\alpha_g \in \mathrm{Aut}(G)$ induced by g is defined such that for all $h \in G$ holds $\alpha_g(h) = ghg^{-1}$.

One easily checks that indeed for all $g \in G$ the map $\alpha_g : G \to G$ is an automorphism, i.e., it satisfies $\alpha_g(hh') = \alpha_g(h)\alpha_g(h')$ for all $h, h' \in G$.

15.3 Lie group actions

A familiar concept from algebra is that of the *action* of a group on a set. Since we are working with Lie groups here, we are in particular interested how a Lie group can act on a manifold. Again we demand compatibility of the differentiable and algebraic structures, as in the following definition.

Definition 15.3.1 (Lie group action)**.** Let G be a Lie group with unit element $e \in G$ and M a manifold. A *left Lie group action* is a smooth map $\phi : G \times M \to M$ such that $\phi(e, x) = x$ and $\phi(gh, x) = \phi(g, \phi(h, x))$ for all $g, h \in G$ and $x \in M$. A *right Lie group action* is a smooth map $\theta : M \times G \to M$ such that $\theta(x, e) = x$ and $\theta(x, gh) = \theta(\theta(x, g), h)$ for all $g, h \in G$ and $x \in M$. We also introduce the notations

$$\phi_x : \begin{array}{ccc} G & \to & M \\ g & \mapsto & \phi(g,x) \end{array}, \qquad \phi^g : \begin{array}{ccc} M & \to & M \\ x & \mapsto & \phi(g,x) \end{array}, \tag{15.3.1a}$$

$$\theta_x : \begin{array}{ccc} G & \to & M \\ g & \mapsto & \theta(x,g) \end{array}, \qquad \theta^g : \begin{array}{ccc} M & \to & M \\ x & \mapsto & \theta(x,g) \end{array} \tag{15.3.1b}$$

for all $g \in G$ and $x \in M$.

We also say that a group G *acts from the left / right* on a manifold M. The following statement follows immediately from the definition above.

Theorem 15.3.1. *Let $\phi : G \times M \to M$ be a left Lie group action. For each $g \in G$ the map $\phi^g : x \mapsto \phi(g, x)$ is a diffeomorphism on M with inverse given by $(\phi^g)^{-1} = \phi^{g^{-1}}$. The same holds for right Lie group actions.*

Proof. The maps $M \to G \times M, x \mapsto (g, x)$ and $\phi : G \times M \to M$, and hence also their composition ϕ^g is smooth. Further,

$$(\phi^{g^{-1}} \circ \phi^g)(x) = \phi(g^{-1}, \phi(g, x)) = \phi(g^{-1}g, x) = \phi(e, x) = x\,, \tag{15.3.2}$$

and so $(\phi^g)^{-1} = \phi^{g^{-1}}$. Exchanging g and g^{-1}, one finds that also $\phi^{g^{-1}}$ is smooth. Hence, ϕ^g is a diffeomorphism. ∎

We further distinguish between different types of Lie group actions.

Definition 15.3.2 (Types of Lie group actions)**.** Let G be a Lie group and M a manifold. A left Lie group action $\phi : G \times M \to M$ is called ...

- ...*transitive* if for all $x, y \in M$ there exists a $g \in G$ such that $\phi(g, x) = y$.
- ...*effective* (or *faithful*) if for all distinct $g, h \in G$ there exists $x \in M$ such that $\phi(g, x) \neq \phi(h, x)$.
- ...*free* if for all distinct $g, h \in G$ and for all $x \in M$ holds $\phi(g, x) \neq \phi(h, x)$.

The same naming is used for right Lie group actions.

It follows immediately that every free action is also effective. Of course there are many examples of group actions which appear in physics.

Example 15.3.1. Each of the matrix groups G from example 15.1.3 in the previous section acts from the left on $\mathbb{R}^n$ (or $\mathbb{C}^n$ in the case of matrix groups over the complex numbers) via multiplication. This group action is effective, but neither transitive nor free.

Example 15.3.2. Every Lie group G acts on itself from the left by left multiplication $\phi(g, x) = gx$ and from the right by right multiplication $\theta(x, g) = xg$. Both actions are free and transitive.

The last example is of particular interest, because it is a property of every Lie group. The diffeomorphisms obtained from these actions deserve their own names.

Definition 15.3.3 (Translation maps)**.** Let G be a Lie group. For $g \in G$ the *left translation* is the map $L_g : G \to G, h \mapsto gh$, while the right translation is the map $R_g : G \to G, h \mapsto hg$.

We further introduce the following concepts, which will help us analyze the structure of Lie group actions. Here we define them for left Lie group actions, but their definition for right Lie group actions is completely analogous.

Definition 15.3.4 (Invariant subset)**.** Let $\phi : G \times M \to M$ be a left Lie group action. A subset $U \subseteq M$ is called *invariant* under ϕ if and only if $\phi(g, x) \in U$ for every $g \in G$ and $x \in U$.

Of particular importance is a special case of the invariant subset, known as the orbit. It is defined as follows.

Definition 15.3.5 (Orbit)**.** Let $\phi : G \times M \to M$ be a left Lie group action. For $x \in M$ the *orbit* is the set

$$\mathcal{O}_x^\phi = \{\phi(g, x), g \in G\} \subseteq M\,. \tag{15.3.3}$$

We clarify the special role of the orbit as an invariant subset in the following few statements, which are closely related to each other. First, we show that orbits are invariant subsets.

Theorem 15.3.2. *Let $\phi : G \times M \to M$ be a left Lie group action. For every $x \in M$ the orbit $\mathcal{O}_x^\phi \subseteq M$ is an invariant subset under ϕ.*

Proof. Let $y \in \mathcal{O}_x^\phi$. Then, by definition of the orbit, there exists $g \in G$ such that $y = \phi(g, x)$. For $g' \in G$, we then have

$$\phi(g', y) = \phi(g', \phi(g, x)) = \phi(g'g, x) \in \mathcal{O}_x^\phi \,. \tag{15.3.4}$$

This holds for all $y \in \mathcal{O}_x^\phi$ and $g' \in G$, and so $\mathcal{O}_x^\phi$ is an invariant subset. ■

In fact, we can even use orbits to classify invariant subsets. This can be stated as follows.

Theorem 15.3.3. *Let $\phi : G \times M \to M$ be a left Lie group action. A set $U \subseteq M$ is an invariant subset under ϕ if and only if $\mathcal{O}_x^\phi \subseteq U$ for every $x \in U$.*

Proof. By definition, U is an invariant subset if and only if $\phi(g, x) \in U$ for all $g \in G$ and $x \in U$, hence if and only if for all $x \in U$ holds:

$$U \supseteq \{\phi(g, x), g \in G\} = \mathcal{O}_x^\phi \,. \tag{15.3.5}$$

■

The previous statement suggests that a group action allows us to divide M into distinct orbits. This can most comprehensively be stated as follows.

Theorem 15.3.4. *A group action $\phi : G \times M \to M$ equips M with an equivalence relation $\overset{\phi}{\sim}$ defined such that $x \overset{\phi}{\sim} y$ if and only if $x \in \mathcal{O}_y^\phi$.*

Proof. According to the definition above, we write $x \overset{\phi}{\sim} y$ if and only if there exists $g \in G$ such that $x = \phi(g, y)$. Obviously, $\overset{\phi}{\sim}$ is reflexive, since $x = \phi(e, x)$, and so $x \overset{\phi}{\sim} x$. Further, it is symmetric, since $x = \phi(g, y)$ implied $y = \phi(g^{-1}, x)$ and vice versa. Finally, it is transitive, since from $x = \phi(g, y)$ and $y = \phi(h, z)$ for $x, y, z \in M$ and $g, h \in G$ follows

$$x = \phi(g, \phi(h, z)) = \phi(gh, z) \,, \tag{15.3.6}$$

and so $x \overset{\phi}{\sim} y$ and $y \overset{\phi}{\sim} z$ implies $x \overset{\phi}{\sim} z$. ■

So far we have focused only on the fact that ϕ is a group action, but not made use of the additional structure implied by the condition that it is a Lie group action. We now take this additional structure into account as well. This allows us to show the following.

Theorem 15.3.5. *Let $\phi : G \times M \to M$ be a left Lie group action. For every $x \in M$ the orbit $\mathcal{O}_x^\phi \subseteq M$ is an immersed submanifold.*

Proof. ►...◄ ■

We now come back to definition 15.3.2 of different classes of Lie group actions. It turns out that we can also make use of the notion of the orbit here.

Theorem 15.3.6. *A left Lie group action $\phi : G \times M \to M$ is transitive if and only if the orbit satisfies $\mathcal{O}_x^\phi = M$ for some (and hence for all) $x \in M$.*

Proof. ϕ is transitive if for all $x, y \in M$ there exists $g \in G$ such that $y = \phi(g, x)$, and hence $y \in \mathcal{O}_x^\phi$. ■

Next, we come to a notion which is in some sense dual to that of the orbit. While the orbit answers the questions which points can be reached from a given point $x \in M$ by acting with all group elements, the following notion answers the questions which group elements leave the point x unchanged.

Definition 15.3.6 (Stabilizer). Let $\phi : G \times M \to M$ be a left Lie group action. For $x \in M$ the *stabilizer* is the subgroup

$$\mathcal{S}_x^\phi = \{g \in G \,|\, \phi(g, x) = x\} \subseteq G\,. \tag{15.3.7}$$

As the definition already suggests, the stabilizer is a subgroup. We prove this together with another property, which follows from the topological properties of Lie group actions.

Theorem 15.3.7. *Let $\phi : G \times M \to M$ be a left Lie group action. For every $x \in M$ the stabilizer $\mathcal{S}_x^\phi \subseteq M$ is a closed subgroup of G.*

Proof. For fixed $x \in M$, the map $\phi_x : G \to M, g \mapsto \phi(g, x)$ is continuous. Further, $\{x\} \subset M$ is closed. Hence, $\mathcal{S}_x^\phi = \phi_x^{-1}(\{x\})$ is also closed. Further, for $g, g' \in \mathcal{S}_x^\phi$, one has

$$\phi(gg', x) = \phi(g, \phi(g', x)) = \phi(g, x) = x\,, \tag{15.3.8}$$

as well as

$$\phi(g^{-1}, x) = \phi(g^{-1}, \phi(g, x)) = \phi(g^{-1}g, x) = \phi(e, x) = x\,, \tag{15.3.9}$$

where $e \in G$ is the unit element. Thus, $gg' \in \mathcal{S}_x^\phi$ and $g^{-1} \in \mathcal{S}_x^\phi$, and so $\mathcal{S}_x^\phi$ is a subgroup. ■

Now we can also use the stabilizer in order to characterize the different types of group actions, which we introduced in definition 15.3.2. We start with the notion of a free action.

Theorem 15.3.8. *A left Lie group action $\phi : G \times M \to M$ is free if and only if $\mathcal{S}_x^\phi = \{e\}$ for all $x \in M$.*

Proof. If ϕ is free, then for all $x \in M$ and $g \in G \setminus \{e\}$ holds

$$\phi(g, x) \neq \phi(e, x) = x\,, \tag{15.3.10}$$

and so $\mathcal{S}_x^\phi = \{e\}$. Conversely, if $\mathcal{S}_x^\phi = \{e\}$ for all $x \in M$, then for all distinct $g, h \in G$ holds

$$\phi(g^{-1}, \phi(h, x)) = \phi(g^{-1}h, x) \neq x = \phi(e, x) = \phi(g^{-1}g, x) = \phi(g^{-1}, \phi(g, x))\,, \tag{15.3.11}$$

and hence also $\phi(g, x) \neq \phi(h, x)$, so that ϕ is free. ■

We then continue with the notion of an effective action, which is a weaker condition.

Theorem 15.3.9. *A left Lie group action $\phi : G \times M \to M$ is effective if and only if*

$$\bigcap_{x \in M} \mathcal{S}_x^\phi = \{e\}\,. \tag{15.3.12}$$

Proof. If ϕ is effective, then for every $g \in G \setminus \{e\}$ there exists $x \in M$ such that

$$\phi(g, x) \neq \phi(e, x) = x\,, \tag{15.3.13}$$

and so $g \notin \mathcal{S}_x^\phi$. On the other hand, $e \in \mathcal{S}_x^\phi$ for all $x \in M$, and so (15.3.12) follows. Conversely, if (15.3.12) holds, then for all distinct $g, h \in G$ there exists $x \in M$ such that $g^{-1}h \notin \mathcal{S}_x^\phi$, so that

$$\phi(g^{-1}, \phi(h, x)) = \phi(g^{-1}h, x) \neq x = \phi(e, x) = \phi(g^{-1}g, x) = \phi(g^{-1}, \phi(g, x))\,, \tag{15.3.14}$$

and hence also $\phi(g, x) \neq \phi(h, x)$, so that ϕ is effective. ■

There is another notion which is in some sense "dual" to that of the stabilizer. For the latter we have fixed a point $x \in M$ and posed the question which group elements $g \in G$ leave this point unchanged. Conversely, we may also fix an element $g \in G$ and pose the question which points $x \in M$ are left unchanged by this group element. This leads to the following definition.

Definition 15.3.7 (Fixed point)**.** Let $\phi : G \times M \to M$ be a left Lie group action and $g \in G$. A *fixed point* of g is an element $x \in M$ satisfying $\phi(g, x) = x$. The set of all fixed points of g is the *fixed point set*

$$\mathcal{F}_g^\phi = \{x \in M \,|\, \phi(g, x) = x\}\,. \tag{15.3.15}$$

Given that the fixed point set is closely related to the stabilizer, it is not surprising that it also allows us to express the two aforementioned conditions on Lie group actions. First, we discuss this for a free Lie group action.

Theorem 15.3.10. *A left Lie group action $\phi : G \times M \to M$ is free if and only if $\mathcal{F}_g^\phi = \emptyset$ for all $g \in G \setminus \{e\}$.*

Proof. If ϕ is free, then for all $x \in M$ and $g \in G \setminus \{e\}$ holds

$$\phi(g, x) \neq \phi(e, x) = x\,, \tag{15.3.16}$$

and so $x \notin \mathcal{F}_g^\phi$; hence, $\mathcal{F}_g^\phi = \emptyset$. Conversely, if $\mathcal{F}_g^\phi = \emptyset$, then for all distinct $g, h \in G$ and $x \in M$ holds $x \notin \mathcal{F}_{g^{-1}h}^\phi$. Thus,

$$\phi(g^{-1}, \phi(h, x)) = \phi(g^{-1}h, x) \neq x = \phi(e, x) = \phi(g^{-1}g, x) = \phi(g^{-1}, \phi(g, x))\,, \tag{15.3.17}$$

and hence also $\phi(g, x) \neq \phi(h, x)$, so that ϕ is free. ■

For an effective Lie group action, a similar statement holds.

Theorem 15.3.11. *A left Lie group action $\phi : G \times M \to M$ is effective if and only if $\mathcal{F}_g^\phi \neq M$ for all $g \in G \setminus \{e\}$.*

Proof. If ϕ is effective, then for every $g \in G \setminus \{e\}$ there exists $x \in M$ such that

$$\phi(g, x) \neq \phi(e, x) = x\,, \tag{15.3.18}$$

and so $x \notin \mathcal{F}_g^\phi$; hence, $\mathcal{F}_g^\phi \neq M$. Conversely, if $\mathcal{F}_g^\phi \neq M$ for all $g \in G \setminus \{e\}$, then for distinct $g, h \in G$ there exists $x \in M \setminus \mathcal{F}_{g^{-1}h}^\phi$, for which holds

$$\phi(g^{-1}, \phi(h, x)) = \phi(g^{-1}h, x) \neq x = \phi(e, x) = \phi(g^{-1}g, x) = \phi(g^{-1}, \phi(g, x))\,, \tag{15.3.19}$$

and hence also $\phi(g, x) \neq \phi(h, x)$, so that ϕ is effective. ■

We illustrate these notions with an example.

Example 15.3.3. Let $\phi : \mathrm{SO}(3) \times \mathbb{R}^3 \to \mathbb{R}^3$ be the left action given by matrix multiplication. Then we have the following relations:

1. For $x \in \mathbb{R}^3$ with $x \neq 0$ the orbit is the sphere with radius $\|x\|$ around the origin. For $x = 0$ the orbit contains only the origin itself.

2. For $x \in \mathbb{R}^3$ with $x \neq 0$ the stabilizer is the subgroup of rotations around the axis $x\mathbb{R}$. For $x = 0$ the stabilizer is $\mathrm{SO}(3)$ itself.

3. For $g \in \mathrm{SO}(3)$ with $g \neq e$ the fixed point set is the axis of rotation of g. For $g = e$ the fixed point set is $\mathbb{R}^3$ itself.

It follows that ϕ is effective, but neither transitive nor free.

We finally discuss the question how we can obtain another Lie group action from a given one, together with a Lie group homomorphism:

Theorem 15.3.12. *Let $\phi : H \times M \to M$ a left action ($\theta : M \times H \to M$ a right action) of a Lie group H on a manifold M and $\varphi : G \to H$ a Lie group homomorphism. Then*

$$\tilde{\phi}(g, x) = \phi(\varphi(g), x)\,, \quad \tilde{\theta}(x, g) = \theta(x, \varphi(g)) \tag{15.3.20}$$

defines a left action (right action) of G on M.

Proof. We show the proof for a left action; the proof for a right action proceeds analogously. First, we check that

$$\tilde{\phi}(e_G, x) = \phi(\varphi(e_G), x) = \phi(e_H, x) = x \tag{15.3.21}$$

for all $x \in M$, where e_G and e_H are the unit elements of G and H, respectively. Further, we have

$$\tilde{\phi}(gg', x) = \phi(\varphi(gg'), x) = \phi(\varphi(g)\varphi(g'), x) = \phi(\varphi(g), \phi(\varphi(g'), x)) = \tilde{\phi}(g, \tilde{\phi}(g', x))\,. \tag{15.3.22}$$

Finally, $\tilde{\phi} = \phi \circ (\varphi, \mathrm{id}_M)$ is a composition of smooth maps, and hence smooth. ■

15.4 Quotient spaces

In theorem 15.3.4 we have shown that the action ϕ of a Lie group G on a manifold M induces an equivalence relation on M, whose equivalence classes are the orbits $\mathcal{O}_x^\phi$ for $x \in M$. It is thus possible to define a quotient space, i.e., the space of all orbits of ϕ. We now come to the question under which circumstances this quotient space carries the structure of a manifold.

Theorem 15.4.1. *Let $\phi : G \times M \to M$ a left Lie group action which is free and proper. Then the space of orbits of ϕ carries the structure of a smooth manifold, such that the projection $x \mapsto \mathcal{O}_x^\phi$ is a smooth, surjective submersion.*

Proof. ►...◄ ■

Definition 15.4.1 (Coset space)**.** Let G be a Lie group and $H \subset G$ a closed subgroup. The *left coset* of $g \in G$ is the equivalence class

$$gH = \{gh, h \in H\} \subset G\,, \tag{15.4.1}$$

while its *right coset* is

$$Hg = \{hg, h \in H\} \subset G\,. \tag{15.4.2}$$

The *left coset space* G/H is the space of all left cosets,

$$G/H = \{gH, g \in G\}\,, \tag{15.4.3}$$

while the *right coset space* $H\backslash G$ is

$$H\backslash G = \{Hg, g \in G\}\,. \tag{15.4.4}$$

In other words, the coset gH is the orbit of g under the right action $G \times H \to G, (g, h) \mapsto gh$ of H on G, while G/H is the corresponding orbit space. The same holds for right cosets.

Theorem 15.4.2. *Let G be a Lie group and $H \subset G$ a closed subgroup. Then the coset spaces G/H and $H\backslash G$ carry the structure of smooth manifolds, such that the projections $g \mapsto gH$ and $g \mapsto Hg$ are smooth, surjective submersions.*

Proof. ►...◄ ■

Theorem 15.4.3 (Orbit-stabilizer theorem)**.** *Let $\phi : G \times M \to M$ a left action of a Lie group G on a manifold M and $x \in M$. Then $\mathcal{O}_x^\phi \cong G/\mathcal{S}_x^\phi$.*

Proof. ►...◄ ■

15.5 Equivariant maps

Often one needs to relate different manifolds which carry actions of the same Lie group. For this purpose it is useful to introduce a particular class of maps between such manifolds. These maps will be defined as follows.

Definition 15.5.1 (Equivariant map)**.** Let G be a Lie group and M, N manifolds which carry Lie group actions of G. A map $\varphi : M \to N$ is called *G-equivariant* if for all $g \in G$ and $x \in M$

- $\varphi(\rho_M(g, x)) = \rho_N(g, \varphi(x))$ if both $\rho_M : G \times M \to M$ and $\rho_N : G \times N \to N$ are left actions,
- $\varphi(\theta_M(x, g)) = \theta_N(\varphi(x), g)$ if both $\theta_M : M \times G \to M$ and $\theta_N : N \times G \to N$ are right actions,
- $\varphi(\rho_M(g, x)) = \theta_N(\varphi(x), g^{-1})$ if $\rho_M : G \times M \to M$ is a left action and $\theta_N : N \times G \to N$ is a right action,
- $\varphi(\theta_M(x, g)) = \rho_N(g^{-1}, \varphi(x))$ if $\theta_M : M \times G \to M$ is a right action and $\rho_N : G \times N \to N$ is a left action.

We denote the space of space of G-equivariant maps by $C_G^\infty(M, N)$.

This can be illustrated by a simple example.

Example 15.5.1. Consider the Lie group $G = \mathrm{SO}(3)$. Let $M = \mathbb{R}^3 \times \mathbb{R}^3$ with left action $\rho_M(g, (x, y)) = (gx, gy)$ and $N = \mathbb{R}^3$ with left action $\rho_N(g, x) = gx$, where gx denotes the multiplication of a matrix and a vector. Then the *cross product* $\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ is an equivariant map.

We can also visualize this notion in terms of commutative diagrams. Let us first assume the case of two left actions ρ_M, ρ_N. Further, denote by (id_G, φ) the map

$$(\mathrm{id}_G, \varphi) \quad : \quad \begin{array}{ccc} G \times M & \to & G \times N \\ (g, x) & \mapsto & (g, \varphi(x)) \end{array} . \tag{15.5.1}$$

Then φ is equivariant if and only if the diagram

$$\begin{array}{ccc} G \times M & \xrightarrow{(\mathrm{id}_G, \varphi)} & G \times N \\ {\scriptstyle \rho_M}\downarrow & & \downarrow{\scriptstyle \rho_N} \\ M & \xrightarrow[\varphi]{} & N \end{array} \tag{15.5.2}$$

commutes. Analogously, for two right actions θ_M, θ_N one demands that the diagram

$$\begin{array}{ccc} M \times G & \xrightarrow{(\varphi, \mathrm{id}_G)} & N \times G \\ {\scriptstyle \theta_M}\downarrow & & \downarrow{\scriptstyle \theta_N} \\ M & \xrightarrow[\varphi]{} & N \end{array} \tag{15.5.3}$$

commutes. For the mixed cases, one has to replace the identity id_G by the inversion map $\bullet^{-1}$ and swap the order of the pairs (x, g) to (g, x) and vice versa.

One may ask why it is necessary to take the inverse group element in the mixed case, where one has a left Lie group action on one manifold and a right action on the other. This can be seen from the properties of Lie group actions, from which follows

$$\begin{aligned} \theta_N(\varphi(x), (gh)^{-1}) &= \varphi(\rho_M(gh, x)) \\ &= \varphi(\rho_M(g, \rho_M(h, x))) \\ &= \theta_N(\varphi(\rho_M(h, x)), g^{-1}) \\ &= \theta_N(\theta_N(\varphi(x), h^{-1}), g^{-1}) \\ &= \theta_N(\varphi(x), h^{-1}g^{-1}) \\ &= \theta_N(\varphi(x), (gh)^{-1}) \,. \end{aligned} \tag{15.5.4}$$

We see that by taking the inverse, the first and the last line are consistent without imposing any restrictions on the Lie group actions or the equivariant map. If we had not taken the inverse, we would have obtained

$$\begin{aligned} \theta_N(\varphi(x), gh) &= \varphi(\rho_M(gh, x)) \\ &= \varphi(\rho_M(g, \rho_M(h, x))) \\ &= \theta_N(\varphi(\rho_M(h, x)), g) \\ &= \theta_N(\theta_N(\varphi(x), h), g) \\ &= \theta_N(\varphi(x), hg) \end{aligned} \tag{15.5.5}$$

instead, which can be satisfied only if $\theta_N(y, gh) = \theta_N(y, hg)$ for all y in the image of φ, and which does not hold for general Lie group actions.

15.6 Lie algebras

So far we have introduced the basic structure of Lie groups and their actions on manifolds. We now consider particular classes of vector fields and differential forms on Lie groups, which play an important role in physics. We start with the following definition.

Definition 15.6.1 (Invariant vector field)**.** Let G be a Lie group. A vector field X on G is called *left invariant* if its pullback along the diffeomorphism L_g for all $g \in G$ satisfies $L_g^*(X) = X$. Similarly, it is called *right invariant* if $R_g^*(X) = X$ for all $g \in G$.

Note that in general the left- and right invariant field vector fields on a Lie group are not the same, unless the group is abelian. We illustrate this with an example.

Example 15.6.1 (Invariant vector fields on the one-dimensional affine group). We consider the group $G = \mathrm{IGL}(1, \mathbb{R}) = \mathbb{R} \rtimes \mathrm{GL}(1, \mathbb{R})$ of affine transformations of the real line. We can make use of the block form (15.1.4) to write the group elements in the form

$$g = \begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix} \tag{15.6.1}$$

with $x, y \in \mathbb{R}$ and $x \neq 0$. This allows us to use $(x, y) \in \mathbb{R} \setminus \{0\} \times \mathbb{R}$ as coordinates on a global chart of G. In these coordinates, the left and right translations are given by

$$\begin{array}{cccc} L_{(\tilde{x},\tilde{y})} & : & G & \to & G \\ & & (x, y) & \mapsto & (x\tilde{x}, y\tilde{x} + \tilde{y}) \end{array} \tag{15.6.2}$$

and

$$\begin{array}{cccc} R_{(\tilde{x},\tilde{y})} & : & G & \to & G \\ & & (x, y) & \mapsto & (x\tilde{x}, x\tilde{y} + y) \end{array} . \tag{15.6.3}$$

Introducing coordinates (x, y, u, v) on TG such that a tangent vector is written as $u\partial_x + v\partial_y \in T_{(x,y)}G$, the differentials of these maps are given by

$$\begin{array}{cccc} L_{(\tilde{x},\tilde{y})*} & : & TG & \to & TG \\ & & (x, y, u, v) & \mapsto & (x\tilde{x}, y\tilde{x} + \tilde{y}, \tilde{x}u, \tilde{x}v) \end{array} \tag{15.6.4}$$

and

$$\begin{array}{cccc} R_{(\tilde{x},\tilde{y})*} & : & TG & \to & TG \\ & & (x, y, u, v) & \mapsto & (x\tilde{x}, x\tilde{y} + y, \tilde{x}u, \tilde{y}u + v) \end{array} . \tag{15.6.5}$$

We then construct the invariant vector fields as follows. Let $\tilde{u}\partial_x + \tilde{v}\partial_y \in T_{(1,0)}G$ be a tangent vector at the unit element $(1, 0)$, which we write in coordinates as $(1, 0, \tilde{u}, \tilde{v})$. We then define vector fields $X_{(\tilde{u},\tilde{v})}, Y_{(\tilde{u},\tilde{v})} \in \mathrm{Vect}(G)$ such that

$$\begin{array}{cccc} X_{(\tilde{u},\tilde{v})} & : & G & \to & TG \\ & & (x, y) & \mapsto & L_{(x,y)*}(\tilde{u}\partial_x + \tilde{v}\partial_y) = (x, y, x\tilde{u}, x\tilde{v}) \end{array} \tag{15.6.6}$$

and

$$\begin{array}{cccc} Y_{(\tilde{u},\tilde{v})} & : & G & \to & TG \\ & & (x, y) & \mapsto & R_{(x,y)*}(\tilde{u}\partial_x + \tilde{v}\partial_y) = (x, y, x\tilde{u}, y\tilde{u} + \tilde{v}) \end{array} . \tag{15.6.7}$$

To show that $X_{(\tilde{u},\tilde{v})}$ is left invariant, let us define

$$(x', y') = L_{(\tilde{x},\tilde{y})}(x, y) = (x\tilde{x}, y\tilde{x} + \tilde{y}) \quad \Rightarrow \quad (x, y) = \left(\frac{x'}{\tilde{x}}, \frac{y' - \tilde{y}}{\tilde{x}}\right) . \tag{15.6.8}$$

The pullback $L^*_{(\tilde{x},\tilde{y})}X_{(\tilde{u},\tilde{v})}$ of $X_{(\tilde{u},\tilde{v})}$ along $L_{(\tilde{x},\tilde{y})}$, evaluated at (x, y), is given by

$$\begin{aligned} \left(L^*_{(\tilde{x},\tilde{y})}X_{(\tilde{u},\tilde{v})}\right)(x, y) &= (\partial_x, \partial_y) \cdot \begin{pmatrix} \frac{\partial x}{\partial x'} & \frac{\partial x}{\partial y'} \\ \frac{\partial y}{\partial x'} & \frac{\partial y}{\partial y'} \end{pmatrix} \cdot \begin{pmatrix} X^x_{(\tilde{u},\tilde{v})}(x', y') \\ X^y_{(\tilde{u},\tilde{v})}(x', y') \end{pmatrix} \\ &= (\partial_x, \partial_y) \cdot \begin{pmatrix} \frac{1}{\tilde{x}} & 0 \\ 0 & \frac{1}{\tilde{x}} \end{pmatrix} \cdot \begin{pmatrix} x'\tilde{u} \\ x'\tilde{v} \end{pmatrix} \\ &= x\tilde{u}\partial_x + x\tilde{v}\partial_y \\ &= X_{(\tilde{u},\tilde{v})}(x, y) , \end{aligned} \tag{15.6.9}$$

so that $X_{(\tilde{u},\tilde{v})}$ is indeed left invariant. Similarly, if we define

$$(x', y') = R_{(\tilde{x},\tilde{y})}(x, y) = (x\tilde{x}, x\tilde{y} + y) \quad \Rightarrow \quad (x, y) = \left(\frac{x'}{\tilde{x}}, y' - x'\frac{\tilde{y}}{\tilde{x}}\right) , \tag{15.6.10}$$

we find that

$$\begin{aligned}\left(R^*_{(\tilde{x},\tilde{y})}Y_{(\tilde{u},\tilde{v})}\right)(x,y) &= (\partial_x,\partial_y)\cdot\begin{pmatrix}\frac{\partial x}{\partial x'} & \frac{\partial x}{\partial y'}\\ \frac{\partial y}{\partial x'} & \frac{\partial y}{\partial y'}\end{pmatrix}\cdot\begin{pmatrix}Y^x_{(\tilde{u},\tilde{v})}(x',y')\\ Y^y_{(\tilde{u},\tilde{v})}(x',y')\end{pmatrix}\\ &= (\partial_x,\partial_y)\cdot\begin{pmatrix}\frac{1}{\tilde{x}} & 0\\ -\frac{\tilde{y}}{\tilde{x}} & 1\end{pmatrix}\cdot\begin{pmatrix}x'\tilde{u}\\ y'\tilde{u}+\tilde{v}\end{pmatrix}\\ &= x\tilde{u}\partial_x + (y\tilde{u}+\tilde{v})\partial_y\\ &= Y_{(\tilde{u},\tilde{v})}(x,y)\,,\end{aligned} \tag{15.6.11}$$

and hence $Y_{(\tilde{u},\tilde{v})}$ is indeed right invariant.

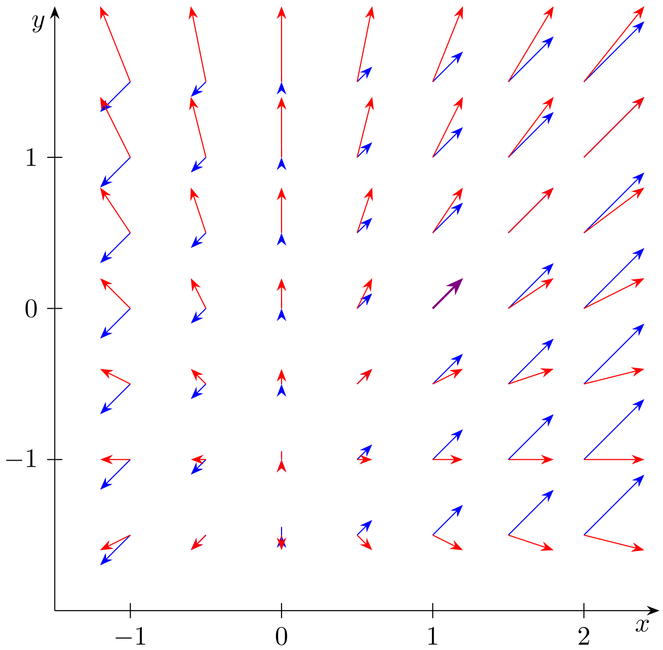


Figure 15.1: Example of a left and right invariant vector field on the affine group $G = \mathrm{IGL}(1,\mathbb{R})$; see example 15.6.1 for an explanation of the coordinates. Even though $X_{(1,1/2)}$ (blue) and $Y_{(1,1/2)}$ (red) agree at the unit element $(1,0)$, they differ almost everywhere else.

From the fact that diffeomorphisms preserve the Lie bracket follows the following property.

Theorem 15.6.1. *Let X,Y be left (right) invariant vector fields on a Lie group G. Then also their Lie bracket $[X,Y]$ is left (right) invariant.*

Proof. ►. . . ◄ ■

In the following we will use the standard convention and work with left invariant vector fields in order to be consistent with the literature. The statement above then tells us that the left invariant vector fields together with the Lie bracket form a Lie algebra, which plays a fundamental role.

Definition 15.6.2 (Lie algebra)**.** Let G be a Lie group. Its *Lie algebra* is the Lie algebra $\mathfrak{g} = \mathrm{Lie}(G)$ defined by the left invariant vector fields together with the Lie bracket of vector fields.

The question arises whether this Lie algebra is finite-dimensional, and what is its dimension. The following theorem answers both of these questions.

Theorem 15.6.2. *The Lie algebra $\mathfrak{g}$ of a Lie group G is canonically isomorphic as a vector space to the tangent space T_eG at the unit element $e \in G$.*

Proof. This is easy to see. Given a left-invariant vector field X, one can simply evaluate it at e to obtain $X(e) \in T_eG$. Conversely, given $v \in T_eG$, one can uniquely construct a left invariant vector field X as $X(g) = L_{g*}(v) \in T_gG$. ■

It thus follows immediately that the dimension of the Lie algebra $\mathfrak{g}$ is the same as the dimension of the Lie group G, and we can simply identify $\mathfrak{g}$ and T_eG. This allows us to construct the Lie algebras of the matrix groups shown in the first section.

Example 15.6.2. The Lie algebras of the matrix groups for $n \in \mathbb{N}$ are as follows, where the Lie bracket $[A, B]$ is given by the matrix commutator $AB - BA$:

- The (real) *general linear algebra* $\mathfrak{gl}(n, \mathbb{R})$ (or simply $\mathfrak{gl}(n)$) of real $n \times n$ matrices is a Lie algebra of dimension n^2.
- The complex *general linear algebra* $\mathfrak{gl}(n, \mathbb{C})$ of complex $n \times n$ matrices is a complex Lie algebra of complex dimension n^2, or a real Lie algebra of real dimension $2n^2$.
- The (real) *special linear algebra* $\mathfrak{sl}(n, \mathbb{R})$ (or simply $\mathfrak{sl}(n)$) of real, trace-free $n \times n$ matrices is a Lie algebra of dimension $n^2 - 1$.
- The complex *special linear algebra* $\mathfrak{sl}(n, \mathbb{C})$ of complex, trace-free $n \times n$ matrices is a complex Lie algebra of complex dimension $n^2 - 1$, or a real Lie algebra of real dimension $2n^2 - 2$.
- The *orthogonal algebra* $\mathfrak{o}(n)$, which is the same as the *special orthogonal algebra* $\mathfrak{so}(n)$, of real, antisymmetric $n \times n$ matrices, $A = -A^t$, is a Lie algebra of dimension $n(n - 1)/2$.
- The *indefinite orthogonal algebra* $\mathfrak{o}(p, q)$ with $n = p + q$, which is the same as the *indefinite special orthogonal algebra* $\mathfrak{so}(p, q)$, of real $n \times n$ matrices satisfying $A = -\eta A^t \eta$, is a Lie algebra of dimension $n(n - 1)/2$.
- The (real) *symplectic algebra* $\mathfrak{sp}(2n, \mathbb{R})$ (or simply $\mathfrak{sp}(2n)$) of real $2n \times 2n$ matrices satisfying $\Omega A + A^t \Omega = 0$ is a Lie algebra of dimension $n(2n + 1)$.
- The complex *symplectic algebra* $\mathfrak{sp}(2n, \mathbb{C})$ of complex $2n \times 2n$ matrices satisfying $\Omega A + A^t \Omega = 0$ is a complex Lie algebra of dimension $n(2n + 1)$, or a real Lie algebra of dimension $2n(2n + 1)$.
- The *unitary algebra* $\mathfrak{u}(n)$ of complex, anti-hermitian $n \times n$ matrices, $A = -A^\dagger$, is a Lie algebra of dimension n^2.
- The *special unitary algebra* $\mathfrak{su}(n)$ of complex, anti-hermitian $n \times n$ matrices, $A = -A^\dagger$, with trace 0 is a Lie algebra of dimension $n^2 - 1$.

15.7 Exponential map

To further explore the relationship between Lie groups and their Lie algebras, we define the following.

Definition 15.7.1 (One-parameter subgroup). A *one-parameter subgroup* of a Lie group G is a Lie group homomorphism $\varphi : (\mathbb{R}, +) \to G$.

In other words, a one-parameter subgroup is a curve φ on G such that $\varphi(s+t) = \varphi(s)\varphi(t)$ for all $s, t \in \mathbb{R}$. In particular it follows that $\varphi(0) = e$ is the unit element of G. A one-parameter subgroup thus defines an element $\dot{\varphi}(0) \in T_eG \cong \mathfrak{g}$. The following theorem states that also the converse is true.

Theorem 15.7.1. *Let G be a Lie group and $X \in \mathfrak{g}$ a left invariant vector field. Then there exists a unique one-parameter subgroup φ_X such that $\dot{\varphi}_X(t) = X(\varphi_X(t))$ for all $t \in \mathbb{R}$.*

Proof. ▶...◀ ■

The proof is a bit lengthy, but simple, so we will omit it here. This theorem allows us to finally define another important concept.

Definition 15.7.2 (Exponential map). Let G be a Lie group and $\mathfrak{g}$ its Lie algebra. The *exponential map* is the map

$$\begin{array}{rcccc} \exp & : & \mathfrak{g} & \to & G \\ & & X & \mapsto & \varphi_X(1) \end{array}, \tag{15.7.1}$$

where φ_X is the unique one-parameter subgroup such that $\dot{\varphi}_X(t) = X(\varphi_X(t))$ for all $t \in \mathbb{R}$.

We will continue with a few properties of the exponential map.

Theorem 15.7.2. *The exponential map satisfies:*

- *It maps the zero element $0 \in \mathfrak{g}$ to the unit e of the Lie group:* $\exp(0) = e$.
- *For all $X \in \mathfrak{g}$ holds* $\exp(-X) = \exp(X)^{-1}$.
- *For all $X \in \mathfrak{g}$ and $s, t \in \mathbb{R}$ holds* $\exp((s+t)X) = \exp(sX)\exp(tX)$.

Proof. ▶...◀ ■

One may be tempted to conclude that the usual law for the exponential function of complex numbers holds, so that one may simply replace the exponential of the sum of two Lie algebra elements by the product of their exponentials. However, this is *not* the case for general Lie algebras - one has $\exp(X+Y) \neq \exp(X)\exp(Y)$ in general! Instead, one has to apply the Baker-Campbell-Hausdorff formula.

15.8 Lie algebra homomorphisms

As with Lie groups, given two Lie algebras $\mathfrak{g}_1, \mathfrak{g}_2$, the question arises which maps between them preserve not only the underlying manifold structure, in this case inherited from the fiber of the tangent bundle over the unit element, but also the algebraic structure. For such kind of maps we introduce the following terminology.

Definition 15.8.1 (Lie algebra homomorphism / isomorphism)**.** Let $\mathfrak{g}_1$ and $\mathfrak{g}_2$ be Lie algebras. A *Lie algebra homomorphism* from $\mathfrak{g}_1$ to $\mathfrak{g}_2$ is a smooth, linear map $\varphi : \mathfrak{g}_1 \to \mathfrak{g}_2$ such that $\varphi([X, Y]) = [\varphi(X), \varphi(Y)]$ for all $X, Y \in \mathfrak{g}_1$. If it is also a diffeomorphism, it is called a *Lie algebra isomorphism*.

As with Lie group isomorphisms, also Lie algebra isomorphisms have various applications. Again we consider which class of isomorphisms arises when both algebras are identical, $\mathfrak{g}_1 = \mathfrak{g}_2 = \mathfrak{g}$:

Definition 15.8.2 (Lie algebra automorphism)**.** Let $\mathfrak{g}$ be a Lie algebra. An *automorphism* of $\mathfrak{g}$ is a Lie algebra isomorphism $\varphi : \mathfrak{g} \to \mathfrak{g}$. The group of all Lie algebra automorphisms of $\mathfrak{g}$ is denoted $\mathrm{Aut}(\mathfrak{g})$.

Since we derived Lie algebras from Lie groups, one may naturally expect that also Lie group homomorphisms induce Lie algebra homomorphisms between their corresponding Lie algebras. We now show that this is indeed the case.

Theorem 15.8.1. *Let G_1, G_2 be Lie groups and $\phi : G_1 \to G_2$ a Lie group homomorphism. Then $\varphi = \phi_*|_{e_1} : T_{e_1}G_1 \to T_{e_2}G_2$ is a Lie algebra homomorphism, and φ is a Lie algebra isomorphism if and only if ϕ is a Lie group isomorphism.*

Proof. Note first that a Lie group homomorphism $\phi : G_1 \to G_2$ in particular satisfies $\phi(e_1) = e_2$. Thus, the differential ϕ_* maps $T_{e_1}G_1 \cong \mathfrak{g}_1$ to $T_{e_2}G_2 \cong \mathfrak{g}_2$, and we made use of this fact when we identified the Lie algebras with the tangent spaces of the unit elements in the statement of the theorem. ▶. . . ◀ ■

15.9 Adjoint representation

A particularly important class of Lie algebra automorphisms, which we will use later, can be constructed from the inner automorphisms α_g of a Lie group by application of theorem 15.8.1. We can thus define the following notion:

Definition 15.9.1 (Adjoint representation)**.** For a Lie group G, the *adjoint representation* is the map $\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g})$ defined by $\mathrm{Ad}_g = \alpha_{g*}$.

One easily checks that for all $g \in G$ the map $\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$ is indeed an automorphism, i.e., it satisfies $\mathrm{Ad}_g[X, Y] = [\mathrm{Ad}_g(X), \mathrm{Ad}_g(Y)]$ for all $X, Y \in \mathfrak{g}$.

15.10 Lie algebra valued differential forms

To further study Lie groups and their properties, it is often useful to work with differential k-forms, which do not yield a real number when they are evaluated at k vectors at a point, but to an element of a Lie algebra. In this section we introduce these objects and discuss a number of operations acting on them. We start with a formal definition.

Definition 15.10.1 (Lie algebra valued differential form)**.** Let M be a manifold and $\mathfrak{g}$ be a Lie algebra. The space $\Omega^k(M, \mathfrak{g})$ of *Lie algebra valued differential k-forms* is the space $\Gamma(E \otimes \Lambda^k T^*M)$ of sections of the tensor product bundle $E \otimes \Lambda^k T^*M$, where $E = M \times \mathfrak{g}$ is the trivial bundle with base M and fiber $\mathfrak{g}$.

Given a basis $(\mathbb{e}_\mu)$ of $\mathfrak{g}$ and coordinates (x^a) on M, one can thus write a $\mathfrak{g}$-valued k-form $\omega \in \Omega^k(M, \mathfrak{g})$ as

$$\omega = \omega^\mu \mathbb{e}_\mu = \frac{1}{k!} \omega^\mu{}_{a_1 \cdots a_k} \mathbb{e}_\mu \otimes \mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k} \,, \tag{15.10.1}$$

thus generalizing the basis expansion (9.1.2) of ordinary differential forms. A number of operations which are defined on differential forms, in particular those which are linear in their single argument, can be extended to act on Lie algebra valued forms by acting component-wise on each component ω^μ. This includes:

1. the exterior derivative $\mathrm{d}\omega$,
2. the interior product $\iota_X \omega$ with a vector field $X \in \mathrm{Vect}(M)$,
3. the pullback $\varphi^*(\omega)$ along a map $\varphi : N \to M$,
4. the Lie derivative $\mathcal{L}_X \omega$ along a vector field $X \in \mathrm{Vect}(M)$.

More care must be taken for the exterior product $\omega \wedge \sigma$, which is linear in each of its arguments by definition. The only possible generalization to a component-wise operation arising from this linearity is to allow *one* of the two factors ω, σ to take values in a Lie algebra, while the other one remains an ordinary differential form. More often, however, one encounters a product involving two Lie algebra valued differential forms. It can formally be defined as follows.

Definition 15.10.2 (Lie algebra valued exterior product)**.** Let M be a manifold and $\mathfrak{g}$ be a Lie algebra. The *exterior product of Lie algebra valued forms* is the unique bilinear function $[\bullet \wedge \bullet] : \Omega^k(M, \mathfrak{g}) \times \Omega^l(M, \mathfrak{g}) \to \Omega^{k+l}(M, \mathfrak{g})$, such that for all $\alpha \in \Omega^k(M), \beta \in \Omega^l(M)$ and $a, b \in \mathfrak{g}$ holds

$$[(a \otimes \alpha) \wedge (b \otimes \beta)] = [a, b] \otimes (\alpha \wedge \beta) \,. \tag{15.10.2}$$

Here we make use of the fact that any Lie algebra is equipped with the bilinear Lie bracket $[\bullet, \bullet] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, which can be used to combine two elements into one. The notation $[\bullet \wedge \bullet]$ we introduced here shows that both an exterior product and this Lie bracket are involved. In other literature one also finds simply the bracket notation $[\bullet, \bullet]$ here. Demanding linearity in the definition above ensures that the operation is defined on the whole space of Lie algebra valued differential forms by linear extension from the basis elements. Given

$$\omega = \frac{1}{k!} \omega^\mu{}_{a_1 \cdots a_k} \mathbb{e}_\mu \otimes \mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k} \in \Omega^k(M, \mathfrak{g}) \,, \tag{15.10.3a}$$

$$\sigma = \frac{1}{l!} \sigma^\mu{}_{a_1 \cdots a_l} \mathbb{e}_\mu \otimes \mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_l} \in \Omega^l(M, \mathfrak{g}) \,, \tag{15.10.3b}$$

one has

$$[\omega \wedge \sigma] = \frac{1}{k! l!} \omega^\mu{}_{a_1 \cdots a_k} \sigma^\nu{}_{b_1 \cdots b_l} [\mathbb{e}_\mu, \mathbb{e}_\nu] \otimes \mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k} \wedge \mathrm{d}x^{b_1} \wedge \ldots \wedge \mathrm{d}x^{b_l} \in \Omega^{k+l}(M, \mathfrak{g}) \,. \tag{15.10.4}$$

From the properties of the exterior product and the Lie bracket follow a number of useful relations also for the Lie algebra valued exterior product, which we summarize as follows:

Theorem 15.10.1. *The Lie algebra valued exterior product satisfies the graded antisymmetry*

$$[\omega \wedge \sigma] = -(-1)^{kl}[\sigma \wedge \omega] \tag{15.10.5}$$

and graded Jacobi identity

$$[\omega \wedge [\sigma \wedge \tau]] = [[\omega \wedge \sigma] \wedge \tau] + (-1)^{kl}[\sigma \wedge [\omega \wedge \tau]] \tag{15.10.6}$$

for $\omega \in \Omega^k(M, \mathfrak{g})$, $\sigma \in \Omega^l(M, \mathfrak{g})$ and $\tau \in \Omega^m(M, \mathfrak{g})$.

Proof. Since the Lie algebra valued exterior product is linear in each argument, it suffices to show these relations for tensor products of the form $\omega = a \otimes \alpha$, $\sigma = b \otimes \beta$ and $\tau = c \otimes \gamma$, where $\alpha \in \Omega^k(M)$, $\beta \in \Omega^l(M)$, $\gamma \in \Omega^m(M)$ and $a, b, c \in \mathfrak{g}$, and then to conclude by linearity. For the first relation, the symmetry properties of the ordinary exterior product and the Lie bracket then yield

$$[(a\otimes\alpha)\wedge(b\otimes\beta)] = [a,b]\otimes(\alpha\wedge\beta) = -[b,a]\otimes(-1)^{kl}(\beta\wedge\alpha) = -(-1)^{kl}[(a\otimes\alpha)\wedge(b\otimes\beta)]\,. \tag{15.10.7}$$

Similarly, one calculates the Jacobi identity as

$$\begin{aligned}
[(a \otimes \alpha) \wedge [(b \otimes \beta) \wedge (c \otimes \gamma)]] &= [a, [b, c]] \otimes (\alpha \wedge \beta \wedge \gamma) \\
&= -([b, [c, a]] + [c, [a, b]]) \otimes (\alpha \wedge \beta \wedge \gamma) \\
&= ([[a, b], c] + [b, [a, c]]) \otimes (\alpha \wedge \beta \wedge \gamma) \\
&= [[a, b], c] \otimes (\alpha \wedge \beta \wedge \gamma) + [b, [a, c]] \otimes (\alpha \wedge \beta \wedge \gamma) \\
&= [[a, b], c] \otimes (\alpha \wedge \beta \wedge \gamma) + [b, [a, c]] \otimes (-1)^{kl}(\beta \wedge \alpha \wedge \gamma) \\
&= [[(a \otimes \alpha) \wedge (b \otimes \beta)] \wedge (c \otimes \gamma)] \\
&\quad + (-1)^{kl}[(b \otimes \beta) \wedge [(a \otimes \alpha) \wedge (c \otimes \gamma)]]\,. \quad \blacksquare
\end{aligned} \tag{15.10.8}$$

Another comment regarding the relation to the exterior derivative and the interior product is in order. Since these act component-wise on the differential form part only while leaving the Lie algebra part inert, they satisfy the usual relations

$$\mathrm{d}[\omega \wedge \sigma] = [\mathrm{d}\omega \wedge \sigma] + (-1)^k[\omega \wedge \mathrm{d}\sigma] \tag{15.10.9}$$

and

$$\iota_X[\omega \wedge \sigma] = [\iota_X\omega \wedge \sigma] + (-1)^k[\omega \wedge \iota_X\sigma] \tag{15.10.10}$$

for $\omega \in \Omega^k(M, \mathfrak{g})$, $\sigma \in \Omega^l(M, \mathfrak{g})$ and $X \in \mathrm{Vect}(M)$. Also note that for Lie algebra valued one-forms $\omega, \sigma \in \Omega^1(M, \mathfrak{g})$ and vector fields $X, Y \in \mathrm{Vect}(M)$ we have

$$\iota_Y\iota_X[\omega \wedge \sigma] = [\iota_X\omega, \iota_Y\sigma] - [\iota_Y\omega, \iota_X\sigma]\,. \tag{15.10.11}$$

In particular, for $\omega = \sigma$ thus follows

$$\iota_Y\iota_X[\omega \wedge \omega] = [\iota_X\omega, \iota_Y\omega] - [\iota_Y\omega, \iota_X\omega] = 2[\iota_X\omega, \iota_Y\omega]\,. \tag{15.10.12}$$

To illustrate the Lie algebra valued exterior product, we finally consider the case of matrix Lie algebras, such as those given in example 15.6.2.

> *Example* 15.10.1. For a matrix Lie algebra $\mathfrak{g} \subset M_{n,n}$, one can represent elements $a, b \in \mathfrak{g}$ by their matrix components $a^\mu{}_\nu, b^\mu{}_\nu$. The Lie bracket is given by the commutator, and thus has components
>
> $$[a, b]^\mu{}_\nu = a^\mu{}_\rho b^\rho{}_\nu - b^\mu{}_\rho a^\rho{}_\nu\,. \tag{15.10.13}$$

When considering Lie algebra valued differential forms $\omega \in \Omega^k(M, \mathfrak{g})$ and $\sigma \in \Omega^l(M, \mathfrak{g})$, each matrix component represents an ordinary differential form, and instead of the commutative product of real numbers one has the graded exterior product, hence

$$[\omega \wedge \sigma]^\mu{}_\nu = \omega^\mu{}_\rho \wedge \sigma^\rho{}_\nu - \omega^\rho{}_\nu \wedge \sigma^\mu{}_\rho = \omega^\mu{}_\rho \wedge \sigma^\rho{}_\nu - (-1)^{kl} \sigma^\mu{}_\rho \wedge \omega^\rho{}_\nu \,. \tag{15.10.14}$$

In the literature, one sometimes finds the indices omitted, and simply reads $[\omega \wedge \sigma] = \omega \wedge \sigma - (-1)^{kl} \sigma \wedge \omega$, where matrix multiplication is implied; we do not use such notation here to avoid confusion. In particular, for $\omega = \sigma$ one has

$$[\omega \wedge \omega]^\mu{}_\nu = \left(1 - (-1)^{k^2}\right) \omega^\mu{}_\rho \wedge \omega^\rho{}_\nu \,, \tag{15.10.15}$$

where the factor in brackets equals 0 if k is even and 2 if k is odd.

15.11 Maurer-Cartan form

We now come to a particular example for a Lie algebra valued differential form, as discussed in the previous section. It is canonically defined on any Lie group as follows.

Definition 15.11.1 (Maurer-Cartan form)**.** Let G be a Lie group with Lie algebra $\mathfrak{g}$. Its *Maurer-Cartan form* is the $\mathfrak{g}$-valued one-form $\mu \in \Omega^1(G, \mathfrak{g})$ such that for all $g \in G$ and $v \in T_g G$ the element $X = \mu(g)(v) \in \mathfrak{g} \subset \mathrm{Vect}(G)$ is the unique left invariant vector field satisfying $X(g) = v$.

A few explanations are in order. Since the Maurer-Cartan form μ is a $\mathfrak{g}$-valued one-form, it can be seen as a map $\mu : TG \to \mathfrak{g}$ which is linear on every fiber, or in other words a linear map $\mu(g) : T_g G \to \mathfrak{g}$ for each $g \in G$. In our definition we have identified $\mathfrak{g}$ with the left invariant vector fields on G, and so this linear map is the unique inverse of the evaluation map

$$\begin{array}{cccc} \bullet(g) & : & \mathfrak{g} & \to & T_g G \\ & & X & \mapsto & X(g) \end{array} \,. \tag{15.11.1}$$

Alternatively, one may identify $\mathfrak{g}$ with the tangent space $T_e G$; the Maurer-Cartan form is then defined as $\mu(g)(v) = (L_{g^{-1}*} v) \in T_e G$. Again we see that this mirrors the definition of the left-invariant vector fields X, since these satisfy

$$X(g) = L_{g*}(X(e)) \,. \tag{15.11.2}$$

It follows from these properties that the Maurer-Cartan form is itself left-invariant; one can prove this and another equivariance property as follows.

Theorem 15.11.1. *The Maurer-Cartan form μ on any Lie group G satisfies the equivalence properties*

$$L_g^*(\mu) = \mu \,, \quad R_g^*(\mu) = \mathrm{Ad}_{g^{-1}} \circ \mu \tag{15.11.3}$$

for all $g \in G$.

Proof. Here we will use the identification of the Lie algebra $\mathfrak{g}$ with the tangent space $T_e G$. Let $h \in G$ and $v \in T_h G$. Then we have $L_g(h) = gh$ and thus

$$L_g^*(\mu)(v) = \mu(L_{g*}(v)) = L_{(gh)^{-1}*}(L_{g*}(v)) = L_{h^{-1}*}(v) = \mu(v) \,, \tag{15.11.4}$$

where we used the property that the left translation is a left action and thus satisfies

$$L_{(gh)^{-1}} \circ L_g = L_{h^{-1}g^{-1}} \circ L_g = L_{h^{-1}} . \tag{15.11.5}$$

Similarly, for the right translations we have $R_g(h) = hg$, so that we find

$$\begin{aligned} R_g^*(\mu)(v) &= \mu(R_{g*}(v)) \\ &= L_{(hg)^{-1}*}(R_{g*}(v)) \\ &= L_{g^{-1}*}(R_{g*}(L_{h^{-1}*}(v))) \\ &= \alpha_{g^{-1}*}(L_{h^{-1}*}(v)) \\ &= \mathrm{Ad}_{g^{-1}}(\mu(v)) , \end{aligned} \tag{15.11.6}$$

where we now used the fact that left and right translations commute to write

$$L_{(hg)^{-1}} \circ R_g = L_{g^{-1}h^{-1}} \circ R_g = L_{g^{-1}} \circ R_g \circ L_{h^{-1}} , \tag{15.11.7}$$

as well as the definition 15.2.3 of the inner automorphism

$$\alpha_{g^{-1}} = R_g = L_{g^{-1}} \circ R_g \tag{15.11.8}$$

and the adjoint representation

$$\mathrm{Ad}_{g^{-1}} = \alpha_{g^{-1}*}\big|_{T_eG} . \tag{15.11.9}$$

The latter applies since $\mu(v) \in T_eG$. ■

Another interesting consequence is the following:

Theorem 15.11.2. *The Maurer-Cartan form μ is a Lie algebra homomorphism on the left-invariant vector fields, i.e., it satisfies*

$$\iota_{[X,Y]}\mu = [\iota_X\mu, \iota_Y\mu] \tag{15.11.10}$$

for all left-invariant vector fields $X, Y \in \mathrm{Vect}(G)$.

Proof. Given a left-invariant vector field $X \in \mathfrak{g}$, it follows that $\iota_X\mu \in \Omega^0(G, \mathfrak{g})$ is the constant Lie algebra valued function $g \mapsto X$ which assigns $X \in \mathfrak{g}$ to every $g \in G$. The bracket on the right hand side is thus to be interpreted pointwise, as

$$[\iota_X\mu, \iota_Y\mu](g) = [(\iota_X\mu)(g), (\iota_Y\mu)(g)] = [X, Y] = (\iota_{[X,Y]}\mu)(g) . \tag{15.11.11}$$

Since this holds for all $g \in G$, the statement follows. ■

This can now immediately be used to prove the following important property of the Maurer-Cartan form.

Theorem 15.11.3. *The Maurer-Cartan form μ satisfies the Maurer-Cartan equation*

$$\mathrm{d}\mu + \frac{1}{2}[\mu \wedge \mu] = 0 . \tag{15.11.12}$$

Proof. Let $X, Y \in \mathfrak{g}$ be two left-invariant vector fields on G. From the relation (9.4.5), which holds for arbitrary vector fields and one-forms, follows by component-wise linear extension to Lie algebra valued forms:

$$\iota_Y\iota_X\mathrm{d}\mu = X(\iota_Y\mu) - Y(\iota_X\mu) - \iota_{[X,Y]}\mu . \tag{15.11.13}$$

Since we have chosen X, Y to be left-invariant, it follows that $\iota_X\mu$ and $\iota_Y\mu$ are constant, and so their derivatives along X and Y vanish,

$$X(\iota_Y\mu) = Y(\iota_X\mu) = 0 . \tag{15.11.14}$$

Further, using the homomorphism property (15.11.10) and the property (15.10.12) of the exterior product of Lie algebra valued differential forms we have

$$\iota_{[X,Y]}\mu = [\iota_X\mu, \iota_Y\mu] = \frac{1}{2}\iota_Y\iota_X[\mu \wedge \mu] \,. \tag{15.11.15}$$

This implies

$$\iota_Y\iota_X\left(\mathrm{d}\mu + \frac{1}{2}[\mu \wedge \mu]\right) = 0 \,. \tag{15.11.16}$$

Since the left-invariant vector fields span the tangent space at every point, and this equation involves only the pointwise values of the vector fields through the interior product, but not their derivatives, it follows that the two-form in brackets must vanish identically. ■

We can use this definition to construct a coordinate expression for the Maurer-Cartan form on matrix Lie groups, such as those given in example 15.1.3.

> *Example* 15.11.1 (Maurer-Cartan form on matrix Lie groups). Let $G \subset M_{n,n}$ be a matrix group, and introduce the matrix components $(g^a{}_b)$ as coordinates on $M_{n,n}$. By imposing suitable restrictions on them, we obtain coordinates on G. For a fixed $g \in G$, the left translation $L_g : \tilde{g} \mapsto g\tilde{g}$ can be written as $\tilde{g}^a{}_b \mapsto g^a{}_c\tilde{g}^c{}_b$. It thus follows that the pushforward of a tangent vector $v = v^a{}_b\partial_a{}^b \in T_gG$ along $L_{g^{-1}}$ is given by
>
> $$\mu(g)(v) = L_{g^{-1}*}(v) = (g^{-1})^a{}_c v^c{}_b\partial_a{}^b \in T_eG \cong \mathfrak{g} \,. \tag{15.11.17}$$
>
> Thus, the Maurer-Cartan form is given by $\mu(g)^a{}_b = (g^{-1})^a{}_c \,\mathrm{d}g^c{}_b$. One often finds the shorthand notation $\mu = g^{-1}\,\mathrm{d}g$.

15.12 Fundamental vector fields

We finally define a helpful class of vector fields, which are defined by a group action G on a manifold M. For this purpose, recall that the Lie algebra $\mathfrak{g} = \operatorname{Lie} G$ of a Lie group G is given by the left invariant vector fields on G. These vector fields define vector fields on P as follows.

Definition 15.12.1 (Fundamental vector fields)**.** Let G be a Lie group with Lie algebra $\mathfrak{g}$ and M a manifold M. For a left action $\phi : G \times M \to M$ and a left invariant vector field $X \in \mathfrak{g}$ we define the *fundamental vector field* $\tilde{X}$ as the map

$$\begin{array}{rccc} \tilde{X} & : & M & \to & TM \\ & & x & \mapsto & (\phi_x)_*(-X(e)) \end{array} \,, \tag{15.12.1}$$

where $\phi_x : G \to M, g \mapsto \phi(g, x)$. For a right action $\theta : M \times G \to M$, we instead define

$$\begin{array}{rccc} \tilde{X} & : & M & \to & TM \\ & & x & \mapsto & (\theta_x)_*(X(e)) \end{array} \,, \tag{15.12.2}$$

where $\theta_x : G \to M, g \mapsto \theta(x, g)$.

To see that $\tilde{X}$ is indeed a vector field, we have to check that $\tau \circ \tilde{X} = \mathrm{id}_M$, where $\tau : TM \to M$ is the tangent bundle projection, and so $\tilde{X}(x) \in T_xM$ for all $x \in M$. To see this, note first that $X(e) \in T_eG$. Further, the differentials of ϕ_x and θ_x cover their respective defining maps, so that

$$\tau(\tilde{X}(x)) = \tau((\phi_x)_*(-X(e)) = \phi_x(\hat{\tau}(-X(e))) = \phi_x(e) = \phi(e, x) = x \,, \tag{15.12.3}$$

where we wrote $\hat{\tau} : TG \to G$ for the tangent bundle projection of G. Analogously, one checks the case of a right group action θ.

The fundamental vector fields satisfy an important relation.

Theorem 15.12.1. *The function* $\tilde{\bullet} : \mathfrak{g} \to \mathrm{Vect}(M)$*, which assigns to every Lie algebra element its fundamental vector field with respect to a given action on* M*, is a Lie algebra homomorphism, i.e.,*

$$\widetilde{[X,Y]} = [\tilde{X}, \tilde{Y}] \tag{15.12.4}$$

for all $X, Y \in \mathfrak{g}$.

Proof. ▶. . . ◀ ■

Here the square brackets on the left hand side denote the Lie bracket of $\mathfrak{g}$, i.e., the commutator of left invariant vector fields on G, while the square brackets on the right hand side denote the commutator of vector fields on M. The proof is not difficult, but we will omit it here.

Chapter 16

Lie derivative and flow

16.1 Flows of vector fields

In section 15.3 we have discussed the action of Lie groups on manifolds. We will now restrict ourselves to the action of a particular group, namely that of the real line $(\mathbb{R}, +)$. This is in fact an important special case, since any one-parameter subgroup $\varphi : \mathbb{R} \to G$ of a Lie group G acting on a manifold M induces an action of $(\mathbb{R}, +)$ on M, and all of these actions together describe (locally) the action of G. To study the local behavior of Lie group actions, and in particular those of $(\mathbb{R}, +)$, we start with a definition.

Definition 16.1.1 (Integral curve)**.** Let M be a manifold and X a vector field on M. An *integral curve* of X is a curve $\gamma \in C^\infty((a, b), M)$ with $a, b \in \mathbb{R}$ such that $\dot{\gamma}(t) = X(\gamma(t))$ for all $t \in (a, b)$.

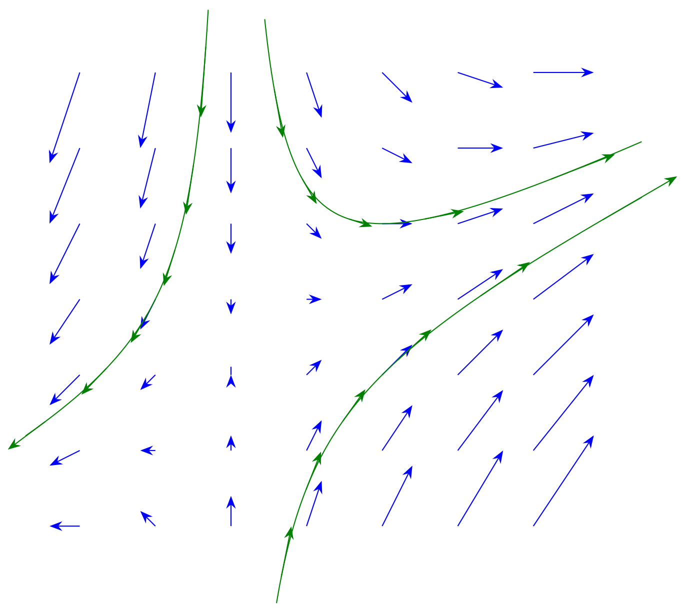


Figure 16.1: A vector field and some of its integral curves.

By making use of the canonical lift from definition 7.3.2, we may also state that an integral curve of a vector field X is a curve γ such that $\dot{\gamma} = X \circ \gamma$. One may ask whether such integral curves exist for any vector field. This is indeed the case, and is guaranteed by the following theorem, which comes from the theory of differential equations.

Theorem 16.1.1. *Let M be a manifold and X a vector field on M. For each $x \in M$ there exists an open set $U \subset M$ containing x, $\epsilon > 0$ and a map $\gamma : (-\epsilon, \epsilon) \times U \to M, (t, y) \mapsto \gamma_y(t)$ such that for all $y \in U$ the curve γ_y is an integral curve of X with $\gamma_y(0) = y$.*

Proof. ▶... ◀ ■

A case of particular interest is given when an integral curve can be defined for all on $\mathbb{R}$. For this case we define the following notion.

Definition 16.1.2 (Complete vector field)**.** A vector field X on a manifold M is called *complete* if for each $x \in M$ there exists an integral curve $\gamma \in C^\infty(\mathbb{R}, M)$ of X with $\gamma(0) = x$.

Note that not every vector field is complete, as the following example shows.

Example 16.1.1. Let $M = (0,1) \times (0,1)$ and $X = \partial_1$. This vector field is not complete.

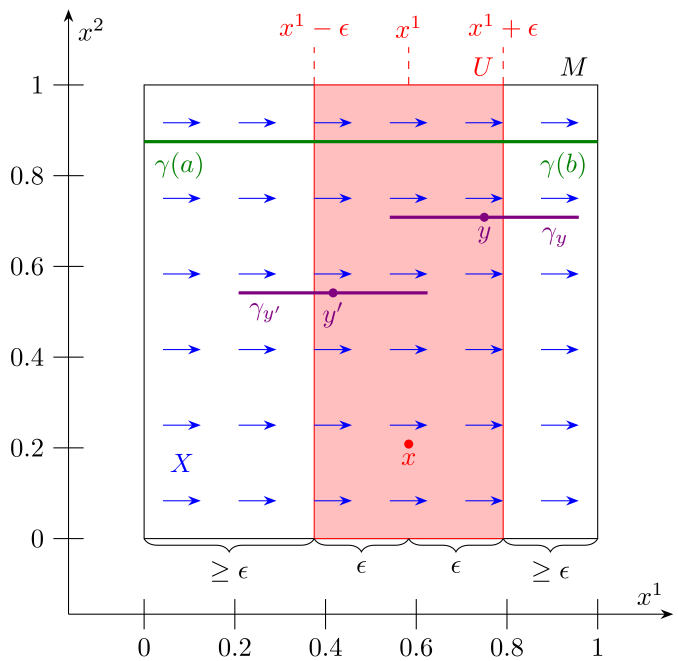


Figure 16.2: The vector field $X = \partial_1$ on the manifold $M = (0,1) \times (0,1)$ is not complete, since any integral curve γ has maximal domain $(a, b) \subset \mathbb{R}$ with $b - a \leq 1$, which cannot be extended to $\mathbb{R}$. However, around every point $x \in M$ exists an open set $U = (x^1 - \epsilon, x^1 + \epsilon) \times (0, 1)$, where $\epsilon = \frac{1}{2}\min(x^1, 1 - x^1)$, so that through every $y \in U$ one can find an integral curve γ_y with domain at least $(-\epsilon, \epsilon)$ and $\gamma(0) = y$. Note that ϵ depends only on x, but not on y, and that the integral curves do not necessarily lie inside of U.

Given a complete vector field, we can define the following notion.

Definition 16.1.3 (Flow). Let M be a manifold and X a complete vector field on M. The *flow* of X is the unique map $\phi : \mathbb{R} \times M \to M$ such that for each $x \in M$ the map $\phi_\bullet(x) : \mathbb{R} \to M$ is an integral curve of X and $\phi_0(x) = x$.

In fact, the flow can also be defined *locally* for a non-complete vector field. In this case it is simply a map from an open subset $U \in \mathbb{R} \times M$ to M, where $\{0\} \times M \subset U$. The flow has a number of nice properties, one of which can be written most nicely for complete vector fields.

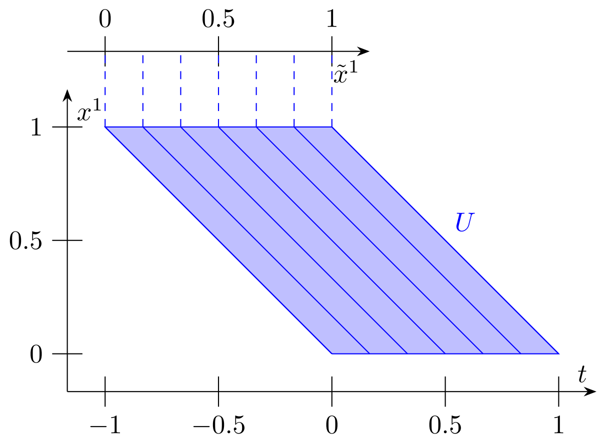


Figure 16.3: Visualization of the (local) flow $(\tilde{x}^1, \tilde{x}^2) = \phi_t(x^1, x^2) = (x^1 + t, x^2)$ of the vector field $X = \partial_1$ on $M = (0, 1) \times (0, 1)$ and its domain $U = \{(t, x) \in \mathbb{R} \times M, 0 < x^1 + t < 1\}$. The coordinate x^2 is not shown to simplify the diagram.

Theorem 16.1.2. *The flow of a complete vector field X is both a left and a right Lie group action of $(\mathbb{R}, +)$ on M.*

Proof. Since $(\mathbb{R}, +)$ is abelian, every left action is also a right action. We thus simply have to check that $\phi : \mathbb{R} \times M \to M$ is a smooth map such that $\phi_{s+t}(x) = \phi_s(\phi_t(x))$ for all $s, t \in \mathbb{R}$ and $x \in M$. We will not check the smoothness here. For fixed $t \in \mathbb{R}$ and $x \in M$ the maps $\gamma_1 : s \mapsto \phi_{s+t}(x)$ and $\gamma_2 : s \mapsto \phi_s(\phi_t(x))$ define curves on M. For these curves we have

$$\dot{\gamma}_1(s) = X(\phi_{s+t}(x)) = X(\gamma_1(s)), \tag{16.1.1a}$$

$$\dot{\gamma}_2(s) = X(\phi_s(\phi_t(x))) = X(\gamma_2(s)), \tag{16.1.1b}$$

so that both of them are integral curves of X. Further, they have the same initial point $\gamma_1(0) = \phi_t(x) = \gamma_2(0)$. Since integral curves are unique, it thus follows that $\gamma_1(s) = \gamma_2(s)$ for all $s \in \mathbb{R}$, and therefore $\phi_{s+t}(x) = \phi_s(\phi_t(x))$. ■

In fact, the relation $\phi_{s+t}(x) = \phi_s(\phi_t(x))$ holds also for local flows, whenever both sides are well-defined. This will be sufficient for the constructions in this chapter. However, note that the flow is a group action only for complete vector fields.

16.2 Lie derivative of tensor fields

Using the tools from the previous section we can now define a useful and important object in differential geometry; see [Yan57] for a thorough treatment.

Definition 16.2.1 (Lie derivative)**.** Let $T \in \Gamma(T^r_s M)$ be a tensor field and $X \in \mathrm{Vect}(M)$ a vector field on a manifold M. Let $\phi : \mathbb{R} \times M \supseteq U \to M$ be the flow of X. The *Lie derivative* of T with respect to X is the tensor field defined by

$$\mathcal{L}_X T = \lim_{t\to 0} \frac{\phi_t^* T - T}{t} . \tag{16.2.1}$$

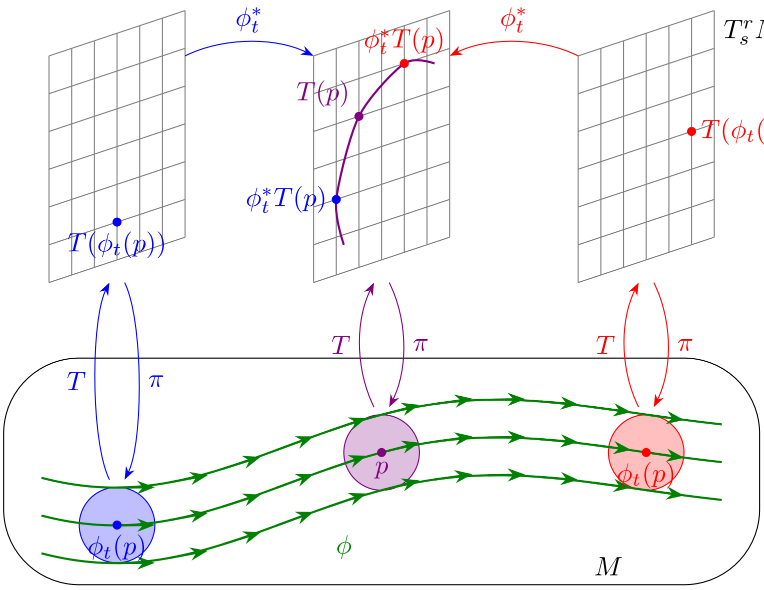


Figure 16.4: Illustration of the Lie derivative. For every t, pull back the tensor field T from $\phi_t(p)$ along ϕ_t. Evaluating this tensor field $\phi_t^* T$ at p yields an element $\phi_t^* T(p) \in T^r_{s\,p} M$. These elements form a curve $t \mapsto \phi_t^* T(p)$. The derivative of this curve with respect to t at $t = 0$ is the Lie derivative $\mathcal{L}_X T(p)$.

We see that the Lie derivative can be seen as the infinitesimal change of the tensor field T along the flow of X: starting from a point $x \in M$ one follows the flow line of X, takes the tensor field at that point $\phi_t(x)$, pulls it back along ϕ_t to obtain a tensor at the original point x and then measures how much this tensor at x changes with t. Note that for this it is not enough to know only the integral curve along which to move, but also the flow ϕ in a neighborhood around p is needed, since the pullback ϕ_t^* depends on the *derivatives* of ϕ.

Of course one has to show that the limit (16.2.1) really exists and that it yields a smooth tensor field. Instead of proving this here in a rigorous, coordinate-free way, we only illustrate the definition and derive the coordinate expression of the Lie derivative.

Using coordinates (x^a) on M, let $X = X^a \partial_a$ be a vector field and

$$T = T^{a_1\cdots a_r}{}_{b_1\cdots b_s} \partial_{a_1} \otimes \ldots \otimes \partial_{a_r} \otimes \mathrm{d}x^{b_1} \otimes \ldots \otimes \mathrm{d}x^{b_s} . \tag{16.2.2}$$

Writing the pullback $\phi_t^* T$ in the same coordinates as

$$\phi_t^* T = T'_t = T'^{a_1\cdots a_r}_t{}_{b_1\cdots b_s} \partial_{a_1} \otimes \ldots \otimes \partial_{a_r} \otimes \mathrm{d}x^{b_1} \otimes \ldots \otimes \mathrm{d}x^{b_s} . \tag{16.2.3}$$

With this notation the coordinate expression for the Lie derivative takes the form

$$(\mathcal{L}_X T)^{a_1\cdots a_r}{}_{b_1\cdots b_s}(x) = \lim_{t\to 0} \frac{T'^{a_1\cdots a_r}_t{}_{b_1\cdots b_s}(x) - T^{a_1\cdots a_r}{}_{b_1\cdots b_s}(x)}{t} = \frac{\mathrm{d}}{\mathrm{d}t} T'^{a_1\cdots a_r}_t{}_{b_1\cdots b_s}(x)\Big|_{t=0} . \tag{16.2.4}$$

To evaluate this derivative, recall that the pullback of a tensor field by a diffeomorphism is given by

$$T_t'^{a_1\cdots a_r}{}_{b_1\cdots b_s}(x) = T^{c_1\cdots c_r}{}_{d_1\cdots d_s}(x_t'(x))\frac{\partial x^{a_1}}{\partial x_t'^{c_1}}(x_t'(x))\cdots\frac{\partial x^{a_r}}{\partial x_t'^{c_r}}(x_t'(x))\frac{\partial x_t'^{d_1}}{\partial x^{b_1}}(x)\cdots\frac{\partial x_t'^{d_s}}{\partial x^{b_s}}(x)\,, \tag{16.2.5}$$

where we wrote the flow ϕ of X in the form $x_t'(x)$. It is related to the vector field X via the flow equation

$$X^a(x) = \left.\frac{\mathrm{d}}{\mathrm{d}t}x_t'^a(x)\right|_{t=0}. \tag{16.2.6}$$

This equation together with the chain rule is used to evaluate

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}T^{a_1\cdots a_r}{}_{b_1\cdots b_s}(x_t'(x))\right|_{t=0} = X^c(x)\partial_c T^{a_1\cdots a_r}{}_{b_1\cdots b_s}(x)\,. \tag{16.2.7}$$

From the fact that partial derivatives commute follows that

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial x_t'^b}{\partial x^a}(x)\right|_{t=0} = \partial_a\left.\frac{\mathrm{d}}{\mathrm{d}t}x_t'^b(x)\right|_{t=0} = \partial_a X^b(x)\,. \tag{16.2.8}$$

To evaluate the remaining term, we further use the fact that $\phi_t^{-1} = \phi_{-t}$, from which follows that

$$\frac{\partial x^a}{\partial x_t'^b}(x_t'(x)) = \frac{\partial x_{-t}'^a}{\partial x^b}(x_t'(x))\,. \tag{16.2.9}$$

When we differentiate this term with respect to t, we must pay attention that it appears both in the argument $x_t'(x)$ and explicitly in the function $\partial x_{-t}'^a/\partial x^b$ itself, so that we obtain the two terms

$$\begin{aligned}\left.\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial x_{-t}'^a}{\partial x^b}(x_t'(x))\right|_{t=0} &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial x_{-t}'^a}{\partial x^b}(x_0'(x))\right|_{t=0} + \frac{\partial}{\partial x_c}\frac{\partial x_0'^a}{\partial x^b}(x)\left.\frac{\mathrm{d}}{\mathrm{d}t}x_t'^c(x)\right|_{t=0}\\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial x_{-t}'^a}{\partial x^b}(x)\right|_{t=0} + \partial_c\delta_b^a X_t^c(x)\,.\end{aligned} \tag{16.2.10}$$

The second term vanishes, since the Kronecker symbol δ_b^a is constant, and so $\partial_c\delta_b^a = 0$. For the first term, we can change the order of differentiation and find

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial x_{-t}'^a}{\partial x^b}(x)\right|_{t=0} = \frac{\partial}{\partial x^b}\left.\frac{\mathrm{d}}{\mathrm{d}t}x_{-t}'^a(x)\right|_{t=0} = -\partial_b X^a(x)\,. \tag{16.2.11}$$

Putting everything together we finally find the coordinate expression for the Lie derivative as

$$\begin{aligned}(\mathcal{L}_X T)^{a_1\cdots a_r}{}_{b_1\cdots b_s} = {} & X^c\partial_c T^{a_1\cdots a_r}{}_{b_1\cdots b_s}\\ & - \partial_c X^{a_1}T^{ca_2\cdots a_r}{}_{b_1\cdots b_s} - \ldots - \partial_c X^{a_r}T^{a_1\cdots a_{r-1}c}{}_{b_1\cdots b_s}\\ & + \partial_{b_1}X^c T^{a_1\cdots a_r}{}_{cb_2\cdots b_s} + \ldots + \partial_{b_s}X^c T^{a_1\cdots a_r}{}_{b_1\cdots b_{s-1}c}\,.\end{aligned} \tag{16.2.12}$$

The Lie derivative of tensor fields has a few helpful and important properties. We start with a few properties which concern its application to tensor fields, which are constructed from other tensor fields, as follows.

Theorem 16.2.1. *Let M be a manifold, S,T tensor fields on M, X a vector field on M, $\mu,\nu \in \mathbb{R}$ and $k,l \in \mathbb{R}$. The Lie derivative satisfies:*

1. *Linearity in the tensor fields:*

$$\mathcal{L}_X(\mu S + \nu T) = \mu\mathcal{L}_X S + \nu\mathcal{L}_X T\,. \tag{16.2.13}$$

2. *Leibniz rule:*

$$\mathcal{L}_X(S\otimes T) = (\mathcal{L}_X S)\otimes T + S\otimes(\mathcal{L}_X T)\,. \tag{16.2.14}$$

3. *Compatibility with contraction:*

$$\mathcal{L}_X(\mathrm{tr}_l^k\, T) = \mathrm{tr}_l^k(\mathcal{L}_X T)\,. \tag{16.2.15}$$

Proof. These properties are easily derived from the properties of the pullback of tensor fields. The first property follows immediately from the linearity of the pullback, while the third one follows from the fact that it commutes with the contraction. Finally, the second one follows from the Leibniz rule

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(\phi_t^*(S)\otimes\phi_t^*(T))\Big|_{t=0} &= \frac{\mathrm{d}}{\mathrm{d}t}\phi_t^*(S)\Big|_{t=0}\otimes\phi_0^*(T)+\phi_0^*(S)\otimes\frac{\mathrm{d}}{\mathrm{d}t}\phi_t^*(T)\Big|_{t=0}\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\phi_t^*(S)\Big|_{t=0}\otimes T+S\otimes\frac{\mathrm{d}}{\mathrm{d}t}\phi_t^*(T)\Big|_{t=0}\,,\end{aligned} \tag{16.2.16}$$

using $\phi_0 = \mathrm{id}_M$. ■

These properties are easily illustrated using the coordinate expression 16.2.12. First, note that the latter is linear in the components of the tensor field T. To illustrate the Leibniz rule for the tensor product, consider a vector field $Y = Y^a\partial_a$ and a one-form $\omega = \omega_a \mathrm{d}x^a$. In this case the Lie derivative reads

$$\begin{aligned}\mathcal{L}_X(Y\otimes\omega) &= \mathcal{L}_X(Y^a\omega_b\partial_a\otimes\mathrm{d}x^b)\\ &= [X^c\partial_c(Y^a\omega_b)-\partial_cX^aY^c\omega_b+\partial_bX^cY^a\omega_c]\partial_a\otimes\mathrm{d}x^b\\ &= [X^c\partial_cY^a\omega_b-\partial_cX^aY^c\omega_b+X^cY^a\partial_c\omega_b+\partial_bX^cY^a\omega_c]\partial_a\otimes\mathrm{d}x^b\\ &= \mathcal{L}_XY\otimes\omega+Y\otimes\mathcal{L}_X\omega\,.\end{aligned} \tag{16.2.17}$$

Finally, for the contraction, consider a tensor field $T = T^a{}_b\partial_a\otimes\mathrm{d}x^b$, and observe that

$$\begin{aligned}\mathrm{tr}_1^1\,\mathcal{L}_XT &= \mathrm{tr}_1^1\,\mathcal{L}_X(T^a{}_b\partial_a\otimes\mathrm{d}x^b)\\ &= \mathrm{tr}_1^1[(X^c\partial_cT^a{}_b-\partial_cX^aT^c{}_b+\partial_bX^cT^a{}_c)\partial_a\otimes\mathrm{d}x^b]\\ &= X^c\partial_cT^d{}_d-\partial_cX^dT^c{}_d+\partial_dX^cT^d{}_c\\ &= X^c\partial_cT^d{}_d\\ &= \mathcal{L}_X\,\mathrm{tr}_1^1\,T\,.\end{aligned} \tag{16.2.18}$$

Further, we may also construct a new vector field from other vector fields. Also the Lie derivatives with respect to these different vector fields are related, as we show below.

Theorem 16.2.2. *Let M be a manifold, T a tensor field on M, X, Y vector fields on M and $\mu,\nu\in\mathbb{R}$. The Lie derivative satisfies:*

1. *Linearity in the vector fields:*

$$\mathcal{L}_{\mu X+\nu Y}T = \mu\mathcal{L}_XT+\nu\mathcal{L}_YT\,. \tag{16.2.19}$$

2. *Commutator:*

$$\mathcal{L}_{[X,Y]}T = \mathcal{L}_X\mathcal{L}_YT-\mathcal{L}_Y\mathcal{L}_XT\,. \tag{16.2.20}$$

In other words, the Lie derivative is a Lie algebra homomorphism from the Lie algebra Vect(M) *of vector fields on M to the algebra of linear operators acting on tensor fields on M, where the latter is equipped with the commutator as Lie bracket.*

Proof. ►. . .◄ ■

►Show in coordinates.◄

In the following sections, we will discuss a few examples for the Lie derivative of particular tensor fields.

16.3 Lie derivative of real functions

The simplest possible tensor field is of course a tensor field of type $(0,0)$, which is the same as a real function. In this case the Lie derivative has a very simple form.

Theorem 16.3.1. *For a function $f \in C^\infty(M,\mathbb{R})$ the Lie derivative with respect to a vector field X is given by*

$$\mathcal{L}_X f = Xf\,. \tag{16.3.1}$$

Proof. For a function f the pullback along the flow ϕ of X is given by

$$(\phi_t^* f)(p) = f(\phi_t(p)) \tag{16.3.2}$$

for every $p \in M$ and $t \in \mathbb{R}$, such that (t,p) lies in the domain of ϕ. Denoting by $\gamma_p : t \mapsto \phi_t(p)$ the flow line which passes through p at $t = 0$, we thus have

$$(\mathcal{L}_X f)(p) = \frac{\mathrm{d}}{\mathrm{d}t}(\phi_t^* f)(p)\bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}f(\phi_t(p))\bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}f(\gamma_p(t))\bigg|_{t=0} = \dot{\gamma}_p(0)(f) = X(p)(f)\,, \tag{16.3.3}$$

where we used the definition of the tangent vector of a curve and the flow $\dot{\gamma}_p(0) = X(p)$. Since this holds for all $p \in M$, one has $\mathcal{L}_X f = Xf$. ■

In other words, the Lie derivative of a function reduces to the action of a vector field. From this follow a few useful properties of the Lie derivative in this special case.

Theorem 16.3.2. *For a vector field $X \in \mathrm{Vect}(M)$ and real functions $f, g \in C^\infty(M,\mathbb{R})$ on a manifold M the Lie derivative satisfies:*

1. *Leibniz rule:*
$$\mathcal{L}_X(fg) = \mathcal{L}_X f \cdot g + f \cdot \mathcal{L}_X g\,. \tag{16.3.4}$$

2. *Multiplication of the vector field:*
$$\mathcal{L}_{gX} f = g \cdot \mathcal{L}_X f\,. \tag{16.3.5}$$

Proof. The Leibniz rule follows immediately from the Leibniz rule for tensor fields, since for real functions we simply have $f \otimes g = fg$. The second property follows from the fact that $(gX)(p) = g(p)X(p)$ and hence $((gX)f)(p) = g(p)(Xf)(p)$ for all $p \in M$, since the multiplication is performed pointwise. ■

It is important to note that the second property holds only for functions, i.e., tensors of rank $(0,0)$, and not for other tensor fields.

16.4 Lie derivative of vector fields

As the next example we discuss the Lie derivative of vector fields. Also in this case it reduces to a familiar object as follows.

Theorem 16.4.1. *For a vector field $Y \in \mathrm{Vect}(M)$ the Lie derivative with respect to a vector field X is given by*

$$\mathcal{L}_X Y = [X,Y]\,. \tag{16.4.1}$$

Proof. From the fact that the Lie derivative of a tensor fields yields a tensor field of the same rank follows that $\mathcal{L}_X Y$ is again a vector field. Hence, it is fully determined by its action on functions. Let $f \in C^\infty(M, \mathbb{R})$. Then we have

$$\begin{aligned}
(\mathcal{L}_X Y)f &= \frac{\mathrm{d}}{\mathrm{d}t}\phi_t^* Y\bigg|_{t=0} f \\
&= \frac{\mathrm{d}}{\mathrm{d}t}(\phi_t^* Y)f\bigg|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}(Y(f \circ \phi_t^{-1})) \circ \phi_t\bigg|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}(Y(f \circ \phi_{-t})) \circ \phi_t\bigg|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}(Yf) \circ \phi_t\bigg|_{t=0} + \frac{\mathrm{d}}{\mathrm{d}t}Y(f \circ \phi_{-t})\bigg|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\phi_t^*(Yf)\bigg|_{t=0} - Y\,\frac{\mathrm{d}}{\mathrm{d}t}\phi_t^* f\bigg|_{t=0} \\
&= \mathcal{L}_X(Yf) - Y\mathcal{L}_X f \\
&= X(Yf) - Y(Xf) \\
&= [X, Y]f\,. \quad ■
\end{aligned} \tag{16.4.2}$$

From the fact that $\mathrm{Vect}(M)$ together with the Lie bracket forms a Lie algebra one can derive the following properties of the Lie derivative of vector fields.

Theorem 16.4.2. *For vector fields $X, Y, Z \in \mathrm{Vect}(M)$ on a manifold M the Lie derivative satisfies:*

1. *Antisymmetry:*
$$\mathcal{L}_X Y = -\mathcal{L}_Y X\,. \tag{16.4.3}$$
2. *Jacobi identity:*
$$\mathcal{L}_X[Y, Z] = [\mathcal{L}_X Y, Z] + [Y, \mathcal{L}_X Z]\,. \tag{16.4.4}$$

Proof. These follow directly from the identification 16.4.1 of the Lie derivative with the commutator of vector fields and the statement 7.5.1 that the latter equips the vector fields with the structure of a Lie algebra, so that the antisymmetry and Jacobi identity hold. ■

The second relation can be brought into various different forms.

16.5 Lie derivative of differential forms

Another special case for the Lie derivative which we discuss in this chapter is the Lie derivative of differential forms. Also in this case there exists a helpful formula for the Lie derivative in terms of objects we have already previously encountered.

Theorem 16.5.1. *For a k-form $\omega \in \Omega^k(M)$ with $k \geq 1$ the Lie derivative with respect to a vector field X is given by "Cartan's magic formula"*

$$\mathcal{L}_X\omega = \iota_X \mathrm{d}\omega + \mathrm{d}\iota_X\omega\,. \tag{16.5.1}$$

Proof. One can use the fact that any differential form can be constructed as a linear combination of exterior products of zero-forms and (exact) one-forms, for example by decomposing it using a partition of unity, and expressing each component in coordinates in the form (9.1.2), where the coordinates and the coefficients of the differential form can be regarded as zero-forms. This allows us to perform the proof inductively in three steps:

1. Let $f \in \Omega^0(M)$ be a real function. Then $\iota_X f = 0$ and $\iota_X \mathrm{d}f = Xf$. Since also $\mathcal{L}_X f = Xf$, the formula holds for zero-forms.

2. For a general k-form ω, we find

$$\mathcal{L}_X \mathrm{d}\omega = \lim_{t\to 0} \frac{\phi_t^* \mathrm{d}\omega - \mathrm{d}\omega}{t} = \lim_{t\to 0} \frac{\mathrm{d}\phi_t^* \omega - \mathrm{d}\omega}{t} = \mathrm{d} \lim_{t\to 0} \frac{\phi_t^* \omega - \omega}{t} = \mathrm{d}\mathcal{L}_X \omega \,, \tag{16.5.2}$$

using the properties 11.3.1 of the pullback of a differential form, and the fact that differentiation with respect to the parameter t and the partial derivatives constituting d commute. Assuming that Cartan's magic formula holds for ω, we then have

$$\begin{aligned} \mathcal{L}_X \mathrm{d}\omega &= \mathrm{d}\mathcal{L}_X \omega \\ &= \mathrm{d}(\iota_X \mathrm{d}\omega + \mathrm{d}\iota_X \omega) \\ &= \mathrm{d}\iota_X \mathrm{d}\omega \\ &= \mathrm{d}\iota_X \mathrm{d}\omega + \iota_X \mathrm{d}\mathrm{d}\omega \,, \end{aligned} \tag{16.5.3}$$

and so it holds for $\mathrm{d}\omega$ as well. Hence, setting $\omega = f \in \Omega^0(M)$, it holds in particular for all exact one-forms $\mathrm{d}f$.

3. Finally, one can show that

$$\mathcal{L}_X(\omega \wedge \sigma) = (\mathcal{L}_X \omega) \wedge \sigma + \omega \wedge (\mathcal{L}_X \sigma) \tag{16.5.4}$$

for any exterior product of two differential forms ω and σ, by again using the properties 11.3.1, and following exactly the same steps as in the proof of the equality (16.2.14). Assuming that Cartan's magic formula holds for ω and σ, one thus has

$$\begin{aligned} \mathcal{L}_X(\omega \wedge \sigma) &= (\mathcal{L}_X \omega) \wedge \sigma + \omega \wedge (\mathcal{L}_X \sigma) \\ &= (\iota_X \mathrm{d}\omega + \mathrm{d}\iota_X \omega) \wedge \sigma + \omega \wedge (\iota_X \mathrm{d}\sigma + \mathrm{d}\iota_X \sigma) \\ &= \iota_X \mathrm{d}(\omega \wedge \sigma) + \mathrm{d}\iota_X(\omega \wedge \sigma) \,, \end{aligned} \tag{16.5.5}$$

where the intermediate steps taken to obtain the last line, which involve only the properties of the interior product and exterior derivative, are shown in detail in the proof of theorem 16.5.2.

Now using the fact that any differential form can be composed as a linear composition of exterior products of zero-forms and exact one-forms, and that Cartan's magic formula is linear on both sides and holds for any of these building blocks, we conclude that it holds for all differential forms. ∎

Cartan's magic formula turns out to be a special case of a more general class of operators acting on differential forms, called graded derivations; these will be discussed in full detail in chapter 17, where we also proof a more general statement. One can directly use the formula 16.5.1 and the properties of the operations on differential forms to derive the following properties of the Lie derivative.

Theorem 16.5.2. *For vector fields $X, Y \in \mathrm{Vect}(M)$, differential forms $\omega \in \Omega^k(M), \sigma \in \Omega^l(M)$ and functions $f \in C^\infty(M, \mathbb{R})$ on a manifold M the Lie derivative satisfies:*

1. *Compatibility with exterior derivative:*

$$\mathrm{d}\mathcal{L}_X \omega = \mathcal{L}_X \mathrm{d}\omega \,. \tag{16.5.6}$$

2. *Leibniz rule with exterior product:*

$$\mathcal{L}_X(\omega \wedge \sigma) = (\mathcal{L}_X \omega) \wedge \sigma + \omega \wedge (\mathcal{L}_X \sigma) \,. \tag{16.5.7}$$

3. *Relation with interior product:*

$$\iota_{[X,Y]}\omega = \mathcal{L}_X\iota_Y\omega - \iota_Y\mathcal{L}_X\omega = \iota_X\mathcal{L}_Y\omega - \mathcal{L}_Y\iota_X\omega\,. \tag{16.5.8}$$

4. *Distribution law:*

$$\mathcal{L}_{fX}\omega = f\mathcal{L}_X\omega + \mathrm{d}f \wedge \iota_X\omega\,. \tag{16.5.9}$$

Proof. The first two statements have been shown explicitly in the proof of theorem 16.5.1; Given that Cartan's magic formula is now proven, we can also show them as well as the remaining statements as follows:

1. Keeping in mind that $\mathrm{d}^2 = 0$, we have

$$\mathrm{d}\mathcal{L}_X\omega = \mathrm{d}(\iota_X\mathrm{d}\omega + \mathrm{d}\iota_X\omega) = \mathrm{d}\iota_X\mathrm{d}\omega = \iota_X\mathrm{d}\mathrm{d}\omega + \mathrm{d}\iota_X\mathrm{d}\omega = \mathcal{L}_X\mathrm{d}\omega\,. \tag{16.5.10}$$

2. By direct calculation, we have

$$\begin{aligned}
\mathcal{L}_X(\omega\wedge\sigma) &= \iota_X\mathrm{d}(\omega\wedge\sigma) + \mathrm{d}\iota_X(\omega\wedge\sigma)\\
&= \iota_X[\mathrm{d}\omega\wedge\sigma + (-1)^k\omega\wedge\mathrm{d}\sigma] + \mathrm{d}[\iota_X\omega\wedge\sigma + (-1)^k\omega\wedge\iota_X\sigma]\\
&= \iota_X\mathrm{d}\omega\wedge\sigma - (-1)^k\mathrm{d}\omega\wedge\iota_X\sigma + (-1)^k\iota_X\omega\wedge\mathrm{d}\sigma + \omega\wedge\iota_X\mathrm{d}\sigma\\
&\quad + \mathrm{d}\iota_X\omega\wedge\sigma - (-1)^k\iota_X\omega\wedge\mathrm{d}\sigma + (-1)^k\mathrm{d}\omega\wedge\iota_X\sigma + \omega\wedge\mathrm{d}\iota_X\sigma\\
&= \iota_X\mathrm{d}\omega\wedge\sigma + \omega\wedge\iota_X\mathrm{d}\sigma + \mathrm{d}\iota_X\omega\wedge\sigma + \omega\wedge\mathrm{d}\iota_X\sigma\\
&= (\mathcal{L}_X\omega)\wedge\sigma + \omega\wedge(\mathcal{L}_X\sigma)\,.
\end{aligned} \tag{16.5.11}$$

3. Comparing (16.5.1) with the relation (9.4.6), one immediately obtains

$$\iota_{[X,Y]}\omega = \mathrm{d}\iota_X\iota_Y\omega + \iota_X\mathrm{d}\iota_Y\omega - \iota_Y\mathrm{d}\iota_X\omega - \iota_Y\iota_X\mathrm{d}\omega = \mathcal{L}_X\iota_Y\omega - \iota_Y\mathcal{L}_X\omega\,. \tag{16.5.12}$$

   Alternatively, one may change the order of the interior products, which yields a minus sign, to obtain

$$\iota_{[X,Y]}\omega = -\mathrm{d}\iota_Y\iota_X\omega + \iota_X\mathrm{d}\iota_Y\omega - \iota_Y\mathrm{d}\iota_X\omega + \iota_X\iota_Y\mathrm{d}\omega = \iota_X\mathcal{L}_Y\omega - \mathcal{L}_Y\iota_X\omega\,. \tag{16.5.13}$$

4. Using $\iota_{fX}\omega = f\iota_X\omega$, we have

$$\begin{aligned}
\mathcal{L}_{fX}\omega &= \iota_{fX}\mathrm{d}\omega + \mathrm{d}\iota_{fX}\omega\\
&= f\iota_X\mathrm{d}\omega + \mathrm{d}(f\iota_X\omega)\\
&= f\iota_X\mathrm{d}\omega + \mathrm{d}f\wedge\iota_X\omega + f\mathrm{d}\iota_X\omega\\
&= f\mathcal{L}_X\omega + \mathrm{d}f\wedge\iota_X\omega\,. \quad \blacksquare
\end{aligned} \tag{16.5.14}$$

Note in particular that if ω is a zero-form, one has $\iota_X\omega = 0$, and so the last relation reduces to (16.3.5).

16.6 Lie derivative of endomorphisms

We then come to the Lie derivative of sections of the homomorphism bundle $\mathrm{End}(TM) = \mathrm{Hom}(TM, TM) \cong TM \otimes T^*M$. As for the other cases, the Lie derivative $\mathcal{L}_XF$ of a section $F \in \Gamma(\mathrm{End}(TM))$ of this bundle is again a tensor field of the same type. Further, recall from section 4.6 that one can apply such a section to a vector field $Y \in \mathrm{Vect}(M)$. One may naturally ask whether it is possible to express $(\mathcal{L}_XF)Y$ in terms of FY. This is indeed the case, and we state as follows.

Theorem 16.6.1. *Given vector fields $X, Y \in \mathrm{Vect}(M)$ and a section $F \in \Gamma(\mathrm{End}(TM))$ of the endomorphism bundle, the Lie derivative satisfies*

$$(\mathcal{L}_X F)Y = [X, FY] - F[X, Y] \,. \tag{16.6.1}$$

Proof. Viewing F as a section of $TM \otimes T^*M$, one can express FY by first taking the tensor product $F \otimes Y \in \Gamma(TM \otimes T^*M \otimes TM)$, and then contracting the second upper with the first lower index, $FY = \mathrm{tr}_1^2(F \otimes Y)$. The latter commutes with the Lie derivative, while the former obeys the Leibniz rule, so that one finds

$$\mathcal{L}_X(\mathrm{tr}_1^2(F \otimes Y)) = \mathrm{tr}_1^2(\mathcal{L}_X F \otimes Y) + \mathrm{tr}_1^2(F \otimes \mathcal{L}_X Y) = (\mathcal{L}_X F)Y + F(\mathcal{L}_X Y) \,. \tag{16.6.2}$$

Solving for the first term on the right hand side and using theorem 16.4.1 one finds

$$(\mathcal{L}_X F)Y = \mathcal{L}_X(FY) - F(\mathcal{L}_X Y) = [X, FY] - F[X, Y] \,. \tag{16.6.3}$$

■

This formula is easily illustrated in coordinates (x^a). Writing

$$X = X^a \partial_a \,, \quad Y = Y^a \partial_a \,, \quad F = F^a{}_b \partial_a \otimes \mathrm{d}x^b \,, \tag{16.6.4}$$

one has

$$\begin{aligned} [X, FY] - F[X, Y] &= [X^b \partial_b (F^a{}_c Y^c) - F^b{}_c Y^c \partial_b X^a - F^a{}_b (X^c \partial_c Y^b - Y^c \partial_c X^b)] \partial_a \\ &= (X^b \partial_b F^a{}_c - \partial_b X^a F^b{}_c + \partial_c X^b F^a{}_b) Y^c \partial_a \,, \end{aligned} \tag{16.6.5}$$

since the derivatives acting on the components of Y cancel. The term in brackets corresponds to the Lie derivative

$$\mathcal{L}_X F = (X^c \partial_c F^a{}_b - \partial_c X^a F^c{}_b + \partial_b X^c F^a{}_c) \partial_a \otimes \mathrm{d}x^b \,, \tag{16.6.6}$$

as follows from the relation (16.2.12).

Chapter 17

Graded derivations

17.1 Graded derivations

We have encountered several operations which act on differential forms, such as the exterior derivative, interior product and Lie derivative. We will now show that these are special cases of a more general set of operations, essentially following the treatment in [KSM93, sec. 8]. Other references are [SLK14] and [Ant03, ch. 11]. We define these operations as follows.

Definition 17.1.1 (Graded derivation)**.** Let M be a manifold and $\Omega^\bullet(M)$ the differential forms on M. A *graded derivation* of degree $n \in \mathbb{Z}$ on M is a linear function $D : \Omega^\bullet(M) \to \Omega^\bullet(M)$ such that for all $\omega \in \Omega^k(M), \sigma \in \Omega^l(M)$ holds $D\omega \in \Omega^{k+n}(M)$ and

$$D(\omega \wedge \sigma) = D\omega \wedge \sigma + (-1)^{kn}\omega \wedge D\sigma\,. \tag{17.1.1}$$

The space of all graded derivations of degree n on M is denoted $\operatorname{Der}_n \Omega(M)$, and the space of all derivations is denoted $\operatorname{Der}_\bullet \Omega(M)$.

In fact, we have already encountered several graded derivations. This we can state as follows.

Theorem 17.1.1. *The following operations are graded derivations on the differential forms on a manifold M:*

1. *The exterior derivative* d *is a graded derivation of degree* 1.
2. *For any vector field $X \in \operatorname{Vect}(M)$, the interior product ι_X is a graded derivation of degree* -1.
3. *For any vector field $X \in \operatorname{Vect}(M)$, the Lie derivative $\mathcal{L}_X$ is a graded derivation of degree* 0.

Proof. The three mentioned operations are linear by definition. They further satisfy:

1. By definition 9.3.1, the exterior derivative increases the degree of a form by 1 and satisfies the graded Leibniz rule.
2. By definition 9.4.1, the same holds for the interior product, except that it decreases the degree of a form by 1.

3. The Lie derivative of a tensor field is a section of the same bundle as the original vector field, and hence the degree does not change. The graded Leibniz rule (which is just the Leibniz rule, since $n = 0$ is even), follows from theorem 16.5.2. ■

In the definition 17.1.1 we have allowed arbitrary integers as the degree of a graded derivation. One may ask whether any non-trivial derivations exist for any given degree. Before answering this question, we will show a helpful statement, which we will also be able to use later.

Theorem 17.1.2. *Any graded derivation $D \in \mathrm{Der}_\bullet \Omega(M)$ is uniquely determined by its action on smooth functions $f \in \Omega^0(M)$ and their differentials $\mathrm{d}f$.*

Proof. Recall that we can write any differential k-form ω in a basis expansion of the form (9.1.2), where the coefficients $\omega_{a_1 \cdots a_k}$ are zero-forms while the basis elements $\mathrm{d}x^a$ are one-forms, which can be seen as the differentials of the coordinate functions. Since the action of a graded derivation on the exterior product of differential forms is uniquely determined by the action on the individual factors, it follows that it is uniquely determined by the action on the coefficient zero-forms and differentials of coordinate functions. ■

We can now use this result in the following proof.

Theorem 17.1.3. *Any graded derivation $D \in \mathrm{Der}_n \Omega(M)$ of degree n on a manifold of dimension $\dim M = m$ with $n < -1$ or $n > m$ is trivial.*

Proof. For $n > m$ it follows that $D\omega \in \Omega^{k+n}(M)$ for a k-form $\omega \in \Omega^k(M)$ has $k + n > m$ and must therefore vanish. Similarly, for $n < -1$ follows that D must act trivially on $\Omega^0(M)$ and $\Omega^1(M)$, and hence must be trivial by theorem 17.1.2. ■

17.2 Graded commutator

One may pose the question how the space of graded derivations can be described and whether it carries any structure. The latter is indeed the case. To see this, we continue with the following statement.

Theorem 17.2.1. *For any derivations $D \in \mathrm{Der}_m \Omega(M)$ and $D' \in \mathrm{Der}_n \Omega(M)$, one has*

$$D \circ D' - (-1)^{mn} D' \circ D \in \mathrm{Der}_{m+n} \Omega(M)\,. \tag{17.2.1}$$

Proof. The expression given above is obviously linear, since the composition of linear functions is linear, and increases the degree of a form by $m + n$. This leaves to check the graded Leibniz rule, which reads

$$\begin{aligned}
DD'(\omega \wedge \sigma) - (-1)^{mn} D'D(\omega \wedge \sigma) &= D[D'\omega \wedge \sigma + (-1)^{kn} \omega \wedge D'\sigma] \\
&\quad - (-1)^{mn} D'[D\omega \wedge \sigma + (-1)^{km} \omega \wedge D\sigma] \\
&= DD'\omega \wedge \sigma + (-1)^{m(k+n)} D'\omega \wedge D\sigma \\
&\quad + (-1)^{kn} D\omega \wedge D'\sigma + (-1)^{k(m+n)} \omega \wedge DD'\sigma \\
&\quad - (-1)^{mn} D'D\omega \wedge \sigma - (-1)^{kn} D\omega \wedge D'\sigma \\
&\quad - (-1)^{m(k+n)} D'\omega \wedge D\sigma - (-1)^{km+kn+mn} \omega \wedge D'D\sigma \\
&= (DD' - (-1)^{mn} D'D)\omega \wedge \sigma \\
&\quad + (-1)^{k(m+n)} \omega \wedge (DD' - (-1)^{mn} D'D)\sigma \\
&= [D, D']\omega \wedge \sigma + (-1)^{k(m+n)} \omega \wedge [D, D']\sigma
\end{aligned} \tag{17.2.2}$$

for $\omega \in \Omega^k(M)$ and $\sigma \in \Omega^l(M)$. Hence, it is a graded derivation of degree $m + n$. ■

The expression given in the previous theorem carries a name and notation.

Definition 17.2.1 (Graded commutator)**.** Let $D \in \mathrm{Der}_m \Omega(M)$ and $D' \in \mathrm{Der}_n \Omega(M)$ be two graded derivations on a manifold M. Their *graded commutator* is the graded derivation

$$[D, D'] = D \circ D' - (-1)^{mn} D' \circ D \in \mathrm{Der}_{m+n} \Omega(M) . \tag{17.2.3}$$

With this definition in place, we can now study the algebraic structure of the space of derivations. We find the following statement.

Theorem 17.2.2. *The space*

$$\mathrm{Der}\, \Omega(M) = \bigoplus_{n \in \mathbb{Z}} \mathrm{Der}_n \Omega(M) \tag{17.2.4}$$

of graded derivations on a manifold M together with the graded commutator is a graded Lie algebra, i.e., the graded commutator is bilinear and satisfies:

1. *graded antisymmetry:*
$$[D_1, D_2] = -(-1)^{n_1 n_2} [D_2, D_1] , \tag{17.2.5}$$
2. *graded Jacobi identity:*
$$[D_1, [D_2, D_3]] = [[D_1, D_2], D_3] + (-1)^{n_1 n_2} [D_2, [D_1, D_3]] , \tag{17.2.6}$$

where $D_1 \in \mathrm{Der}_{n_1} \Omega(M), D_2 \in \mathrm{Der}_{n_2} \Omega(M), D_3 \in \mathrm{Der}_{n_3} \Omega(M)$.

Proof. Bilinearity is obvious from the definition 17.2.1, and so we will only check the symmetry and the Jacobi identity. For the former we find by direct computation:

$$\begin{aligned} [D_1, D_2] &= D_1 \circ D_2 - (-1)^{n_1 n_2} D_2 \circ D_1 \\ &= -(-1)^{n_1 n_2} (D_2 \circ D_1 - (-1)^{n_1 n_2} D_1 \circ D_2) \\ &= -(-1)^{n_1 n_2} [D_2, D_1] . \end{aligned} \tag{17.2.7}$$

Similarly, we find the Jacobi identity

$$\begin{aligned} [D_1, [D_2, D_3]] &= D_1 \circ [D_2, D_3] - (-1)^{n_1(n_2+n_3)} [D_2, D_3] \circ D_1 \\ &= D_1 \circ D_2 \circ D_3 - (-1)^{n_2 n_3} \underline{D_1 \circ D_3 \circ D_2} \\ &\quad - (-1)^{n_1(n_2+n_3)} (\underline{D_2 \circ D_3 \circ D_1} - (-1)^{n_2 n_3} \underline{\underline{D_3 \circ D_2 \circ D_1}}) \\ &= D_1 \circ D_2 \circ D_3 - (-1)^{n_1 n_2} \underline{D_2 \circ D_1 \circ D_3} \\ &\quad - (-1)^{(n_1+n_2)n_3} (\underline{D_3 \circ D_1 \circ D_2} - (-1)^{n_1 n_2} \underline{\underline{D_3 \circ D_2 \circ D_1}}) \\ &\quad + (-1)^{n_1 n_2} (\underline{D_2 \circ D_1 \circ D_3} - (-1)^{n_1 n_3} \underline{D_2 \circ D_3 \circ D_1}) \\ &\quad - (-1)^{n_2 n_3} (\underline{D_1 \circ D_3 \circ D_2} - (-1)^{n_1 n_3} \underline{D_3 \circ D_1 \circ D_2}) \\ &= [D_1, D_2] \circ D_3 - (-1)^{(n_1+n_2)n_3} D_3 \circ [D_1, D_2] \\ &\quad + (-1)^{n_1 n_2} (D_2 \circ [D_1, D_3] - (-1)^{n_2(n_1+n_3)} [D_1, D_3] \circ D_2) \\ &= [[D_1, D_2], D_3] + (-1)^{n_1 n_2} [D_2, [D_1, D_3]] , \end{aligned} \tag{17.2.8}$$

where dotted and dashed terms cancel each other, and we have used different underlines for the remaining, matching terms for clarity. ■

In fact, we have already encountered an example for the graded commutator of derivations, namely the following.

Theorem 17.2.3. *For any vector fields $X, Y \in \mathrm{Vect}(M)$, the graded commutators are given by:*

$$[\mathrm{d}, \mathrm{d}] = 0\,, \tag{17.2.9a}$$
$$[\iota_X, \iota_Y] = 0\,, \tag{17.2.9b}$$
$$[\mathcal{L}_X, \mathcal{L}_Y] = \mathcal{L}_{[X,Y]}\,, \tag{17.2.9c}$$
$$[\iota_X, \mathrm{d}] = \mathcal{L}_X\,, \tag{17.2.9d}$$
$$[\mathcal{L}_X, \mathrm{d}] = 0\,, \tag{17.2.9e}$$
$$[\mathcal{L}_X, \iota_Y] = \iota_{[X,Y]}\,. \tag{17.2.9f}$$

Proof. By direct calculation with $\omega \in \Omega^k(M)$ we find:

1. For the exterior derivative holds
$$[\mathrm{d}, \mathrm{d}]\omega = \mathrm{d}\mathrm{d}\omega + \mathrm{d}\mathrm{d}\omega = 0\,. \tag{17.2.10}$$
2. The interior product satisfies
$$[\iota_X, \iota_Y]\omega = \iota_X\iota_Y\omega + \iota_Y\iota_X\omega\,, \tag{17.2.11}$$
since it is antisymmetric.
3. For the Lie derivative one finds
$$[\mathcal{L}_X, \mathcal{L}_Y]\omega = \mathcal{L}_X\mathcal{L}_Y\omega - \mathcal{L}_Y\mathcal{L}_X\omega = \mathcal{L}_{[X,Y]}\omega\,, \tag{17.2.12}$$
using the commutator (16.2.20).
4. The interior product and exterior derivative yield
$$[\iota_X, \mathrm{d}]\omega = \iota_X\mathrm{d}\omega + \mathrm{d}\iota_X\omega = \mathcal{L}_X\omega \tag{17.2.13}$$
due to Cartan's formula (16.5.1).
5. The Lie derivative and exterior derivative commute,
$$[\mathcal{L}_X, \mathrm{d}]\omega = \mathcal{L}_X\mathrm{d}\omega - \mathrm{d}\mathcal{L}_X\omega = 0\,, \tag{17.2.14}$$
as given by equation (16.5.6) .
6. Finally, for the Lie derivative and exterior product one has
$$[\mathcal{L}_X, \iota_Y]\omega = \mathcal{L}_X\iota_Y\omega - \iota_Y\mathcal{L}_X\omega = \iota_{[X,Y]}\omega\,, \tag{17.2.15}$$
due to the rule (16.5.8). ■

We will use this result in section 17.5 in order to generalize the notion of the Lie derivative.

17.3 Algebraic derivations

We now come to a particular class of graded derivations, namely those which vanish if they are applied to a scalar function, i.e., to an element $f \in \Omega^0(M) \cong C^\infty(M, \mathbb{R})$. We denote these derivations as follows.

Definition 17.3.1 (Algebraic derivation)**.** A graded derivation $D \in \mathrm{Der}_n \Omega(M)$ of degree n on a manifold M is called *algebraic* if it satisfies $Df = 0$ for all $f \in \Omega^0(M)$. The space of all algebraic derivations of degree n is denoted $\mathrm{Der}^0_n \Omega(M)$, while the space of all algebraic derivations is denoted $\mathrm{Der}^0_\bullet \Omega(M)$.

The question arises whether the algebraic graded derivations are closed under taking the graded commutator, and thus form a graded Lie subalgebra. This is indeed the case, which we see as follows.

Theorem 17.3.1. *The graded commutator $[D, D']$ of two algebraic graded derivations D, D' is again algebraic.*

Proof. By definition, we have

$$[D, D']f = DD'f - (-1)^{mn} D'Df = 0 \tag{17.3.1}$$

for $D \in \mathrm{Der}^0_m \Omega(M)$ and $D' \in \mathrm{Der}^0_n \Omega(M)$. ■

We have already encountered a class of algebraic graded derivations. By definition, the interior product $\iota_X f$ of a vector field $X \in \mathrm{Vect}(M)$ and a function $f \in \Omega^0(M)$ vanishes, and thus turns ι_X into an algebraic graded derivation of degree -1. We now construct an extension of the interior product. Instead of vector fields $X \in \mathrm{Vect}(M) = \Omega^0(M, TM)$, we shall consider arbitrary vector-valued differential forms. For these we can define as follows.

Definition 17.3.2 (Insertion operator)**.** Let M be a manifold. The *insertion operator* is the function

$$\begin{array}{rcccc} \iota & : & \Omega^k(M, TM) \times \Omega^l(M) & \to & \Omega^{k+l-1}(M) \\ & & (K, \omega) & \mapsto & \iota_K \omega \end{array}, \tag{17.3.2}$$

such that for any $K \in \Omega^k(M, TM)$, $\omega \in \Omega^l(M)$ and vector fields $X_1, \ldots, X_{k+l-1} \in \mathrm{Vect}(M)$ holds

$$\begin{aligned} &\iota_K \omega(X_1, \ldots, X_{k+l-1}) = \\ &\qquad \frac{1}{k!(l-1)!} \sum_{\sigma \in S_{k+l-1}} \mathrm{sgn}(\sigma) \omega(K(X_{\sigma(1)}, \ldots, X_{\sigma(k)}), X_{\sigma(k+1)}, \ldots, X_{\sigma(k+l-1)}) . \end{aligned} \tag{17.3.3}$$

Besides the name *insertion operator* [KSM93, sec. 7.7], also the name *substitution operator* (see [SLK14, rem. 3.3.29] or [Ant03, sec. 11.1.E]) is used in the literature. From its definition, we can find a helpful formula by first considering the case $k = 0$, where we $K = X \in \Omega^0(M, TM) = \mathrm{Vect}(M)$ is a vector field. In this case we have that

$$\iota_X \omega(X_1, \ldots, X_{l-1}) = \frac{1}{(l-1)!} \sum_{\sigma \in S_{l-1}} \mathrm{sgn}(\sigma) \omega(X, X_{\sigma(1)}, \ldots, X_{\sigma(l-1)}) = \omega(X, X_1, \ldots, X_{l-1}) \tag{17.3.4}$$

is the usual interior product. Setting

$$X = K(X_{\sigma(1)}, \ldots, X_{\sigma(k)}) \tag{17.3.5}$$

in the general formula, we find

$$\iota_K\omega(X_1,\ldots,X_{k+l-1}) = \frac{1}{k!(l-1)!}\sum_{\sigma\in S_{k+l-1}}\operatorname{sgn}(\sigma)\left[K(X_{\sigma(1)},\ldots,X_{\sigma(k)}) \lrcorner\, \omega\right](X_{\sigma(k+1)},\ldots,X_{\sigma(k+l-1)})\,. \tag{17.3.6}$$

The question arises whether also ι_K for $K \in \Omega^k(M,TM)$ is an algebraic graded derivation. We answer this question together with an even stronger statement, claiming that every algebraic graded derivation is of the form ι_K.

Theorem 17.3.2. *The insertion operator ι establishes a one-to-one correspondence between vector-valued k-forms $K \in \Omega^k(M,TM)$ on a manifold M and algebraic graded derivations $D \in \operatorname{Der}^0_{k-1}\Omega(M)$.*

Proof. We proceed in three steps:

1. Let $K \in \Omega^k(M,TM)$ be a vector-valued k-form and denote by ι_K the insertion operator from definition 17.3.3. This operator obviously acts linearly on differential forms $\omega \in \Omega^l(M)$, and yields a differential form $\iota_K\omega \in \Omega^{k+l-1}(M)$. To see that it is indeed a graded derivation, we still need to check the Leibniz rule. For $\omega \in \Omega^l(M), \sigma \in \Omega^m(M)$ we have

$$\iota_K(\omega\wedge\sigma)(X_1,\ldots,X_{k+l+m-1}) = \frac{1}{k!(l+m-1)!}\cdot \sum_{\sigma\in S_{k+l+m-1}}\operatorname{sgn}(\sigma)\left[K(X_{\sigma(1)},\ldots,X_{\sigma(k)}) \lrcorner\, (\omega\wedge\sigma)\right](X_{\sigma(k+1)},\ldots,X_{\sigma(k+l+m-1)})\,. \tag{17.3.7}$$

We then consider the differential form in square brackets, which is given by

$$\begin{aligned} K_X \lrcorner\, (\omega\wedge\sigma) &= (K_X \lrcorner\, \omega)\wedge\sigma + (-1)^l\omega\wedge(K_X \lrcorner\, \sigma)\\ &= (K_X \lrcorner\, \omega)\wedge\sigma + (-1)^l(-1)^{l(m-1)}(K_X \lrcorner\, \sigma)\wedge\omega\\ &= (K_X \lrcorner\, \omega)\wedge\sigma + (-1)^{lm}(K_X \lrcorner\, \sigma)\wedge\omega\,, \end{aligned} \tag{17.3.8}$$

where we abbreviated

$$K_X = K(X_{\sigma(1)},\ldots,X_{\sigma(k)})\,. \tag{17.3.9}$$

Substituting back into the defining formula for the insertion operator, and factoring out the second factor, we thus obtain

$$\begin{aligned} \iota_K(\omega\wedge\sigma) &= \iota_K\omega\wedge\sigma + (-1)^{lm}\iota_K\sigma\wedge\omega\\ &= \iota_K\omega\wedge\sigma + (-1)^{lm}(-1)^{l(k+m-1)}\omega\wedge\iota_K\sigma\\ &= \iota_K\omega\wedge\sigma + (-1)^{l(k-1)}\omega\wedge\iota_K\sigma\,. \end{aligned} \tag{17.3.10}$$

Hence, $\iota_K \in \operatorname{Der}_{k-1}\Omega(M)$ is a derivation of degree $k-1$. Finally, $\iota_K f = 0$ for any $f \in \Omega^0(M)$, and so $\iota_K \in \operatorname{Der}^0_{k-1}\Omega(M)$ is an algebraic derivation.

2. To show the converse direction, let $D \in \operatorname{Der}^0_{k-1}\Omega(M)$ be an algebraic derivation. Let $K \in \Omega^k(M,TM)$ be the unique vector-valued k-form such that for all one-forms $\omega \in \Omega^1(M)$ and vector fields $X_1,\ldots,X_k \in \operatorname{Vect}(M)$ holds

$$\omega(K(X_1,\ldots,X_k)) = (D\omega)(X_1,\ldots,X_k)\,. \tag{17.3.11}$$

To see that this is well-defined, we have to show that the left hand side is linear (with respect to multiplication by functions $f \in C^\infty(M)$) in ω and $X_1,\ldots,X_k$, as well as antisymmetric in the latter. For the vector fields this is obvious, since $D\omega \in \Omega^k(M)$. For ω, one uses the fact that D is an algebraic derivation, and so

$$D(f\omega) = Df\wedge\omega + fD\omega = fD\omega\,. \tag{17.3.12}$$

Hence, K indeed defines a vector-valued k-form. One also sees from definition 17.3.2 that the left hand side is simply

$$\omega(K(X_1,\ldots,X_k)) = \iota_K\omega(X_1,\ldots,X_k)\,, \tag{17.3.13}$$

as intended. Hence, we have $D\omega = \iota_K\omega$. Since this holds for all one-forms ω, and an algebraic derivation is uniquely defined by its action on one-forms, it follows that $D = \iota_K$.

3. We finally need to show that the vector-valued k-form K constructed above is unique. If we have $K' \in \Omega^k(M, TM)$ with $\iota_{K'} = D = \iota_K$, then it follows immediately from the last equation that

$$\omega(K'(X_1,\ldots,X_k)) = \iota_{K'}\omega(X_1,\ldots,X_k) = \iota_K\omega(X_1,\ldots,X_k) = \omega(K(X_1,\ldots,X_k)) \tag{17.3.14}$$

for all one-form ω and vector fields $X_1,\ldots,X_k$, and so $K = K'$. ■

Hence, we can identify the space of algebraic graded derivations with the space of vector-valued differential forms. To illustrate this construction, we take a look at a particular example. Recall that the tensor product bundle $\Omega^1(M, TM) \cong TM \otimes T^*M \cong \mathrm{End}(TM)$ admits in particular the unit section δ, given in definition 5.5.1 by the identity map id_{T_xM} on every tangent space. In this case we find the following result.

Theorem 17.3.3. *For the unit section $\delta \in \Omega^1(M, TM)$ and $\omega \in \Omega^k(M)$ holds $\iota_\delta\omega = k\omega$.*

Proof. By direct calculation we find

$$\begin{aligned}\iota_\delta\omega(X_1,\ldots,X_k) &= \frac{1}{(k-1)!}\sum_{\sigma\in S_k}\mathrm{sgn}(\sigma)\omega(X_{\sigma(1)},\ldots,X_{\sigma(k)})\\ &= \frac{k!}{(k-1)!}\omega(X_1,\ldots,X_k)\\ &= k\omega(X_1,\ldots,X_k)\,. \quad ■\end{aligned} \tag{17.3.15}$$

Also here we find in particular that for $k = 0$ the result vanishes, so that it is indeed an algebraic graded derivation. Another helpful set of formulas is the following.

Theorem 17.3.4. *For all $K \in \Omega^k(M, TM)$, $\omega \in \Omega^l(M)$ and $f \in C^\infty(M)$ holds*

$$\iota_{fK}\omega = \iota_K(f\omega) = f\iota_K\omega\,. \tag{17.3.16}$$

Proof. This follows immediately by direct calculation from the definition 17.3.2. ■

17.4 Nijenhuis-Richardson bracket

We have seen in section 17.3 that the algebraic graded derivations, together with their graded commutator, form a graded Lie algebra. Further, we have seen that there is a one-to-one correspondence between such algebraic graded derivations and vector-valued differential forms. Via this one-to-one correspondence, we may equip the space of vector-valued differential forms itself with an algebra structure, simply by defining an algebra relation as follows.

Definition 17.4.1 (Nijenhuis-Richardson bracket). Let M be a manifold and $K \in \Omega^k(M, TM)$, $L \in \Omega^l(M, TM)$ vector-valued differential forms on M. Their *Nijenhuis-Richardson bracket* (or *algebraic bracket*) is the unique vector-valued form $[K, L]^\wedge \in \Omega^{k+l-1}(M, TM)$ such that

$$\iota_{[K,L]^\wedge} = [\iota_K, \iota_L]\,. \tag{17.4.1}$$

This is well-defined, since the right hand side is, by construction, an algebraic graded derivation, and so there indeed exists a unique vector-valued form $[K,L]^\wedge \in \Omega^{k+l-1}(M,TM)$, which follows from the fact that ι is an isomorphism of vector spaces. The Nijenhuis-Richardson bracket turns ι into an isomorphism of graded Lie algebras. Here the algebra structure on $\Omega^\bullet(M,TM)$ has the following properties.

Theorem 17.4.1. *The Nijenhuis-Richardson bracket equips the space*

$$\Omega^\bullet(M,TM) = \bigoplus_{k=0}^{\dim M} \Omega^k(M,TM) \tag{17.4.2}$$

of vector-valued differential forms on a manifold M with the structure of a graded Lie algebra, such that:

1. *graded antisymmetry:*
$$[K_1,K_2]^\wedge = -(-1)^{(k_1-1)(k_2-1)}[K_2,K_1]^\wedge\,, \tag{17.4.3}$$
2. *graded Jacobi identity:*
$$[K_1,[K_2,K_3]^\wedge]^\wedge = [[K_1,K_2]^\wedge,K_3]^\wedge + (-1)^{(k_1-1)(k_2-1)}[K_2,[K_1,K_3]^\wedge]^\wedge\,, \tag{17.4.4}$$

where $K_1 \in \Omega^{k_1}(M,TM), K_2 \in \Omega^{k_2}(M,TM), K_3 \in \Omega^{k_3}(M,TM)$.

Proof. This follows directly from the definition 17.4.1 of the Nijenhuis-Richardson bracket, from the fact that ι is a vector space isomorphism and that ι_{K_i} is a graded derivation of degree $k_i - 1$. ■

One may pose the question whether the Nijenhuis-Richardson bracket of two vector-valued differential forms can also be expressed directly in terms of operators acting on these forms, without resorting to the graded commutator. It turns out that this is indeed the case. However, in order to arrive at this result, we need another definition, which we state as follows.

Definition 17.4.2 (Insertion in vector-valued forms). Let M be a manifold and $K \in \Omega^k(M,TM)$ a vector-valued differential form. The *insertion* operator $\iota_K : \Omega^l(M) \to \Omega^{k+l-1}(M)$ is extended to act on vector-valued differential forms $\iota_K : \Omega^l(M,TM) \to \Omega^{k+l-1}(M,TM)$ such that for all $L \in \Omega^l(M,TM)$ and vector fields $X_1,\ldots,X_{k+l-1} \in \mathrm{Vect}(M)$ holds

$$\begin{aligned}\iota_K L(X_1,\ldots,X_{k+l-1}) &= \\ &\frac{1}{k!(l-1)!}\sum_{\sigma\in S_{k+l-1}} \mathrm{sgn}(\sigma) L(K(X_{\sigma(1)},\ldots,X_{\sigma(k)}),X_{\sigma(k+1)},\ldots,X_{\sigma(k+l-1)})\,.\end{aligned} \tag{17.4.5}$$

In other words, we define the insertion operator to act on vector-valued forms by acting exactly as in definition 17.3.2, with the only difference that (17.3.3) takes values in $\mathbb{R}$, while (17.4.5) takes values in TM. We give a simple example.

Theorem 17.4.2. *Let $\delta \in \Omega^1(M,TM)$ be the unit section and $K \in \Omega^k(M,TM)$. Then the insertion operator satisfies*

$$\iota_\delta K = kK\,, \quad \iota_K\delta = K\,. \tag{17.4.6}$$

Proof. By direct calculation we find

$$\begin{aligned}\iota_\delta K(X_1,\dots,X_k) &= \frac{1}{(k-1)!}\sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)K(X_{\sigma(1)},\dots,X_{\sigma(k)})\\ &= \frac{k!}{(k-1)!}K(X_1,\dots,X_k)\\ &= kK(X_1,\dots,X_k)\,,\end{aligned}\tag{17.4.7}$$

as well as

$$\begin{aligned}\iota_K\delta(X_1,\dots,X_k) &= \frac{1}{k!}\sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)\delta(K(X_{\sigma(1)},\dots,X_{\sigma(k)}))\\ &= \frac{1}{k!}\sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)K(X_{\sigma(1)},\dots,X_{\sigma(k)})\\ &= K(X_1,\dots,X_k)\,. \quad ■\end{aligned}\tag{17.4.8}$$

As with the usual insertion operator, the following holds.

Theorem 17.4.3. *For all $K\in\Omega^k(M,TM)$, $L\in\Omega^l(M,TM)$ and $f\in C^\infty(M)$ holds*

$$\iota_{fK}L=\iota_K(fL)=f\iota_K L\,.\tag{17.4.9}$$

Proof. This follows immediately by direct calculation from the definition 17.4.2. ■

With this definition in place, we can now give an explicit formula for the Nijenhuis-Richardson bracket.

Theorem 17.4.4. *Let M be a manifold and $K\in\Omega^k(M,TM)$, $L\in\Omega^l(M,TM)$ vector-valued differential forms on M. Their Nijenhuis-Richardson bracket satisfies*

$$[K,L]^\wedge=\iota_K L-(-1)^{(k-1)(l-1)}\iota_L K\,.\tag{17.4.10}$$

Proof. Let $\omega\in\Omega^1(M)$ be a one-form. Then we have

$$\begin{aligned}[\iota_K,\iota_L]\omega &= \iota_K\iota_L\omega-(-1)^{(k-1)(l-1)}\iota_L\iota_K\omega\\ &= \iota_{\iota_K L}\omega-(-1)^{(k-1)(l-1)}\iota_{\iota_L K}\omega\,.\end{aligned}\tag{17.4.11}$$

Since an algebraic derivation is uniquely determined by its action on one-forms, the statement follows. ■

It is also instructive to calculate a simple example.

Theorem 17.4.5. *The Nijenhuis-Richardson bracket of any vector-valued differential form $K\in\Omega^k(M,TM)$ with the unit section $\delta\in\Omega^1(M,TM)$ satisfies $[\delta,K]^\wedge=(k-1)K$.*

Proof. For $\omega\in\Omega^l(M)$ we have

$$\begin{aligned}\iota_{[\delta,K]^\wedge}\omega &= [\iota_\delta,\iota_K]\omega\\ &= \iota_\delta\iota_K\omega-(-1)^{0\cdot(k-1)}\iota_K\iota_\delta\omega\\ &= (k+l-1)\iota_K\omega-l\iota_K\omega\\ &= (k-1)\iota_K\omega\,,\end{aligned}\tag{17.4.12}$$

where we used the fact that ι_δ and ι_K are graded derivations of degree 0 and $k-1$, respectively, the graded commutator 17.2.1, the insertion 17.3.3 of the unit section and the fact that $\iota_K\omega\in\Omega^{k+l-1}(M)$. Hence, by the uniqueness of the Nijenhuis-Richardson bracket, we have $[\delta,K]^\wedge=(k-1)K$. ■

As mentioned in definition 17.4.1, the Nijenhuis-Richardson bracket is also called algebraic bracket. This relates to the fact that its value is determined by the values of the vector-valued forms only at the point where it is evaluated, but not on their derivatives. This is formally expressed by the following statement.

Theorem 17.4.6. *For all $K \in \Omega^k(M,TM)$, $L \in \Omega^l(M,TM)$ and $f \in C^\infty(M)$ holds*

$$[fK,L]^\wedge = [K,fL]^\wedge = f[K,L]^\wedge\,. \tag{17.4.13}$$

Proof. From theorem 17.4.4 and theorem 17.4.3 follows

$$[fK,L]^\wedge = \iota_{fK}L - (-1)^{(k-1)(l-1)}\iota_L(fK) = f(\iota_K L - (-1)^{(k-1)(l-1)}\iota_L K) = f[K,L]^\wedge \tag{17.4.14}$$

and

$$[K,fL]^\wedge = \iota_K(fL) - (-1)^{(k-1)(l-1)}\iota_{fL}K = f(\iota_K L - (-1)^{(k-1)(l-1)}\iota_L K) = f[K,L]^\wedge\,. \tag{17.4.15}$$

■

17.5 Nijenhuis-Lie derivative

In definition 17.3.2 we have defined the insertion operator as a generalization of the interior product from definition 9.4.1: instead of taking a vector field in order to obtain a graded derivation on the space of differential forms, it takes a vector-valued differential form. One may ask whether also the Lie derivative, which similarly takes a vector field and yields a derivation on the space of differential forms following theorem 17.1.1, can be generalized to vector-valued differential forms instead of vector fields. A hint towards a positive answer to this question comes from theorem 17.2.3, which expresses the Lie derivative as a graded commutator. A straightforward generalization arises from replacing the interior product in the commutator by the insertion operator. Hence, we arrive at the following definition.

Definition 17.5.1 (Nijenhuis-Lie derivative)**.** Let M be a manifold and $K \in \Omega^k(M,TM)$ a vector-valued differential form of degree k on M. Its *Nijenhuis-Lie derivative* (or *Lie derivation*) is the derivation

$$\mathcal{L}_K = [\iota_K, \mathrm{d}] \in \mathrm{Der}_k\,\Omega(M)\,. \tag{17.5.1}$$

Obviously, $\mathcal{L}_K$ is a graded derivation of degree k. Also the special case $k = 0$ reduces straightforwardly to the well-known Lie derivative.

Theorem 17.5.1. *Let M be a manifold and $X \in \mathrm{Vect}(M) = \Omega^0(M,TM)$ a vector field on M. Then the Nijenhuis-Lie derivative reduces to the ordinary Lie derivative $\mathcal{L}_X$ on differential forms.*

Proof. This follows from the fact that in this case the insertion operator in definition 17.5.1 reduces to the interior product, so that theorem 17.2.3 applies. ■

From the usual Lie derivative we know that it commutes with the exterior derivative. Since both the Lie derivative and the exterior derivative are derivations, this statement refers to their graded commutator. We introduce the following nomenclature for this class of derivations.

Definition 17.5.2 (Lie derivation)**.** A derivation $D \in \mathrm{Der}_n\,\Omega(M)$ of degree n is called a *Lie derivation* if and only if $[D,\mathrm{d}] = 0$. The space of all Lie derivations of degree n is denoted $\mathrm{Der}^{\mathrm{d}}_n\,\Omega(M)$, while the space of all Lie derivations is denoted $\mathrm{Der}^{\mathrm{d}}_\bullet\,\Omega(M)$.

In the following we will study the properties of Lie derivations. Recall from theorem 17.1.2 that any graded derivation is uniquely determined by their action on functions and their differentials. For Lie derivatives, an even stronger statement holds.

Theorem 17.5.2. *Any Lie derivation $D \in \operatorname{Der}^{\mathrm{d}}_{\bullet} \Omega(M)$ is uniquely determined by its action on smooth functions $f \in \Omega^0(M)$.*

Proof. For $D \in \operatorname{Der}^{\mathrm{d}}_n \Omega(M)$ it follows from $[D, \mathrm{d}] = 0$ that

$$D \circ \mathrm{d} = (-1)^n \mathrm{d} \circ D\,. \tag{17.5.2}$$

The action of D on the differential $\mathrm{d}f$ of $f \in \Omega^0(M)$ is thus given by

$$D\mathrm{d}f = (-1)^n \mathrm{d}Df\,, \tag{17.5.3}$$

and thus uniquely determined by the action of D on f. Hence, also D is uniquely determined by its action on functions. ■

We have previously seen that the commutator of algebraic derivations is again algebraic. We now ask whether this applies also to Lie derivations.

Theorem 17.5.3. *The graded commutator $[D, D']$ of two Lie derivations D, D' is again a Lie derivation.*

Proof. From the graded Jacobi identity for $D \in \operatorname{Der}^{\mathrm{d}}_k \Omega(M)$ and $D \in \operatorname{Der}^{\mathrm{d}}_l \Omega(M)$ follows

$$[\mathrm{d}, [D, D']] = [[\mathrm{d}, D], D'] + (-1)^{kl}[D', [\mathrm{d}, D']] = 0\,. \tag{17.5.4}$$

■

We may now ask whether also the Nijenhuis-Lie derivative is a Lie derivation. As we have seen for the insertion operator in theorem 17.3.2, an even stronger statement holds, and we find that this is also the case for the Nijenhuis-Lie derivative. We formulate this as follows.

Theorem 17.5.4. *The Nijenhuis-Lie derivative $\mathcal{L}$ establishes a one-to-one correspondence between vector-valued k-forms $K \in \Omega^k(M, TM)$ on a manifold M and Lie derivations $D \in \operatorname{Der}^{\mathrm{d}}_k \Omega(M)$.*

Proof. The proof will be done in three steps:

1. Let $K \in \Omega^k(M, TM)$. Then its Nijenhuis-Lie derivative satisfies

   $$\mathcal{L}_K = [\iota_K, \mathrm{d}] \in \operatorname{Der}_k \Omega(M)\,, \tag{17.5.5}$$

   since $\iota_K \in \operatorname{Der}_{k-1} \Omega(M)$ and $\mathrm{d} \in \operatorname{Der}_1 \Omega(M)$. The commutator with d reads

   $$\begin{aligned}[\mathcal{L}_K, \mathrm{d}] &= [[\iota_K, \mathrm{d}], \mathrm{d}] \\ &= [\iota_K, \mathrm{d}] \circ \mathrm{d} - (-1)^k \mathrm{d} \circ [\iota_K, \mathrm{d}] \\ &= \left(\iota_K \circ \mathrm{d} - (-1)^{k-1} \mathrm{d} \circ \iota_K\right) \circ \mathrm{d} - (-1)^k \mathrm{d} \circ \left(\iota_K \circ \mathrm{d} - (-1)^{k-1} \mathrm{d} \circ \iota_K\right) \\ &= -(-1)^{k-1} \mathrm{d} \circ \iota_K \circ \mathrm{d} - (-1)^k \mathrm{d} \circ \iota_K \circ \mathrm{d} \\ &= 0\,,\end{aligned} \tag{17.5.6}$$

   and so $\mathcal{L}_K \in \operatorname{Der}^{\mathrm{d}}_k \Omega(M)$.

2. To show the converse direction, let $D \in \operatorname{Der}^{\mathrm{d}}_k \Omega(M)$ and define $K \in \Omega^k(M, TM)$ as the vector-valued k-form which assigns to vector fields $X_1, \ldots, X_k$ the vector field $K(X_1, \ldots, X_k)$ acting on functions by

   $$K(X_1, \ldots, X_k)(f) = (Df)(X_1, \ldots, X_k)\,. \tag{17.5.7}$$

Note that the right hand side indeed defines the action of a vector field on f, since D is linear and satisfies the Leibniz rule, and the insertion of vector fields in the k-form Df is linear. To see that D is the Nijenhuis-Lie derivative of K, we calculate

$$\begin{aligned}(\mathcal{L}_K f)(X_1, \ldots, X_k) &= (\iota_K \mathrm{d}f)(X_1, \ldots, X_k) \\ &= K(X_1, \ldots, X_k) \lrcorner \mathrm{d}f \\ &= K(X_1, \ldots, X_k)(f) \\ &= (Df)(X_1, \ldots, X_k).\end{aligned} \tag{17.5.8}$$

Since $\mathcal{L}_K$ and D are Lie derivations, and thus uniquely determined by their action on functions, it follows that $\mathcal{L}_K = D$.

3. We finally need to show that K above is the unique vector-valued k-form whose Nijenhuis-Lie derivative is D. Let $K' \in \Omega^k(M, TM)$ such that $\mathcal{L}_{K'} = D = \mathcal{L}_K$. From the derivation above then follows

$$K(X_1, \ldots, X_k)(f) = (Df)(X_1, \ldots, X_k) = K'(X_1, \ldots, X_k)(f). \tag{17.5.9}$$

Since this holds for all functions f, we see that

$$K(X_1, \ldots, X_k) = K'(X_1, \ldots, X_k) \tag{17.5.10}$$

is the same vector field. Since this holds for all vector fields $X_1, \ldots, X_k$, we find that $K = K'$. ■

To study an explicit case, we may consider the same example as for the insertion operator, and pose the question what is the Nijenhuis-Lie derivative with respect to the canonical unit section of $\Omega^1(M, TM)$. Here we find the following result.

Theorem 17.5.5. *For the unit section $\delta \in \Omega^1(M, TM)$ holds $\mathcal{L}_\delta = \mathrm{d}$.*

Proof. By direct calculation, we find

$$\begin{aligned}\mathcal{L}_\delta \omega &= [\iota_\delta, \mathrm{d}]\omega \\ &= \iota_\delta \mathrm{d}\omega - \mathrm{d}\iota_\delta \omega \\ &= (k+1)\mathrm{d}\omega - k\mathrm{d}\omega \\ &= \mathrm{d}\omega\end{aligned} \tag{17.5.11}$$

for $\omega \in \Omega^k(M)$, where we use the result 17.3.3 and the fact that $\mathrm{d}\omega \in \Omega^{k+1}(M)$, since d is a graded derivation of degree 1. ■

We know that the ordinary Lie derivative satisfies a number of useful formulas for its commutator with itself or the interior product, as well as multiplication with scalar functions. Here we take a look at the latter, and calculate the following.

Theorem 17.5.6. *For all $K \in \Omega^k(M, TM)$, $\omega \in \Omega^l(M)$ and $f \in C^\infty(M)$ holds*

$$\mathcal{L}_K(f\omega) = \mathcal{L}_K f \wedge \omega + f\mathcal{L}_K \omega \tag{17.5.12}$$

and

$$\mathcal{L}_{fK}\omega = f\mathcal{L}_K\omega + (-1)^k \mathrm{d}f \wedge \iota_K \omega\,. \tag{17.5.13}$$

Proof. By direct calculation and using theorem 17.3.4 we find

$$\begin{aligned}\mathcal{L}_K(f\omega) &= [\iota_K, \mathrm{d}](f\omega) \\ &= \iota_K \mathrm{d}(f\omega) - (-1)^{k-1}\mathrm{d}\iota_K(f\omega) \\ &= \iota_K(\mathrm{d}f \wedge \omega + f\mathrm{d}\omega) - (-1)^{k-1}\mathrm{d}(f\iota_K\omega) \\ &= \iota_K \mathrm{d}f \wedge \omega + (-1)^{k-1}\mathrm{d}f \wedge \iota_K\omega + f\iota_K\mathrm{d}\omega - (-1)^{k-1}(\mathrm{d}f \wedge \iota_K\omega + f\mathrm{d}\iota_K\omega) \\ &= \iota_K\mathrm{d}f \wedge \omega + f[\iota_K\mathrm{d}\omega - (-1)^{k-1}\mathrm{d}\iota_K\omega] \\ &= \mathcal{L}_K f \wedge \omega + f\mathcal{L}_K\omega\end{aligned} \tag{17.5.14}$$

and

$$\begin{aligned}
\mathcal{L}_{fK}\omega &= [\iota_{fK}, \mathrm{d}]\omega \\
&= \iota_{fK}\mathrm{d}(f\omega) - (-1)^{k-1}\mathrm{d}\iota_{fK}\omega \\
&= f\iota_K\mathrm{d}(f\omega) - (-1)^{k-1}\mathrm{d}(f\iota_K\omega) \\
&= f\iota_K\mathrm{d}(f\omega) - (-1)^{k-1}(\mathrm{d}f \wedge \iota_K\omega + f\mathrm{d}\iota_K\omega) \\
&= f[\iota_K\mathrm{d}(f\omega) - (-1)^{k-1} + \mathrm{d}\iota_K\omega] + (-1)^k\mathrm{d}f \wedge \iota_K\omega \\
&= f\mathcal{L}_K\omega + (-1)^k\mathrm{d}f \wedge \iota_K\omega\,. \quad \blacksquare
\end{aligned} \tag{17.5.15}$$

17.6 Frölicher-Nijenhuis bracket

We have seen in section 17.4 that the insertion operator allows to identify the graded algebra of algebraic graded derivations with the vector-valued differential forms, where the algebra structure on the latter is given by the Nijenhuis-Richardson bracket. In section 17.5 we have seen that also the Nijenhuis-Lie derivative establishes an isomorphism between vector spaces, namely between graded derivations which commute with the exterior derivative and, again, vector-valued differential forms. We now promote also this vector space isomorphism to an isomorphism of graded Lie algebras. We thus need to define a graded Lie bracket on the space $\Omega^\bullet(M, TM)$, which is obtained by carrying the graded commutator over using the Nijenhuis-Lie derivative, in analogy to definition 17.4.1. Hence, we define this bracket as follows.

Definition 17.6.1 (Frölicher-Nijenhuis bracket)**.** Let M be a manifold and $K \in \Omega^k(M, TM)$, $L \in \Omega^l(M, TM)$ vector-valued differential forms on M. Their *Frölicher-Nijenhuis bracket* $[\![K, L]\!] \in \Omega^{k+l}(M, TM)$ is defined such that

$$\mathcal{L}_{[\![K,L]\!]} = [\mathcal{L}_K, \mathcal{L}_L]\,. \tag{17.6.1}$$

To see that this is well-defined, we must check that such an element $[\![K, L]\!]$ exists and is unique. To see that this is the case, note that both $\mathcal{L}_K$ and $\mathcal{L}_L$ commute with d, as a consequence of theorem 17.5.4. From the graded Jacobi identity then follows

$$[\mathrm{d}, [\mathcal{L}_K, \mathcal{L}_L]] = [[\mathrm{d}, \mathcal{L}_K], \mathcal{L}_L] + (-1)^{kl}[\mathcal{L}_K, [\mathrm{d}, \mathcal{L}_L]] = 0\,. \tag{17.6.2}$$

Hence, also their graded commutator commutes with d. Thus, again by theorem 17.5.4, a unique element $[\![K, L]\!]$ with the desired properties indeed exists, and we can conclude as follows.

Theorem 17.6.1. *The Frölicher-Nijenhuis bracket equips the space*

$$\Omega^\bullet(M, TM) = \bigoplus_{k=0}^{\dim M} \Omega^k(M, TM) \tag{17.6.3}$$

of vector-valued differential forms on a manifold M with the structure of a graded Lie algebra, such that:

1. *graded antisymmetry:*
$$[\![K_1, K_2]\!] = -(-1)^{k_1k_2}[\![K_2, K_1]\!]\,, \tag{17.6.4}$$
2. *graded Jacobi identity:*
$$[\![K_1, [\![K_2, K_3]\!]]\!] = [\![[\![K_1, K_2]\!], K_3]\!] + (-1)^{k_1k_2}[\![K_2, [\![K_1, K_3]\!]]\!]\,, \tag{17.6.5}$$

where $K_1 \in \Omega^{k_1}(M,TM), K_2 \in \Omega^{k_2}(M,TM), K_3 \in \Omega^{k_3}(M,TM)$.

Proof. This follows directly from the definition 17.6.1 of the Frölicher-Nijenhuis bracket, from the fact that $\mathcal{L}$ is a vector space isomorphism and that $\mathcal{L}_{K_i}$ is a graded derivation of degree k_i. ■

So far, the Frölicher-Nijenhuis bracket appears similar to the Nijenhuis-Richardson bracket. They differ, however, by the fact that the Frölicher-Nijenhuis bracket is not algebraic, in contrast to the Nijenhuis-Richardson bracket. This can be seen as follows.

Theorem 17.6.2. *For all* $K \in \Omega^k(M,TM)$, $L \in \Omega^l(M,TM)$ *and* $f \in C^\infty(M)$ *holds*

$$\blacktriangleright \dots \blacktriangleleft \tag{17.6.6}$$

Proof. With the help of theorem 17.5.6, we apply the commutator to a function $g \in C^\infty(M)$ and calculate

$$\begin{aligned} [\mathcal{L}_{fK}, \mathcal{L}_L]g &= \mathcal{L}_{fK}\mathcal{L}_L g - (-1)^{kl}\mathcal{L}_L\mathcal{L}_{fK}g \\ &= f\mathcal{L}_K\mathcal{L}_L g + (-1)^k \mathrm{d}f \wedge \iota_K \mathcal{L}_L g - (-1)^{kl}\mathcal{L}_L[f\mathcal{L}_K g + (-1)^l \mathrm{d}f \wedge \iota_L g] \\ &= f\mathcal{L}_K\mathcal{L}_L g + (-1)^k \mathrm{d}f \wedge \iota_K \mathcal{L}_L g - (-1)^{kl}(\mathcal{L}_L f \mathcal{L}_K g + f\mathcal{L}_L\mathcal{L}_K g) \\ &= f[\mathcal{L}_K, \mathcal{L}_L]g + (-1)^k \mathrm{d}f \wedge \iota_K \mathcal{L}_L g - (-1)^{kl}\mathcal{L}_L f \mathcal{L}_K g \end{aligned} \tag{17.6.7}$$

■

As in the case of the Nijenhuis-Richardson bracket, we provide an explicit formula also for the Frölicher-Nijenhuis bracket. Here we make use of the fact that it is bilinear and that we can write any vector-valued differential form as a linear combination of tensor products $\omega \otimes X$, where ω is a differential form and X is a vector field. For terms of this form we find the following expression for the Frölicher-Nijenhuis bracket.

Theorem 17.6.3. *Let* $\omega \in \Omega^k(M), \sigma \in \Omega^l(M)$ *be differential forms and* $X, Y \in \mathrm{Vect}(M)$ *be vector fields. The Frölicher-Nijenhuis bracket of their tensor products is given by*

$$\begin{aligned} &[\![\omega \otimes X, \sigma \otimes Y]\!] = \\ &\omega \wedge \sigma \otimes [X,Y] + \omega \wedge \mathcal{L}_X\sigma \otimes Y - \mathcal{L}_Y\omega \wedge \sigma \otimes X + (-1)^k(\mathrm{d}\omega \wedge \iota_X\sigma \otimes Y + \iota_Y\omega \wedge \mathrm{d}\sigma \otimes X)\,. \end{aligned} \tag{17.6.8}$$

Proof. For a function $f \in \Omega^0(M)$ we find

$$\begin{aligned} [\mathcal{L}_{\omega\otimes X}, \mathcal{L}_{\sigma\otimes Y}]f &= \mathcal{L}_{\omega\otimes X}\mathcal{L}_{\sigma\otimes Y}f - (-1)^{kl}\mathcal{L}_{\sigma\otimes Y}\mathcal{L}_{\omega\otimes X}f \\ &= \mathcal{L}_{\omega\otimes X}[Y(f)\sigma] - (-1)^{kl}\mathcal{L}_{\sigma\otimes Y}[X(f)\omega] \\ &= X(Y(f))\omega\wedge\sigma + Y(f)\left[\omega\wedge\iota_X\mathrm{d}\sigma - (-1)^{k-1}\mathrm{d}(\omega\wedge\iota_X\sigma)\right] \\ &\quad - (-1)^{kl}\left\{Y(X(f))\sigma\wedge\omega + X(f)\left[\sigma\wedge\iota_Y\mathrm{d}\omega - (-1)^{l-1}\mathrm{d}(\sigma\wedge\iota_Y\omega)\right]\right\} \\ &= [X,Y](f)\omega\wedge\sigma + Y(f)\left[\omega\wedge\iota_X\mathrm{d}\sigma - (-1)^{k-1}\mathrm{d}\omega\wedge\iota_X\sigma + \omega\wedge\mathrm{d}\iota_X\sigma\right] \\ &\quad - (-1)^{kl}X(f)\left[\sigma\wedge\iota_Y\mathrm{d}\omega - (-1)^{l-1}\mathrm{d}\sigma\wedge\iota_Y\omega + \sigma\wedge\mathrm{d}\iota_Y\omega\right] \\ &= [X,Y](f)\omega\wedge\sigma + Y(f)\left[\omega\wedge\mathcal{L}_X\sigma + (-1)^k\mathrm{d}\omega\wedge\iota_X\sigma\right] \\ &\quad - (-1)^{kl}X(f)\left[\sigma\wedge\mathcal{L}_Y\omega + (-1)^l\mathrm{d}\sigma\wedge\iota_Y\omega\right] \\ &= [X,Y](f)\omega\wedge\sigma + Y(f)\omega\wedge\mathcal{L}_X\sigma - X(f)\mathcal{L}_Y\omega\wedge\sigma \\ &\quad + (-1)^k\left[Y(f)\mathrm{d}\omega\wedge\iota_X\sigma + X(f)\iota_Y\omega\wedge\mathrm{d}\sigma\right]\,, \end{aligned} \tag{17.6.9}$$

from which the statement follows. ■

As a particularly simple example, we may calculate the Frölicher-Nijenhuis bracket of an arbitrary vector-valued differential form and the unit section as follows.

Theorem 17.6.4. *The Frölicher-Nijenhuis bracket of any vector-valued differential form $K \in \Omega^k(M, TM)$ with the unit section $\delta \in \Omega^1(M, TM)$ vanishes, $[\![\delta, K]\!] = 0$.*

Proof. By definition of the Frölicher-Nijenhuis bracket we have

$$\mathcal{L}_{[\![\delta,K]\!]} = [\mathcal{L}_\delta, \mathcal{L}_K] = [\mathrm{d}, \mathcal{L}_K] = 0\,, \tag{17.6.10}$$

using theorems 17.5.5 and 17.5.4. Hence, by the uniqueness of the Frölicher-Nijenhuis bracket, we have $[\![\delta, K]\!] = 0$. ∎

Another case of particular interest is given if $K, L \in \Omega^1(M, TM)$ are vector-valued differential one-forms. In this case the formula from theorem 17.6.3 allows us to derive the following result.

Theorem 17.6.5. *Let $K, L \in \Omega^1(M, TM)$ be vector-valued differential one-forms and $X, Y \in \mathrm{Vect}(M)$ be vector fields. Then the Frölicher-Nijenhuis bracket satisfies*

$$\begin{aligned}&[\![K, L]\!](X, Y) = \\ &[KX, LY] + [LX, KY] + (KL + LK)[X, Y] - K([LX, Y] + [X, LY]) - L([KX, Y] + [X, KY])\,.\end{aligned} \tag{17.6.11}$$

Proof. We will calculate the commutator

$$[\mathcal{L}_K, \mathcal{L}_L]f = \mathcal{L}_K\mathcal{L}_L f + \mathcal{L}_L\mathcal{L}_K f \tag{17.6.12}$$

for a function $f \in \Omega^0(M)$, which yields a two-form, and apply it to two vector fields $X, Y \in \mathrm{Vect}(M)$. For the first term we obtain

$$\begin{aligned}(\mathcal{L}_K\mathcal{L}_L f)(X, Y) &= (\mathcal{L}_K\iota_L\mathrm{d}f)(X, Y) \\ &= (\iota_K\mathrm{d}\iota_L\mathrm{d}f - \mathrm{d}\iota_K\iota_L\mathrm{d}f)(X, Y) \\ &= (\mathrm{d}\iota_L\mathrm{d}f)(KX, Y) - (\mathrm{d}\iota_L\mathrm{d}f)(KY, X) - (\mathrm{d}\iota_K\iota_L\mathrm{d}f)(X, Y) \\ &= (KX)(\langle LY, \mathrm{d}f\rangle) - Y(\langle LKX, \mathrm{d}f\rangle) - \langle L[KX, Y], \mathrm{d}f\rangle \\ &\quad - (KY)(\langle LX, \mathrm{d}f\rangle) + X(\langle LKY, \mathrm{d}f\rangle) + \langle L[KY, X], \mathrm{d}f\rangle \\ &\quad - X(\langle LKY, \mathrm{d}f\rangle) + Y(\langle LKX, \mathrm{d}f\rangle) + \langle LK[X, Y], \mathrm{d}f\rangle \\ &= (KX)((LY)(f)) - (KY)((LX)(f)) \\ &\quad - (L[KX, Y])(f) + (L[KY, X])(f) + (LK[X, Y])(f)\,.\end{aligned} \tag{17.6.13}$$

The second term is identical, with K and L exchanged. Combining them we obtain

$$\begin{aligned}\mathcal{L}_K\mathcal{L}_L f + \mathcal{L}_L\mathcal{L}_K f = {}& [KX, LY](f) + [LX, KY](f) \\ & - (L[KX, Y])(f) - (L[X, KY])(f) + (LK[X, Y])(f) \\ & - (K[LX, Y])(f) - (K[X, LY])(f) + (KL[X, Y])(f)\,.\end{aligned} \tag{17.6.14}$$

We see that the vector field acting on f is the one given in the theorem. ∎

In particular, for $K = L$ we have the following result.

Theorem 17.6.6. *Let $K \in \Omega^1(M, TM)$ be a vector-valued differential one-form and $X, Y \in \mathrm{Vect}(M)$ be vector fields. Then the Frölicher-Nijenhuis bracket satisfies*

$$\frac{1}{2}[\![K, K]\!](X, Y) = [KX, KY] + K^2[X, Y] - K([KX, Y] + [X, KY])\,. \tag{17.6.15}$$

Proof. This follows directly from theorem 17.6.5 for $K = L$. ∎

The expression we encountered here has several interesting applications, which we will encounter later. It thus deserves its own name, and we have the following definition.

Definition 17.6.2 (Nijenhuis tensor). Let $K \in \Omega^1(M, TM)$ be a vector-valued differential one-form. Its *Nijenhuis tensor* is the vector-valued differential two-form

$$N_K = \frac{1}{2}[\![K, K]\!] . \tag{17.6.16}$$

Another special case we remark is obtained if $K = Z \in \Omega^0(M, TM) = \mathrm{Vect}(M)$ is a vector field, while again $L \in \Omega^1(M, TM)$ as in the previous examples. In this case we find the following relation.

Theorem 17.6.7. *Let $L \in \Omega^1(M, TM)$ be a vector-valued differential one-form and $X, Z \in \mathrm{Vect}(M)$ be vector fields. Then the Frölicher-Nijenhuis bracket satisfies*

$$[\![Z, L]\!](X) = (\mathcal{L}_Z L)X = [Z, LX] - L[Z, X] . \tag{17.6.17}$$

Proof. By direct calculation for a function $f \in \Omega^0(M)$ we obtain

$$\begin{aligned}([\mathcal{L}_Z, \mathcal{L}_L]f)(X) &= (\mathcal{L}_Z\mathcal{L}_L f - \mathcal{L}_L\mathcal{L}_Z f)(X)\\ &= (\mathcal{L}_Z\iota_L \mathrm{d}f - \mathcal{L}_L(Z(f)))(X)\\ &= (\iota_Z\mathrm{d}\iota_L\mathrm{d}f + \mathrm{d}\iota_Z\iota_L\mathrm{d}f - \iota_L\mathrm{d}(Z(f)))(X)\\ &= (\mathrm{d}\iota_L\mathrm{d}f)(Z, X) + X((LZ)(f)) - (LX)(Z(f))\\ &= Z(\langle LX, \mathrm{d}f\rangle) - X(\langle LZ, \mathrm{d}f\rangle) - (\langle L[Z, X], \mathrm{d}f\rangle) + X((LZ)(f)) - (LX)(Z(f))\\ &= [Z, LX](f) - (L[Z, X])(f) ,\end{aligned} \tag{17.6.18}$$

which yields the desired result. ■

Finally, we discuss the case of two vector fields.

Theorem 17.6.8. *Let $X, Y \in \Omega^0(M, TM) = \mathrm{Vect}(M)$ be vector fields. Then the Frölicher-Nijenhuis bracket reduces to the Lie bracket,*

$$[\![X, Y]\!] = [X, Y] . \tag{17.6.19}$$

Proof. This can be obtained by direct calculation using the explicit formula in theorem 17.6.3 for setting $\omega = \sigma \in \Omega^0(M)$ to be the constant function $p \mapsto 1$, in which case only the first term is non-vanishing and gives the desired result. Alternatively, one may use the definition of the Frölicher-Nijenhuis bracket, the fact that for vector fields the Nijenhuis-Lie derivative reduces to the ordinary Lie derivative and the commutator is given by theorem 16.2.1. ■

17.7 Graded algebra of derivations

We have seen in section 17.3 that any algebraic graded derivation D on a manifold M can uniquely be identified with a vector-valued form K through the insertion operator as $D = \iota_K$. It turns out that this result may be generalized to arbitrary, i.e., also non-algebraic graded derivations. This can be seen as follows.

Theorem 17.7.1 (Frölicher-Nijenhuis). *Let $D \in \mathrm{Der}_k\,\Omega(M)$ be a derivation of degree k on M. Then there exist unique elements $K \in \Omega^k(M, TM)$ and $L \in \Omega^{k+1}(M, TM)$ such that*

$$D = \mathcal{L}_K + \iota_L . \tag{17.7.1}$$

Proof. For $D \in \mathrm{Der}_k\,\Omega(M)$, define $K \in \Omega^k(M, TM)$ by the relation (17.5.7). It was shown in the proof of theorem 17.5.4 that K is the unique vector-valued k-form such that for any function f holds $\mathcal{L}_K f = Df$. As a consequence, $D - \mathcal{L}_K$ acts trivially on functions, and so it is an algebraic derivation. Following theorem 17.3.2 it then follows that there exists a unique $L \in \Omega^{k+1}(M, TM)$ such that $D - \mathcal{L}_K = \iota_L$. ■

A particular consequence is that the algebra of graded derivations decomposes in the form

$$\mathrm{Der}_\bullet\,\Omega(M) = \mathrm{Der}^0_\bullet\,\Omega(M) \oplus \mathrm{Der}^{\mathrm{d}}_\bullet\,\Omega(M)\,, \tag{17.7.2}$$

or

$$\mathrm{Der}_n\,\Omega(M) = \mathrm{Der}^0_n\,\Omega(M) \oplus \mathrm{Der}^{\mathrm{d}}_n\,\Omega(M)\,, \tag{17.7.3}$$

if we restrict to graded derivations of degree n. Since any derivation can uniquely be expressed in terms of vector-valued differential forms K, L following theorem 17.7.1, one may ask whether there exists any formula to express these two vector-valued differential forms for, e.g., the graded commutator $[D, D']$ of two graded derivations, in terms of the vector-valued differential forms defining the individual derivations D, D'. In fact, we have already encountered two special cases of this formula. If both D and D' are algebraic, i.e., their Nijenhuis-Lie derivative part vanishes, our question is answered by the formula (17.4.1), which defines the Nijenhuis-Richardson bracket. Similarly, if both D and D' are given by pure Nijenhuis-Lie derivatives without an algebraic part, the defining relation (17.6.1) of the Frölicher-Nijenhuis bracket allows us to express their graded commutator again as a Nijenhuis-Lie derivative. Making use of the fact that the graded commutator is bilinear, we are thus only left with the mixed case, i.e., how to express the graded commutator of a Nijenhuis-Lie derivative and an algebraic graded derivation using the decomposition (17.7.1). We find the following answer to this question.

Theorem 17.7.2. *Let M be a manifold and $K \in \Omega^k(M, TM)$, $L \in \Omega^l(M, TM)$ vector-valued differential forms on M. Then the mixed commutator satisfies*

$$[\mathcal{L}_K, \iota_L] = \iota_{[\![K,L]\!]} - (-1)^{k(l-1)}\mathcal{L}_{\iota_L K}\,. \tag{17.7.4}$$

Proof. We will follow the procedure outlined in theorem 17.7.1. First, we apply the left hand side to a function $f \in \Omega^0(M)$. This yields

$$\begin{aligned}
[\mathcal{L}_K, \iota_L]f &= \mathcal{L}_K\iota_L f - (-1)^{k(l-1)}\iota_L\mathcal{L}_K f\\
&= -(-1)^{k(l-1)}\iota_L\iota_K\mathrm{d}f\\
&= -(-1)^{k(l-1)}\iota_{\iota_L K}\mathrm{d}f\\
&= -(-1)^{k(l-1)}\mathcal{L}_{\iota_L K} f\,,
\end{aligned} \tag{17.7.5}$$

which shows the correctness of the second term. We now subtract this term and apply the result to $\mathrm{d}f$, recalling that an algebraic derivation is uniquely determined by its action on differentials of functions. Then we find

$$\begin{aligned}
\left([\mathcal{L}_K, \iota_L] + (-1)^{k(l-1)}\mathcal{L}_{\iota_L K}\right)\mathrm{d}f &= \left(\mathcal{L}_K\iota_L - (-1)^{k(l-1)}\iota_L\mathcal{L}_K + (-1)^{k(l-1)}\mathcal{L}_{\iota_L K}\right)\mathrm{d}f\\
&= \mathcal{L}_K\iota_L\mathrm{d}f - (-1)^{kl}\iota_L\mathrm{d}\mathcal{L}_K f + (-1)^{kl+l-1}\mathrm{d}\mathcal{L}_{\iota_L K} f\\
&= \mathcal{L}_K\iota_L\mathrm{d}f - (-1)^{kl}\left(\iota_L\mathrm{d}\mathcal{L}_K f - (-1)^{l-1}\mathrm{d}\iota_L\mathcal{L}_K f\right)\\
&= \mathcal{L}_K\mathcal{L}_L f - (-1)^{kl}\mathcal{L}_L\mathcal{L}_K f\\
&= [\mathcal{L}_K, \mathcal{L}_L]f\\
&= \mathcal{L}_{[\![K,L]\!]}f\\
&= \iota_{[\![K,L]\!]}\mathrm{d}f\,,
\end{aligned} \tag{17.7.6}$$

which proves the correctness of the first term. ■

With these results in place, we have now revealed the full structure of the algebra graded derivations. We conclude as follows.

Theorem 17.7.3. *The graded derivations* $\mathrm{Der}_\bullet\,\Omega(M)$ *on a manifold* M *of dimension* $\dim M = m$ *decompose as*

$$\mathrm{Der}_\bullet\,\Omega(M) = \bigoplus_{k=-1}^{m-1} \mathrm{Der}^0_k\,\Omega(M) \oplus \bigoplus_{k=0}^{m} \mathrm{Der}^{\mathrm{d}}_k\,\Omega(M)\,, \tag{17.7.7}$$

where

$$\mathrm{Der}^0_k\,\Omega(M) \cong \Omega^{k+1}(M, TM)\,, \quad \mathrm{Der}^{\mathrm{d}}_k\,\Omega(M) \cong \Omega^{k}(M, TM)\,, \tag{17.7.8}$$

with the identification given by the insertion operator 17.3.2 and the Nijenhuis-Lie derivative 17.5.1, respectively, and their graded commutator satisfies

$$\left[\mathrm{Der}^0_\bullet\,\Omega(M), \mathrm{Der}^0_\bullet\,\Omega(M)\right] \subseteq \mathrm{Der}^0_\bullet\,\Omega(M)\,, \tag{17.7.9a}$$

$$\left[\mathrm{Der}^{\mathrm{d}}_\bullet\,\Omega(M), \mathrm{Der}^{\mathrm{d}}_\bullet\,\Omega(M)\right] \subseteq \mathrm{Der}^{\mathrm{d}}_\bullet\,\Omega(M)\,, \tag{17.7.9b}$$

$$\left[\mathrm{Der}_k\,\Omega(M), \mathrm{Der}_l\,\Omega(M)\right] \subseteq \mathrm{Der}_{k+l}\,\Omega(M) \tag{17.7.9c}$$

for all allowed values of k *and* l.

Proof. The identification (17.7.8) is established by theorems 17.3.2 and 17.5.4. The decomposition follows from theorem 17.7.1, together with theorem 17.1.3, the fact that the degree k of vector-valued differential forms on M satisfies $0 \leq k \leq m$, as well as the two previously mentioned theorems establishing their identification with graded derivations of a particular degree. Finally, the relations (17.7.9) follow from theorems 17.3.1, 17.5.3 and 17.2.1 . ■

Chapter 18

Multivector fields

18.1 Schouten-Nijenhuis bracket

Chapter 19

Natural bundles over fiber bundles

19.1 Natural bundles over product manifolds

The most simple example of a fiber bundle, which we introduced as a motivating example in section 2.1, is that of a trivial fiber bundle, so that its total space is that of the direct product $M \times N$ of two manifolds M, N. In this case there are two projection maps $\mathrm{pr}_1 : M \times N \to M$ and $\mathrm{pr}_2 : M \times N \to N$, which allow us to identify certain bundles over the total space with bundles constructed from those over the individual factor manifolds. To study this relationship, we first consider a single tangent space.

Theorem 19.1.1. *Let M, N be manifolds and $M \times N$ their direct product. For each $p \in M$ and $q \in N$, there exists a canonical vector space isomorphism $T_{(p,q)}(M \times N) \cong T_pM \oplus T_qN$.*

Proof. Let $w \in T_{(p,q)}(M \times N)$, and consider the differentials $\mathrm{pr}_{1*} : T(M \times N) \to TM$ and $\mathrm{pr}_{2*} : T(M \times N) \to TN$ of the projections $\mathrm{pr}_1 : M \times N \to M$ and $\mathrm{pr}_2 : M \times N \to N$. From $\mathrm{pr}_1(p, q) = p$ and $\mathrm{pr}_2(p, q) = q$ follows $\mathrm{pr}_{1*}(w) \in T_pM$ and $\mathrm{pr}_{2*}(w) \in T_qN$, so that $(\mathrm{pr}_{1*}(w), \mathrm{pr}_{2*}(w)) \in T_pM \oplus T_qN$. This map is linear, since it is linear in each component, and therefore linear with respect to the vector space structure on the direct sum.

Conversely, let $u \in T_pM$ and $v \in T_qN$, so that $(u, v) \in T_pM \oplus T_qN$. Consider the constant maps $\mu_q : M \to M \times N, x \mapsto (x, q)$ and $\nu_p : N \to M \times N, y \mapsto (p, y)$. Then $\mu_q(p) = \nu_p(q) = (p, q)$, and thus $\mu_{q*}(u) + \nu_{p*}(v) \in T_{(p,q)}(M \times N)$. Also this map is linear by construction, since it is the sum of two linear maps on the two constituting subspaces T_pM and T_qN of $T_pM \oplus T_qN$.

Finally, we need to show that each linear map constructed above is the inverse of the other. First, for $u \in T_pM$ and $v \in T_qN$ we set $w = \mu_{q*}(u) + \nu_{p*}(v)$. We can use the linearity of the pushforward to show that

$$\begin{aligned}
\mathrm{pr}_{1*}(w) &= \mathrm{pr}_{1*}(\mu_{q*}(u) + \nu_{p*}(v)) \\
&= \mathrm{pr}_{1*}(\mu_{q*}(u)) + \mathrm{pr}_{1*}\,\nu_{p*}(v)) \\
&= (\mathrm{pr}_1 \circ \mu_q)_*(u) + (\mathrm{pr}_1 \circ \nu_p)_*(v) \\
&= \mathrm{id}_{M*}(u) + \phi_{p*}(v) \\
&= u\,,
\end{aligned} \tag{19.1.1}$$

where we denoted

$$\begin{array}{cccc} \phi_p & : & N & \to & M \\ & & q & \mapsto & p \end{array} \tag{19.1.2}$$

the constant function, so that $\phi_{p*}(v) = 0$ according to theorem 10.1.5. Analogously, one shows that $\mathrm{pr}_{2*}(w) = v$. It thus follows that for each pair $(u, v) \in T_pM \oplus T_qN$ there exists $w = \mu_{p*}(u) + \nu_{q*}(v)$ such that $(\mathrm{pr}_{1*}(w), \mathrm{pr}_{2*}(w)) = (u, v)$, and so $(\mathrm{pr}_{1*}, \mathrm{pr}_{2*})$ is surjective. Now using the fact that

$$\dim T_{(p,q)}(M \times N) = \dim M \times N = \dim M + \dim N = \dim T_pM + \dim T_qN = \dim T_pM \oplus T_qN \tag{19.1.3}$$

follows that it must also be injective, and thus bijective. Hence, we have constructed a linear bijection, and thus a vector space isomorphism, from $T_{(p,q)}(M \times N)$ to $T_pM \oplus T_qN$. ■

Having established the relation above, we can now conclude on the structure of the tangent bundle.

Theorem 19.1.2. *Let M, N be manifolds and $M \times N$ their direct product. There exists a canonical vector bundle isomorphism $T(M \times N) \cong \mathrm{pr}_1^* TM \oplus \mathrm{pr}_2^* TN$.*

Proof. We start by discussing the structure of the given pullback bundles. By definition, the elements of $\mathrm{pr}_1^* TM$ are the pairs $((p, q), u) \in (M \times N) \times TM$ such that

$$p = \mathrm{pr}_1(p, q) = \tau_M(u)\,, \tag{19.1.4}$$

where $\tau_M : TM \to M$ is the tangent bundle projection. ►...◄ ■

19.2 Vertical tangent bundle

Often one considers tangent vectors to the total space E of a fiber bundle (E, B, π, F). Since E is equipped with some additional structure in this case, the same holds also for its tangent bundle TE. In particular, we can find a particular type of tangent vectors, which we define as follows.

Definition 19.2.1 (Vertical tangent space)**.** Let (E, B, π, F) be a fiber bundle and $e \in E$. A derivation $v \in T_eE$ is called *vertical* if and only if $\pi_*(v) = 0$. The space $V_eE = \ker(\pi_*|_{T_eE})$ of all vertical tangent vectors at e is called the *vertical tangent space* over e.

It is clear from the definition that the vertical tangent space V_eE at $e \in E$ is a vector space, as it arises as the kernel of a linear function, and so naturally the question for the dimension of this space arises. This can easily be found as follows.

Theorem 19.2.1. *Let (E, B, π, F) be a fiber bundle and $e \in E$. The vertical tangent space V_eE is a vector subspace of the tangent space T_eE with dimension $\dim V_eE = \dim F$.*

Proof. Recall from theorem 10.5.5 that the projection $\pi : E \to B$ of a fiber bundle is a submersion. Hence, the restriction of its differential π_* to T_eE is a surjective linear function onto $T_{\pi(e)}B$. Since V_eE is the kernel of this map, it follows that the dimensions are related by

$$\begin{aligned} \dim V_eE &= \dim\ker(\pi_*|_{T_eE}) \\ &= \dim T_eE - \dim\mathrm{im}(\pi_*|_{T_eE}) \\ &= \dim T_eE - \dim T_{\pi(e)}B \\ &= \dim E - \dim B \\ &= \dim F. \quad ■ \end{aligned} \tag{19.2.1}$$

Given a vector space V_eE in every point $e \in E$ of a manifold E, one may attempt to construct a vector bundle. This will be defined as follows.

Definition 19.2.2 (Vertical tangent bundle)**.** Let (E, B, π, F) be a fiber bundle with $\dim F = n$. The *vertical tangent bundle* of E is the vector bundle $(VE, E, \nu, \mathbb{R}^n)$, whose total space is the disjoint union

$$VE = \biguplus_{e \in E} V_e E\,, \tag{19.2.2}$$

and the projection is the function $\nu : VE \to E$ such that $\nu(v) = e$ for $v \in V_e E$.

In order to define the structure of the vector bundle, one also needs to specify the local trivializations. In this case one can use the fact that TE, also being a vector bundle, is already equipped with local trivializations. Furthermore, since E is a total space of a fiber bundle, it is equipped with induced charts. Using these to define the trivializations of TE, one can obtain the following result, hence also defining the trivializations of VE.

Theorem 19.2.2. *The vertical tangent bundle VE of a fiber bundle (E, B, π, F) is a subbundle of rank $\dim F$ of the tangent bundle TE.*

Proof. Let $e \in E$. Then there exists an induced chart (W, ω) around e, which gives rise to adapted coordinates (x^a, y^μ) on $W \subset E$. ▶. . . ◀ ■

Having defined the vector bundle structure of VE and its relation with TE, one can consider sections of these bundles. As for TE, these deserve a particular name, which we define as follows.

Definition 19.2.3 (Vertical vector field)**.** Let (E, B, π, F) be a fiber bundle and VE its vertical tangent bundle. A section $X : E \to VE$ is called a *vertical vector field.*

In order to study the properties of vertical vector fields, it is helpful to find another characterization of vertical vectors. One easily shows the following relation:

Theorem 19.2.3. *Let (E, B, π, F) be a fiber bundle and $v \in TE$. Then v is vertical, $v \in VE$, if and only if $v(f \circ \pi) = 0$ for all $f \in C^\infty(B, \mathbb{R})$.*

Proof. Using the definition 10.1.1 of the differential, we have

$$v(f \circ \pi) = \pi_*(v)(f)\,. \tag{19.2.3}$$

If v is vertical, we have $\pi_*(v) = 0$, and so also $v(f \circ \pi) = 0$ for all $f \in C^\infty(B, \mathbb{R})$. Conversely, if the latter holds, then $\pi_*(v)(f)$ vanishes for all $f \in C^\infty(B, \mathbb{R})$. However, this implies that $\pi_*(v) = 0$, and so $v \in V_e E$. ■

One now easily sees that the action of a vertical vector field X on a function $f \circ \pi$ with $f \in C^\infty(B, \mathbb{R})$ vanishes, i.e., yields the zero function $0 \in C^\infty(E, \mathbb{R})$, since for each $e \in E$, $X(p) \in V_e E$ is vertical, and hence

$$(X(f \circ \pi))(e) = X(e)(f \circ \pi) = 0 \tag{19.2.4}$$

for all $e \in E$. This has an interesting consequence.

Theorem 19.2.4. *The commutator of vertical vector fields is vertical.*

Proof. Let $X, Y \in \Gamma(VE)$ be vertical vector fields. At every point $e \in E$ and for every function $f \in C^\infty(B, \mathbb{R})$ we find

$$[X, Y](e)(f \circ \pi) = X(e)Y(e)(f \circ \pi) - Y(e)X(e)(f \circ \pi) = 0\,, \tag{19.2.5}$$

and so $[X, Y]$ is vertical. ■

Theorem 19.2.5. *The vertical tangent bundle of a fiber bundle $\pi : E \to B$ defines a foliation on E, whose leaves are the fibers of $\pi : E \to B$.*

Proof. ►. . . ◄ ■

19.3 Horizontal cotangent bundle

After constructing a subbundle of the tangent bundle over the total space of a fiber bundle, we can now turn our attention to the cotangent bundle. Also in this case we can find a particular subbundle, which is defined by the bundle projection. Here we employ the following definition for its fibers.

Definition 19.3.1 (Horizontal cotangent space)**.** Let (E, B, π, F) be a fiber bundle and $e \in E$. A covector $\alpha \in T_e^*E$ is called *horizontal* if and only if $\langle v, \alpha\rangle = 0$ for all vertical vectors $v \in V_eE$. The space H_e^*E of all horizontal covectors at e is called the *horizontal cotangent space* over e.

Again it is clear that each space H^*eE is a vector space, and so we will determine its dimension as follows.

Theorem 19.3.1. *Let (E, B, π, F) be a fiber bundle and $e \in E$. The horizontal cotangent space H_e^*E is a vector subspace of the cotangent space T_e^*E with dimension $\dim H_e^*E = \dim B$.*

Proof. Keeping in mind that covectors $\alpha \in T_e^*E$ can be seen as linear functions $\alpha : T_eE \to \mathbb{R}$, there exists a function

$$\begin{array}{rccc} \rho & : & T_e^*E & \to & (V_eE)^* \\ & & \alpha & \mapsto & \alpha|_{V_eE} \end{array}\,. \tag{19.3.1}$$

This function ρ is obviously linear and surjective, and has $\ker\rho = H_e^*E$. It follows that

$$\dim H_e^*E = \dim\ker\rho = \dim T_e^*E - \dim(V_eE)^* = \dim E - \dim F = \dim B\,. \tag{19.3.2}$$

■

Now it is straightforward to define a fiber bundle as follows.

Definition 19.3.2 (Horizontal cotangent bundle)**.** Let (E, B, π, F) be a fiber bundle with $\dim B = n$. The *horizontal cotangent bundle* of E is the vector bundle $(H^*E, E, \bar{\nu}, \mathbb{R}^n)$, whose total space is the disjoint union

$$H^*E = \biguplus_{e \in E} H_e^*E\,, \tag{19.3.3}$$

and the projection is the function $\bar{\nu} : H^*E \to E$ such that $\bar{\nu}(\alpha) = e$ for $\alpha \in H_e^*E$.

As for the vertical tangent bundle, the local trivializations are defined by a suitable restriction of the local trivializations of the containing bundle, which is the cotangent bundle T^*E in this case. One may expect that also in this case one obtains a subbundle. This can be shown as follows.

Theorem 19.3.2. *The horizontal cotangent bundle H^*E of a fiber bundle (E, B, π, F) is a subbundle of rank $\dim B$ of the cotangent bundle T^*E.*

Proof. ►...◄ ■

Now we have a vector bundle $\bar{\nu} : H^*E \to E$, whose rank is given by $\dim B$. Note that one can obtain another vector bundle with the same base and dimension from the (co)tangent bundle TB and T^*B via pullback along π, and so one may ask whether these bundles are related. We now show that this is the case.

Theorem 19.3.3. *The horizontal cotangent bundle H^*E of a fiber bundle (E, B, π, F) is canonically isomorphic to the pullback bundle π^*T^*B.*

Proof. ►...◄ ■

Note that this construction is possible for the horizontal cotangent bundle, but not for the vertical tangent bundle. For the latter, however, we will encounter a similar procedure in the case of vector bundles in section 19.8.

Definition 19.3.3 (Horizontal covector field)**.** Let (E, B, π, F) be a fiber bundle and H^*E its horizontal cotangent bundle. A section $\omega : E \to H^*E$ is called a *horizontal covector field*.

19.4 Horizontal differential forms

Definition 19.4.1 (Horizontal differential form)**.** Let (E, B, π, F) be a fiber bundle. A *horizontal differential form of rank k* (or *π-horizontal k-form*) on E is a section of the exterior power bundle $\Lambda^k H^*E$ for $k \in \mathbb{N}$.

19.5 Horizontal and vertical tensors

19.6 Bundles over fibered products

Theorem 19.6.1. *Let (E_1, B, π_1, F_1) and (E_2, B, π_2, F_2) be fiber bundles and $(E_1 \times_B E_2, B, \pi, F_1 \times F_2)$ their fibered product. For each $e_1 \in E_1$ and $e_2 \in E_2$ with $\pi_1(e_1) = \pi_2(e_2)$ there exists a canonical vector space isomorphism*

$$\begin{aligned} T_{(e_1,e_2)}(E_1 \times_B E_2) &\cong \ker(\pi_1 \circ \mathrm{pr}_1 - \pi_2 \circ \mathrm{pr}_2)|_{(e_1,e_2)} \\ &= \{w \in T_{(e_1,e_2)}(E_1 \times E_2) \mid \pi_{1*}(\mathrm{pr}_{1*}(w)) = \pi_{2*}(\mathrm{pr}_{2*}(w))\}\,. \end{aligned} \tag{19.6.1}$$

Proof. ►...◄ ■

Theorem 19.6.2. *Let (E_1, B, π_1, F_1) and (E_2, B, π_2, F_2) be fiber bundles and $(E_1 \times_B E_2, B, \pi, F_1 \times F_2)$ their fibered product. There exists a canonical vector bundle isomorphism*

$$T(E_1 \times_B E_2) \cong \ker \iota^*(\pi_1 \circ \mathrm{pr}_1 - \pi_2 \circ \mathrm{pr}_2) = \{w \in \iota^* T(E_1 \times E_2) \mid \pi_{1*}(\mathrm{pr}_{1*}(w)) = \pi_{2*}(\mathrm{pr}_{2*}(w))\}, \tag{19.6.2}$$

where $\iota : E_1 \times_B E_2 \hookrightarrow E_1 \times E_2$ is the canonical inclusion.

Proof. ►...◄ ■

Theorem 19.6.3. *Let (E_1, B, π_1, F_1) and (E_2, B, π_2, F_2) be fiber bundles and $(E_1 \times_B E_2, B, \pi, F_1 \times F_2)$ their fibered product. For each $e_1 \in E_1$ and $e_2 \in E_2)$ with $\pi_1(e_1) = \pi_2(e_2)$ there exists a canonical vector space isomorphism $V_{(e_1,e_2)}(E_1 \times_B E_2) \cong V_{e_1} E_1 \oplus V_{e_2} E_2$.*

Proof. ►...◄ ■

Theorem 19.6.4. *Let (E_1, B, π_1, F_1) and (E_2, B, π_2, F_2) be fiber bundles and $(E_1 \times_B E_2, B, \pi, F_1 \times F_2)$ their fibered product. There exists a canonical vector bundle isomorphism $V(E_1 \times_B E_2) \cong \mathrm{pr}_1^* VE_1 \oplus \mathrm{pr}_2^* VE_2$.*

Proof. ►...◄ ■

19.7 Bundles over pullback bundles

Theorem 19.7.1. *Let (E, B, π, F) be a fiber bundle, M a manifold and $\psi : M \to B$ a map, as well as $(\psi^* E, M, \psi^* \pi, F)$ the pullback bundle. For each $m \in M$ and $e \in E$ with $\psi(m) = \pi(e)$ there exists a canonical vector space isomorphism*

$$\begin{aligned} T_{(m,e)} \psi^* E &\cong \ker(\psi_* \circ \mathrm{pr}_{1*} - \pi_* \circ \mathrm{pr}_{2*})|_{(m,e)} \\ &= \{w \in T_{(m,e)}(M \times E) \mid \psi_*(\mathrm{pr}_{1*}(w)) = \pi_*(\mathrm{pr}_{2*}(w))\}. \end{aligned} \tag{19.7.1}$$

Proof. ►...◄ ■

Theorem 19.7.2. *Let (E, B, π, F) be a fiber bundle, M a manifold and $\psi : M \to B$ a map, as well as $(\psi^* E, M, \psi^* \pi, F)$ the pullback bundle. There exists a canonical vector bundle isomorphism*

$$T\psi^* E \cong \ker \iota^*(\psi_* \circ \mathrm{pr}_{1*} - \pi_* \circ \mathrm{pr}_{2*}) = \{w \in \iota^* T(M \times E) \mid \psi_*(\mathrm{pr}_{1*}(w)) = \pi_*(\mathrm{pr}_{2*}(w))\}, \tag{19.7.2}$$

where $\iota : \psi^* E \hookrightarrow M \times E$ is the canonical inclusion.

Proof. ►...◄ ■

Theorem 19.7.3. *Let (E, B, π, F) be a fiber bundle, M a manifold and $\psi : M \to B$ a map, as well as $(\psi^* E, M, \psi^* \pi, F)$ the pullback bundle. For each $m \in M$ and $e \in E$ with $\psi(m) = \pi(e)$ there exists a canonical vector space isomorphism $V_{(m,e)} \psi^* E \cong V_e E$.*

Proof. ►...◄ ■

Theorem 19.7.4. *Let (E, B, π, F) be a fiber bundle, M a manifold and $\psi : M \to B$ a map, as well as $(\psi^* E, M, \psi^* \pi, F)$ the pullback bundle. There exists a canonical vector bundle isomorphism $V\psi^* E \cong \mathrm{pr}_2^* VE$.*

Proof. ►...◄ ■

19.8 Bundles over vector bundles

So far we have not assumed that the fiber bundles (E, B, π, F) over which we constructed vertical and horizontal bundles have any particular structure. Now we come to the particular case of vector bundles, so that each fiber carries the structure of a vector space, which is preserved under local trivializations. In this case, we find the following property of the vertical tangent bundle.

Theorem 19.8.1. *The vertical tangent bundle VE of a vector bundle $(E, B, \pi, \mathbb{R}^k)$ is canonically isomorphic to the pullback bundle π^*E.*

Proof. First, note that VE and π^*E have the same base manifold E. We will construct a vector bundle isomorphism between these bundles covering the identity id_E on E by showing that for each $e \in E$ there is a canonical vector space isomorphism between the fibers V_eE and $(\pi^*E)_e$. Denote $b = \pi(e)$, and let $\tilde{e} \in E_b$. Then consider the curve

$$\begin{array}{cccc} \gamma & : & \mathbb{R} & \to & E \\ & & t & \mapsto & e + t\tilde{e} \end{array}, \tag{19.8.1}$$

which is well-defined since both e and $\tilde{e}$ lie in the same vector space E_b. Clearly, this curve is vertical, since $\pi(\gamma(t)) = b$ for all $t \in \mathbb{R}$, and $\gamma(0) = e$. Hence, also its tangent vector $v = \dot{\gamma}(0) \in V_eE$ is vertical. Doing this for all $\tilde{e} \in E_b$ defines a map $(\pi^*E)_e \cong E_b \to V_eE$, where the former two fiber spaces are isomorphic due to the construction of the pullback bundle. ►Show isomorphism.◄ ■

19.9 Homogeneity and the Liouville vector field

In the following, we will frequently encounter objects, such as functions and tensor fields, which are *homogeneous* in the sense that they are invariant under a particular group action, up to a constant factor. These objects will be defined on the total space of a vector bundle, on which also the mentioned group action is (canonically) defined as follows.

Definition 19.9.1 (Dilatation)**.** Let $(E, B, \pi, \mathbb{R}^k)$ be a vector bundle. The one-parameter diffeomorphism group χ of *dilatations* is the Lie group action of $(\mathbb{R}, +)$ on E defined by

$$\begin{array}{cccc} \chi & : & \mathbb{R} \times E & \to & E \\ & & (\lambda, v) & \mapsto & \chi_\lambda(v) = e^\lambda v \end{array}. \tag{19.9.1}$$

Note that this definition covers only dilatations by a positive factor, which is usually sufficient. One may also consider reflections, by including the action of the group $\mathbb{Z}_2 = \{1, -1\}$, where the action is defined by $(g, v) \mapsto gv$. However, this will not be necessary for our purposes, unless explicitly noted. In particular, in will not be relevant for the following definition.

Definition 19.9.2 (Liouville vector field)**.** Let $(E, B, \pi, \mathbb{R}^k)$ be a vector bundle. The *Liouville vector field* is the generating vector field $\mathbf{c} \in \mathrm{Vect}(E)$ of the dilatations χ.

In other words, the Liouville vector field is the unique vector field which has the dilatations as its flow, as detailed in section 16.1. To further illustrate the Liouville vector field, it is useful to derive its coordinate expression. For this purpose, let $(\epsilon_\mu, \mu = 1, \ldots, k)$ be a basis of the fiber E_p for $p \in B$, so that an element of the fiber can be expressed as $y = y^\mu \epsilon_\mu$. The dilatations can then be expressed in the form $\chi : (\lambda, y^\mu \epsilon_\mu) \mapsto e^\lambda y^\mu \epsilon_\mu$. To derive an expression for the Liouville vector field, we consider this mapping as a curve with parameter λ, which passes through a fixed point y for $\lambda = 0$. From this we obtain

$$\mathbf{c}(y) = \frac{\partial}{\partial \lambda}\left(e^\lambda y^\mu\right)\bigg|_{\lambda=0} \frac{\partial}{\partial y^\mu} = y^\mu \frac{\partial}{\partial y^\mu}\,. \tag{19.9.2}$$

Note that this coordinate expression is independent of the choice of the basis (ϵ_μ).

The dilatations, and hence also the Liouville vector field, find application in the definition of homogeneous structures on vector bundles. The most common case is that of homogeneous tensors, which we define as follows.

Definition 19.9.3 (Homogeneity)**.** Let $(E, B, \pi, \mathbb{R}^k)$ be a vector bundle. A tensor field (or tensor density) Q on E is called (positively) *homogeneous of order* r if and only if

$$\chi_\lambda^* Q = e^{r\lambda} Q \tag{19.9.3}$$

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

Note that some authors define a different order of homogeneity. In [MA94, BM07], the order of homogeneity for vector fields and contravariant tensor fields differs from the one introduced here by one, while for functions and differential forms no such difference is introduced.

Again we consider only positive dilatations, and so the precise notion we defined is that of *positive* homogeneity. Another important aspect in this definition is the fact that $\chi_\lambda : E \to E$ is a diffeomorphism, and so we can take the pullback of an arbitrary tensor field, following definition 12.1.2. Further, the fact that χ is the flow of the Liouville vector field allows to relate this definition to that of the Lie derivative as shown in section 16.2. One obtains the following relation.

Theorem 19.9.1. *A tensor field Q on the total space E of a vector bundle is homogeneous of order r if and only if $\mathcal{L}_{\mathbf{c}} Q = rQ$.*

Proof. Let Q be a r-homogeneous tensor field according to definition 19.9.3. By direct calculation following definition 16.2.1 of the Lie derivative one finds that

$$\begin{aligned}\mathcal{L}_{\mathbf{c}} Q &= \lim_{\lambda \to 0} \frac{\chi_\lambda^* Q - Q}{\lambda} \\ &= \lim_{\lambda \to 0} \frac{e^{r\lambda} Q - Q}{\lambda} \\ &= \frac{\mathrm{d}}{\mathrm{d}\lambda} e^{r\lambda} Q\bigg|_{\lambda=0} \\ &= rQ\,,\end{aligned} \tag{19.9.4}$$

where the step from the first to the second line follows from the assumption that Q is homogeneous of order r.

Conversely, let Q be a tensor field of rank (k, l) on E such that $\mathcal{L}_{\mathbf{c}} Q = rQ$. For $y \in E$, consider the two functions

$$\begin{array}{cccc}\phi_y & : & \mathbb{R} & \to & T^k_{l\,y} E \\ & & \lambda & \mapsto & e^{r\lambda} Q(y)\end{array} \tag{19.9.5}$$

and

$$\begin{array}{cccc}\psi_y & : & \mathbb{R} & \to & T^k_{l\,y} E \\ & & \lambda & \mapsto & (\chi_\lambda^* Q)(y)\end{array}\,. \tag{19.9.6}$$

Setting $\lambda = 0$, we find

$$\phi_y(0) = Q(y) = \psi_y(0)\,. \tag{19.9.7}$$

Further, for their first derivatives we find

$$\begin{aligned}
\psi_y'(\lambda) &= \frac{\mathrm{d}}{\mathrm{d}\lambda}(\chi_\lambda^* Q)(y) \\
&= \lim_{\tilde\lambda\to 0} \frac{(\chi_{\lambda+\tilde\lambda}^* Q)(y) - (\chi_\lambda^* Q)(y)}{\tilde\lambda} \\
&= \chi_\lambda^* \left(\lim_{\tilde\lambda\to 0} \frac{\chi_{\tilde\lambda}^* Q - Q}{\tilde\lambda} \right)(y) \\
&= (\chi_\lambda^* \mathcal{L}_{\mathbf{c}} Q)(y) \\
&= r(\chi_\lambda^* Q)(y) \\
&= r\psi_y(\lambda)\,,
\end{aligned} \tag{19.9.8}$$

while also

$$\phi_y'(\lambda) = \frac{\mathrm{d}}{\mathrm{d}\lambda} e^{r\lambda} Q(y) = r e^{r\lambda} Q(y) = r\phi_y(\lambda)\,. \tag{19.9.9}$$

Hence, ϕ_y and ψ_y are both solutions to the same first order differential equation with the same initial value, and so they must be identical. Hence,

$$(\chi_\lambda^* Q)(y) = \psi_y(\lambda) = \phi_y(\lambda) = e^{r\lambda} Q(y) \tag{19.9.10}$$

for all $y \in E$ and $\lambda \in \mathbb{R}$, so that Q is r-homogeneous. ■

Recalling the Leibniz rule (16.2.14) from theorem 16.2.1, one now readily concludes that homogeneity must be an additive property. This can be stated as follows.

Theorem 19.9.2. *Let Q, Q' be homogeneous tensor fields of order r, r', respectively. Then their tensor product is homogeneous of order $r + r'$.*

Proof. From the Leibniz rule (16.2.14) follows

$$\begin{aligned}
\mathcal{L}_{\mathbf{c}}(Q \otimes Q') &= (\mathcal{L}_{\mathbf{c}} Q) \otimes Q' + Q \otimes (\mathcal{L}_{\mathbf{c}} Q') \\
&= (rQ) \otimes Q' + Q \otimes (r'Q') \\
&= (r + r')(Q \otimes Q')\,. \quad ■
\end{aligned} \tag{19.9.11}$$

A number of similar relations can be derived in the same fashion, and we list them here for reference. We start with the following rule for differential forms, which is derived in full analogy to the previous one.

Theorem 19.9.3. *Let $\omega \in \Omega^k(E), \omega' \in \Omega^{k'}(E)$ be homogeneous differential forms of order r, r', respectively. Then their exterior product is homogeneous of order $r + r'$.*

Proof. From the Leibniz rule (16.5.7) follows

$$\begin{aligned}
\mathcal{L}_{\mathbf{c}}(\omega \wedge \omega') &= (\mathcal{L}_{\mathbf{c}} \omega) \wedge \omega' + \omega \wedge (\mathcal{L}_{\mathbf{c}} \omega') \\
&= (r\omega) \wedge \omega' + \omega \wedge (r'\omega') \\
&= (r + r')(\omega \wedge \omega')\,. \quad ■
\end{aligned} \tag{19.9.12}$$

A similar rule can be derived for the exterior derivative.

Theorem 19.9.4. *Let $\omega \in \Omega^k(E)$ be a homogeneous differential form of order r. Then also the exterior derivative $\mathrm{d}\omega$ is homogeneous of order r.*

Proof. From the relation (16.5.6) between the Lie derivative and exterior derivative follows

$$\mathcal{L}_{\mathbf{c}}\mathrm{d}\omega = \mathrm{d}\mathcal{L}_{\mathbf{c}}\omega = r\,\mathrm{d}\omega\,. \tag{19.9.13}$$

∎

Next, we come to vector fields, where we can formulate the following rule.

Theorem 19.9.5. *Let $X, Y \in \mathrm{Vect}(E)$ be homogeneous vector fields of order r, r', respectively. Then their commutator is homogeneous of order $r + r'$.*

Proof. From the Jacobi identity (16.4.4) follows

$$\begin{aligned}\mathcal{L}_{\mathbf{c}}[X,Y] &= [\mathcal{L}_{\mathbf{c}}X, Y] + [X, \mathcal{L}_{\mathbf{c}}Y]\\ &= [rX, Y] + [X, r'Y]\\ &= (r + r')[X, Y]\,. \quad ∎\end{aligned} \tag{19.9.14}$$

The relation proven above can be seen as a special case of the following rule, which generalizes the vector field Y to an arbitrary tensor field.

Theorem 19.9.6. *Let $X \in \mathrm{Vect}(E)$ be a homogeneous vector field of order r and Q a homogeneous tensor field of order r' on E. Then the Lie derivative $\mathcal{L}_XQ$ is homogeneous of order $r + r'$.*

Proof. From the commutator rule (16.2.20) of Lie derivatives follows

$$\begin{aligned}\mathcal{L}_{\mathbf{c}}\mathcal{L}_XQ &= \mathcal{L}_{[\mathbf{c},X]}Q + \mathcal{L}_X\mathcal{L}_{\mathbf{c}}Q\\ &= \mathcal{L}_{rX}Q + \mathcal{L}_X(r'Q)\\ &= r\mathcal{L}_XQ + r'\mathcal{L}_XQ\\ &= (r + r')\mathcal{L}_XQ + r'\mathcal{L}_XQ\,. \quad ∎\end{aligned} \tag{19.9.15}$$

Finally, relating vector fields and differential forms we obtain the following relation.

Theorem 19.9.7. *Let $\omega \in \Omega^k(E)$ be a homogeneous differential form of order r and $X \in \mathrm{Vect}(E)$ a homogeneous vector field of order r'. Then the interior product $\iota_X\omega$ is homogeneous of order $r + r'$.*

Proof. From the relation (16.5.8) between the Lie derivative and interior product follows

$$\begin{aligned}\mathcal{L}_{\mathbf{c}}\iota_X\omega &= \iota_{[\mathbf{c},X]}\omega + \iota_X\mathcal{L}_{\mathbf{c}}\omega\\ &= \iota_{r'X}\omega + r\iota_X\omega\\ &= (r + r')\iota_X\omega\,. \quad ∎\end{aligned} \tag{19.9.16}$$

The relation shown in theorem 19.9.1 has a number of interesting special cases. The most simple one is obtained in case of a tensor field of rank $(0, 0)$, i.e., a real function $f \in C^\infty(E, \mathbb{R})$. In this case the theorem reduces to the following.

Theorem 19.9.8 (Euler's homogeneous function theorem)**.** *A function $f \in C^\infty(E, \mathbb{R})$ on the total space E of a vector bundle is homogeneous of order r if and only if $\mathbf{c}f = rf$.*

Proof. This is a direct consequence and special case of theorem 19.9.1 for a tensor field of rank $(0, 0)$. ∎

Another case which can easily be derived and understood is the following.

Theorem 19.9.9. *The Liouville vector field $\mathbf{c}$ is homogeneous of order 0.*

Proof. This follows from the Lie derivative for vector fields, which yields

$$\mathcal{L}_{\mathbf{c}}\mathbf{c} = [\mathbf{c}, \mathbf{c}] = 0\,. \tag{19.9.17}$$

■

It is instructive to recall the coordinate expression (19.9.2) for the Liouville vector field. While its components y^μ in the given coordinates are obviously homogeneous of degree 1, since they scale linearly with a dilatation, it follows using theorem 19.9.2 that the coordinate vector fields $\partial/\partial y^\mu$ must be homogeneous of order -1, which can also be shown by direct calculation. Finally, we come to another simple example.

Theorem 19.9.10. *Let $(E, B, \pi, \mathbb{R}^k)$ be a vector bundle and $\omega \in \Omega^p(B)$ a p-form on the base manifold B. Then $\pi^*\omega \in \Omega^p(E)$ is a 0-homogeneous p-form on the total space E.*

Proof. The Liouville vector field $\mathbf{c}$ is a vertical vector field on E, and so its flow χ preserves the fibers. Hence, $\pi \circ \chi_\lambda = \pi$ for all $\lambda \in \mathbb{R}$. From the chain rule for the pullback then follows

$$\chi_\lambda^*\pi^*\omega = (\pi \circ \chi_\lambda)^*\omega = \pi^*\omega = e^{0\lambda}\pi^*\omega\,. \tag{19.9.18}$$

■

The last theorem is an example for a more general class of objects, which are defined on the base manifold B alone, and then lifted to a 0-homogeneous object the total space E. We will encounter further such objects in later sections.

Chapter 20

Bundles with structure groups

20.1 Principal fiber bundles

We will now come to a class of fiber bundles which are of particular importance. They have the property that they carry the action of a Lie group, which is compatible with the fiber bundle structure, i.e., preserves the fibers. In addition, each fiber is diffeomorphic to the acting Lie group. We define this class of fiber bundles as follows.

Definition 20.1.1 (Principal fiber bundle)**.** Let G be a Lie group. A *principal G-bundle* (or *principal fiber bundle with structure group G*) is a fiber bundle $\pi : P \to M$ with a right Lie group action $\cdot : P \times G \to P$ which preserves the fibers and acts freely and transitively on them.

We clarify a few notions used in this definition. A group action is *fiber preserving* if for all $p \in P$ and $g \in G$ holds $\pi(p) = \pi(p \cdot g)$, i.e., p and $p \cdot g$ lie in the same fiber of P. Further, the action should be free and transitive on the fibers, which means that for each $p, p' \in P$ which lie in the same fiber, $\pi(p) = \pi(p')$, there exists a unique $g \in G$ such that $p' = p \cdot g$.

To get a better understanding of the geometry of principal bundles, it is helpful to discuss their local trivializations. As with any fiber bundle, they encode most of the geometric structure, and are a key ingredient in constructing adapted coordinates. For principal fiber bundles they have a particularly nice property, which we may state as follows.

Theorem 20.1.1. *Let $\pi : P \to M$ be a principal G-bundle and $U \subset M$ a trivializing subset of M. Then there exists a one-to-one correspondence between local trivializations $\phi : \pi^{-1}(U) \to U \times G$ and local sections $\sigma \in \Gamma_{|U}(P)$.*

Proof. We only sketch the proof and omit showing smoothness of the involved maps. Let $\sigma \in \Gamma_{|U}(P)$ be a local section. For $p \in \pi^{-1}(U)$ with $\pi(p) = x \in U$, define $\phi(p) = (x, g)$ such that $g \in G$ is the unique element such that $p = \sigma(x) \cdot g$. This defines a bijection between $\pi^{-1}(U)$ and $U \times G$. It can be shown to be a diffeomorphism by using the smoothness of σ and the group action on P.

Conversely, starting from a local trivialization $\phi : \pi^{-1}(U) \to U \times G$ we can define a local section $\sigma \in \Gamma_{|U}(P)$ by $\sigma(x) = \phi^{-1}(x, e)$, where $e \in G$ is the unit element. This is obviously a smooth section. One now easily checks that this construction and the construction from the first half of this proof are indeed inverses of each other, and thus establish a one-to-one correspondence. ∎

This theorem has an interesting and important consequence for the existence of global sections, which can be stated as follows.

Theorem 20.1.2. *A principal fiber bundle is trivial if and only if it admits a global section.*

Proof. Let $\pi : P \to M$ be a principal G-bundle with Lie group G. If P is a trivial fiber bundle, then there exists a diffeomorphism $\phi : M \times G \to P$ such that $\pi \circ \phi = \mathrm{pr}_M$. Then $\sigma : M \to P, x \mapsto \phi(x, e)$ is a global section, where $e \in G$ is the unit element.

Conversely, let $\sigma : M \to P$ be a global section of an arbitrary principal G-bundle $\pi : P \to M$. Then every $p \in P$ can uniquely be written in the form $p = \sigma(x) \cdot g$ for some $x \in M$ and $g \in G$. One easily checks that the map $\phi : M \times G \to P, (x, g) \mapsto \sigma(x) \cdot g$ is a diffeomorphism, so that P is a trivial fiber bundle. ■

An important example of a principal bundle is that of a *coset space* G/H, which we discuss in detail in section 41.1, following definition 15.4.1. Here we restrict ourselves to studying the fundamental vector fields, which we introduced in section 15.12, in the case that G and H are matrix groups. Then we can explicitly construct the fundamental vector fields as follows:

Example 20.1.1 (Fundamental vector fields on coset bundles of matrix groups). Let $G \subset M_{n,n}$ be a matrix group (for example, one of the classical groups from example 15.1.3) and $H \subset M_{n,n}$ a closed subgroup. Since both of them can be written in form of $n \times n$-matrices in $M_{n,n}$, we can see both G and H as submanifolds of $M_{n,n}$ and use the matrix components as coordinates, on which we place suitable restrictions. We denote these coordinates by $(g^a{}_b)$ on G and $(h^a{}_b)$ on H, where $a, b, = 1, \ldots, n$. In these coordinates the unit element $e \in H$ has the form $h^a{}_b = \delta^a{}_b$. An element $X \in \mathfrak{h} = \operatorname{Lie} H$ of the Lie algebra of H is a left invariant vector field on H, which is uniquely defined by its value at e. Using the coordinates on H, this value can be written as

$$X(e) = X^a{}_b \frac{\partial}{\partial h^a{}_b}, \tag{20.1.1}$$

where the matrix $X^a{}_b$ is given by the matrix representation of the Lie algebra element (see example 15.6.2 for the classical matrix Lie algebras). To calculate its fundamental vector field, we consider the right Lie group action $G \times H \ni (g, h) \mapsto g \cdot h = gh \in G$ of H on G. In coordinates it takes the form

$$(g^a{}_b, h^a{}_b) \mapsto (g^a{}_c h^c{}_b). \tag{20.1.2}$$

For a fixed element $g \in G$ the pushforward of $X(e) \in T_e H$ to $T_g G$ is then given by

$$\begin{aligned} (R^g)_*(X(e)) &= (R^g)_* \left(X^a{}_b \frac{\partial}{\partial h^a{}_b} \right) \\ &= X^a{}_b \frac{\partial}{\partial h^a{}_b} (g^c{}_e h^e{}_d) \bigg|_{h=e} \frac{\partial}{\partial g^c{}_d} \\ &= X^a{}_b g^c{}_e \delta^e_a \delta^b_d \frac{\partial}{\partial g^c{}_d} \\ &= g^a{}_c X^c{}_b \frac{\partial}{\partial g^a{}_b} \\ &= (gX)^a{}_b \frac{\partial}{\partial g^a{}_b}. \end{aligned} \tag{20.1.3}$$

This can nicely be generalized to other principal bundles, if the structure group is a matrix group.

Example 20.1.2 (Fundamental vector fields for matrix groups). Let $G \subset M_{n,n}$ be a matrix group and $\pi : P \to M$ a principal G-bundle. We use the same matrix component coordinates $(g^a{}_b)$ on G as in the previous example. On an open subset $U \subset M$ we can find an equivariant local trivialization $\phi : \pi^{-1}(U) \to U \times G$. Writing the induced coordinates on $\pi^{-1}(U)$ as $(x^\mu, p^a{}_b)$, where (x^μ) are coordinates on U and $(p^a{}_b)$ denotes the matrix components on G, we find that the right action of the structure group is given by

$$((x^\mu, p^a{}_b), g^a{}_b) \mapsto (x^\mu, p^a{}_c g^c{}_b)\,, \tag{20.1.4}$$

i.e., by right multiplication of the matrix component coordinates. For a fixed $p \in \pi^{-1}(U)$ and $X \in T_eG$ we then have the fundamental vector field

$$\begin{aligned}(R^p)_*(X(e)) &= (R^p)_*\left(X^a{}_b \frac{\partial}{\partial g^a{}_b}\right)\\ &= X^a{}_b \frac{\partial}{\partial g^a{}_b}(p^c{}_e g^e{}_d)\bigg|_{g=e} \frac{\partial}{\partial p^c{}_d}\\ &= X^a{}_b p^c{}_e \delta^e_a \delta^b_d \frac{\partial}{\partial p^c{}_d}\\ &= g^a{}_c X^c{}_b \frac{\partial}{\partial g^a{}_b}\\ &= (pX)^a{}_b \frac{\partial}{\partial p^a{}_b}\,,\end{aligned} \tag{20.1.5}$$

so that the coordinate expression essentially looks the same as in the case of coset bundles.

In the case of a principal bundle, the fundamental vector fields have an interesting property.

Theorem 20.1.3. *For each $p \in P$, the fundamental vector fields on the principal G-bundle $\pi : P \to M$ define a vector space isomorphism between the Lie algebra $\mathfrak{g}$ of G and the vertical tangent space V_pP.*

This means that at each point $p \in P$ we can canonically identify vertical vectors, i.e., vectors $\xi \in T_pP$ which satisfy $\pi_*(\xi) = 0$, with Lie algebra elements of $\mathfrak{g}$ and vice versa. Further, this allows us to relate the values of fundamental vector fields at different points on the same fiber, which we state as follows.

Theorem 20.1.4. *The fundamental vector fields satisfy*

$$\tilde{X}(p \cdot g) = R_{g*}\left(\widetilde{\mathrm{Ad}_g(X)}(p)\right)\,. \tag{20.1.6}$$

Proof. First note that

$$R^{p\cdot g}(h) = p \cdot gh = p \cdot \alpha_g(h) g = (R_g \circ R^p \circ \alpha_g)(h) \tag{20.1.7}$$

for all $p \in P$ and $g, h \in G$. By direct calculation we then find

$$\begin{aligned}\tilde{X}(p \cdot g) &= R^{p\cdot g}_*(X(e))\\ &= (R_g \circ R^p \circ \alpha_g)_*(X(e))\\ &= R_{g*}(R^p_*(\mathrm{Ad}_g(X(e))))\\ &= R_{g*}\left(\widetilde{\mathrm{Ad}_g(X)}(p)\right)\end{aligned} \tag{20.1.8}$$

for all $p \in P$, $g \in G$ and $X \in \mathfrak{g}$. ■

Theorem 20.1.5. *The fibered product $P_1 \times_M P_2$ of a principal G_1-bundle $\pi_1 : P_1 \to M$ and a principal G_2-bundle $\pi_2 : P_2 \to M$ is a principal $(G_1 \times G_2)$-bundle.*

Proof. ►...◄ ■

20.2 Principal bundle morphisms

There are different possibilities which allow us to relate principal bundles to each other. The most general is to consider two principal bundles with arbitrary, and hence potentially different structure groups. In this case we must also relate the different structure groups to each other. Here this relation is given by a Lie group homomorphism. We thus define the following notion.

Definition 20.2.1 (Principal bundle morphism)**.** Let $\pi : P \to M$ be a principal G-bundle and $\varpi : Q \to N$ a principal H-bundle. A *principal bundle morphism* is a bundle morphism $\phi : P \to Q$ together with a Lie group homomorphism $\theta : G \to H$ such that $\phi(p \cdot g) = \phi(p) \cdot \theta(g)$ for all $p \in P$ and $g \in G$.

Recall that the definition 2.7.1 of a bundle morphism also entails a map $\tilde{\phi} : M \to N$ relating the base manifolds, which is said to be covered by the bundle morphism. The additional requirement for a principal bundle morphism that there exists a Lie group homomorphism $\theta : G \to H$ intertwining the group actions on the fibers of P and Q can be expressed by enlarging the diagram (2.7.1) and demanding that the following diagram commutes:

$$\begin{array}{ccc} P \times G & \xrightarrow{\phi \times \theta} & Q \times H \\ \downarrow{\scriptstyle \cdot} & & \downarrow{\scriptstyle \cdot} \\ P & \xrightarrow{\phi} & Q \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \varpi} \\ M & \xrightarrow{\tilde{\phi}} & N \end{array} \tag{20.2.1}$$

A more pictorial visualization is shown in figure 20.1. Here a group element $g \in G$ acts on the preimage $p \in P$ to yield a new element $p \cdot g \in P$ within the same fiber. The map $\phi : P \to Q$, which maps p to $\phi(p)$, then maps $p \cdot g$ to $\phi(p) \cdot \theta(g)$.

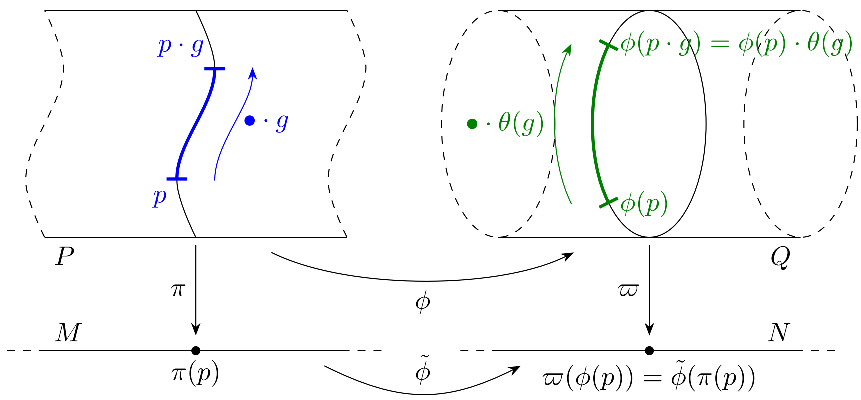


Figure 20.1: Visualization of a principal bundle morphism. If one translates the preimage by $p \mapsto p \cdot g$, the image follows and gets translated by $\phi(p) \mapsto \phi(p) \cdot \theta(g)$.

A special case is found if both bundles have the same structure group. In this case we may additionally demand that the homomorphism between the structure groups is the identity. We thus define the following notion.

Definition 20.2.2 (Principal G-bundle morphism)**.** Let $\pi : P \to M$ and $\varpi : Q \to N$ be principal G-bundles. A *principal G-bundle morphism* is a bundle morphism $\phi : P \to Q$ which is equivariant, i.e., such that $\phi(p \cdot g) = \phi(p) \cdot g$ for all $p \in P$ and $g \in G$.

To see how the notion of equivariance arises, observe that the upper part of the commutative diagram (20.2.1), which represents the defining property of a principal bundle morphism, simply reduces to the diagram (15.5.3) representing an equivariant map, in case the group homomorphism θ is the identity. Going one step further, and demanding that the occurring maps are diffeomorphisms, we arrive at the following notion.

Definition 20.2.3 (Principal G-bundle isomorphism)**.** Let $\pi : P \to M$ and $\varpi : Q \to N$ be principal G-bundles. A *principal G-bundle isomorphism* is a bundle isomorphism $\phi : P \to Q$ which is equivariant, i.e., such that $\phi(p \cdot g) = \phi(p) \cdot g$ for all $p \in P$ and $g \in G$. If a principal G-bundle isomorphism exists, then the two bundles are called isomorphic.

20.3 Associated fiber bundles

In this section we introduce another notion, which is of particular importance in physics. The basic idea behind this construction is to replace the fiber of a principal bundle by another fiber which carries an action of the same Lie group. The resulting bundle, which we call *associated* to the original principal bundle, is defined as follows.

Definition 20.3.1 (Associated fiber bundle)**.** Let G be a Lie group and $\pi : P \to M$ a principal G-bundle. Further, let F be a manifold together with a left Lie group action $\rho : G \times F \to F$. Consider the right action on the Cartesian product $P \times F$ given by

$$(p, f) \cdot g = (p \cdot g, \rho(g^{-1}, f)) \,. \tag{20.3.1}$$

Let $P \times_\rho F$ be the set of orbits of this right action and denote the orbit of $(p, f) \in P \times F$ by $[p, f]$. Finally, let $\pi_\rho : P \times_\rho F \to M$ be the projection map given by $\pi_\rho([p, f]) = \pi(p)$. Then $P \times_\rho F$ is called the *fiber bundle associated* to P via the action ρ.

One can now easily check that for every $x \in M$ the fiber $\pi_\rho^{-1}(x) = P_x \times_\rho F$ is diffeomorphic to F. To see this, and to understand how these fibers are related to each other, we need to find local trivializations. This can be done most easily by using the fact that $\pi : P \to M$ is a (principal) fiber bundle, so that we can use its trivializations. Hence, let us begin with a local trivialization $\phi : \pi^{-1}(U) \to U \times G$ of a trivializing subset $U \subset M$. Following theorem 20.1.1 this is equivalent to a local section $\sigma \in \Gamma_{|U}(P)$, and so we will make use of this section. Let

$$[p, f] = \{(p \cdot g, \rho(g^{-1}, f)), g \in G\} \in \pi_\rho^{-1}(U) \subset P \times_\rho F \tag{20.3.2}$$

be an element of the associated fiber bundle, which is defined as the orbit of the action (20.3.1) of G on $P \times F$ which passes through $(p, f) \in \pi^{-1}(U) \times F$. Let $x = \pi(p)$. By the properties of the principal bundle, there exists a unique $g \in G$ such that $\sigma(x) = p \cdot g$. Using the fact that the orbit also contains the point

$$(p \cdot g, \rho(g^{-1}, f)) = (\sigma(x), \rho(g^{-1}, f)) \,, \tag{20.3.3}$$

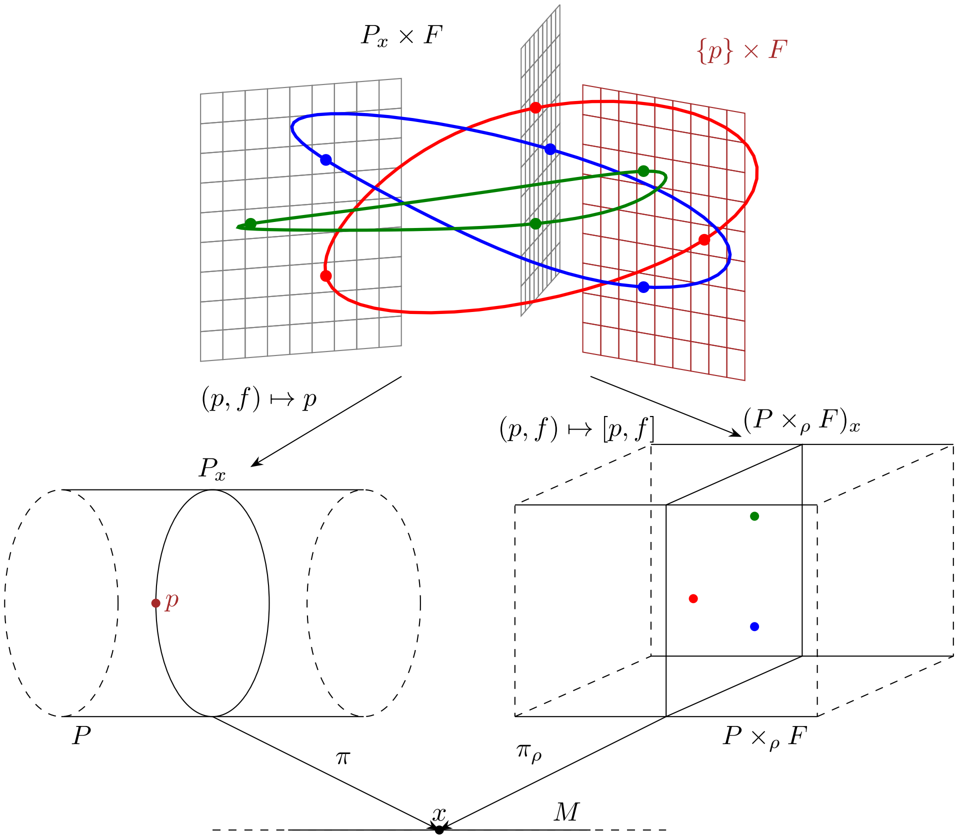


Figure 20.2: Concept of an associated fiber bundle. The elements of the fibers $(P \times_\rho F)_x$ are orbits of the action of the structure group on the product space $P_x \times F$.

and hence

$$[p, f] = [\sigma(x), \rho(g^{-1}, f)]\,, \tag{20.3.4}$$

we obtain a unique element $\rho(g^{-1}, f) \in F$. This element is independent of the original choice of the representative (p, f), and depends only on the orbit $[p, f]$, which intersects $\{\sigma(x)\} \times F$ in exactly one point by construction, as illustrated in 20.2. It is thus well-defined to define $\phi_\rho([p, f]) = (x, \rho(g^{-1}, f)) \in U \times F$. One can also explicitly construct the inverse as ϕ_ρ^{-1} : $U \times F \to \pi_\rho^{-1}(U), (x, f) \mapsto [\sigma(x), f]$. Thus, (U, ϕ_ρ) constitutes a local trivialization of $P \times_\rho F$.

The illustration in figure 20.2 shows that the fibers $P_x \times F$ of the bundle $\pi \circ \mathrm{pr}_1 : P \times F \to M$, whose quotient by the group action of G gives the associated bundle $P \times_\rho F$, are isomorphic to the product $P_x \times (P_x \times_\rho F)$. This suggests that the bundle $\pi \circ \mathrm{pr}_1 : P \times F \to M$ might be related to the fibered product $P \times_M (P \times_\rho F)$. We now show that this is indeed the case.

Theorem 20.3.1. *Let P be a principal G-bundle and $\rho : G \times F \to F$ a left action of G on a manifold F. Then the following hold:*

1. *The fibered product $\pi \times_M \pi_\rho : P \times_M (P \times_\rho F) \to M$ is canonically isomorphic to the bundle $\pi \circ \mathrm{pr}_1 : P \times F \to M$.*
2. $\mathrm{pr}_1 : P \times_M (P \times_\rho F) \to P$ *is trivial.*
3. $\mathrm{pr}_2 : P \times_M (P \times_\rho F) \to P \times_\rho F$ *is a principal G-bundle.*

Proof. Recall from section 2.8 that the elements of $P \times_M (P \times_\rho F)$ are pairs $(p, [p', f'])$ such that $\pi(p) = \pi_\rho([p', f']) = \pi(p') = x$ for some $x \in M$. Since both p and p' belong to the same fiber P_x, there exists a unique $g \in G$ such that $p' = p \cdot g$, and so we can choose a particular representative

$$[p', f'] = [p \cdot g, \rho(g^{-1}, f)] = [p, f]\,, \tag{20.3.5}$$

which uniquely determines the element $f \in F$. Hence, we can write each element $(p, [p', f']) \in P \times_M (P \times_\rho F)$ uniquely as $\vartheta(p, f)$, where we defined

$$\begin{array}{rcccc} \vartheta & : & P \times F & \to & P \times_M (P \times_\rho F) \\ & & (p, f) & \mapsto & (p, [p, f]) \end{array}. \tag{20.3.6}$$

Clearly, ϑ is bijective, and it is smooth since it is a composition of smooth maps given by the projection $(p, f) \mapsto p$ onto the first factor of $P \times F$ and the quotient $(p, f) \mapsto [p, f]$ of the action of G on $P \times F$. Finally, we have

$$(\pi \times_M \pi_\rho)(\vartheta(p, f)) = (\pi \times_M \pi_\rho)(p, [p, f]) = \pi(p) = \pi(\mathrm{pr}_1(p, f))\,, \tag{20.3.7}$$

showing that ϑ is a fiber bundle isomorphism, and it is canonically defined. This proves the first proposition, and we can make use of ϑ to prove the remaining propositions.

For the second proposition, we have

$$\mathrm{pr}_1(\vartheta(p, f)) = \mathrm{pr}_1(p, [p, f]) = p = \mathrm{pr}(p, f)\,, \tag{20.3.8}$$

and so ϑ is also a bundle isomorphism for the bundles $\mathrm{pr}_1 : P \times_M (P \times_\rho F)$ and $\mathrm{pr}_1 : P \times F$. The latter is a trivial bundle, and so also the former must be trivial.

Finally, for the third proposition, we have

$$\mathrm{pr}_2(\vartheta(p, f)) = \mathrm{pr}_2(p, [p, f]) = [p, f]\,, \tag{20.3.9}$$

and so ϑ is also a bundle morphism relating $\mathrm{pr}_2 : P \times_M (P \times_\rho F)$ to the projection $(p, f) \mapsto [p, f]$. The latter defines a principal G-bundle, and hence does also the former. ■

It is also helpful to see what happens if one restricts the local trivialization (U, ϕ_ρ) we constructed on the associated bundle to a single fiber $P_x \times_\rho F$ over $x \in U$. In this case one does not need to specify a local section of P, but only a single point $p \in P$. We can use this to explicitly construct a diffeomorphism as follows.

Definition 20.3.2 (Fiber diffeomorphism)**.** Let $\pi : P \to M$ be a principal G-bundle and $\pi_\rho : P \times_\rho F \to M$ an associated bundle with fiber F. For $p \in P$ the diffeomorphism $[p] : F \to P_{\pi(p)} \times_\rho F, f \mapsto [p, f]$ is called the *fiber diffeomorphism* of p.

One can see that this is related to the previously defined local trivialization ϕ_ρ by setting $p = \sigma(x)$, so that $g = e$ becomes the unit element. It is now easy to see the following property of the fiber diffeomorphisms.

Theorem 20.3.2. *For every $p \in P$, $g \in G$ and $f \in F$ the fiber diffeomorphism defined above satisfies*

$$[p \cdot g](f) = [p](\rho(g, f))\,. \tag{20.3.10}$$

Proof. By definition of the orbits $[p, f]$ we have

$$[p \cdot g](f) = [p \cdot g, f] = [p, \rho(g, f)] = [p](\rho(g, f))\,, \tag{20.3.11}$$

by direct calculation. ■

With the help of this property we can now understand the structure of the space $\Gamma(P \times_\rho F)$ of sections of an associated fiber bundle.

Theorem 20.3.3. *There is a one-to-one correspondence between sections $\sigma \in \Gamma(P \times_\rho F)$ of an associated fiber bundle $P \times_\rho F$ and G-equivariant maps $\phi \in C^\infty_G(P, F)$.*

Proof. Let $\sigma \in \Gamma(P \times_\rho F)$ be a section of $P \times_\rho F$. We then define a map ϕ by

$$\begin{array}{rcccc} \phi & : & P & \to & F \\ & & p & \mapsto & [p]^{-1}(\sigma(\pi(p))) \end{array} . \tag{20.3.12}$$

This map is well-defined, since $\sigma(\pi(p)) \in P_{\pi(p)} \times_\rho F$ and $[p] : F \to P_{\pi(p)} \times_\rho F$ is a diffeomorphism, and thus possesses an inverse. It is G-equivariant, since for all $g \in G$:

$$\phi(p \cdot g) = [p \cdot g]^{-1}(\sigma(\pi(p \cdot g))) = \rho(g^{-1}, [p]^{-1}(\sigma(\pi(p)))) = \rho(g^{-1}, \phi(p)) . \tag{20.3.13}$$

Conversely, let $\phi \in C^\infty_G(P, F)$ be an equivariant map. For $x \in M$ choose $p \in \pi^{-1}(x) \subset P$ and define $\sigma(x) = [p, \phi(p)]$. This definition is independent of the choice of p, since for any other $p' = p \cdot g$ we have

$$[p', \phi(p')] = [p \cdot g, \phi(p \cdot g)] = [p \cdot g, \rho(g^{-1}, \phi(p))] = [p, \phi(p)] . \tag{20.3.14}$$

It is easy to check that σ defines a section of $P \times_\rho F$. ■

20.4 Associated vector bundles

A particularly common case is given if the fiber space F is a vector space and the action $\rho : G \times F \to F$ is linear in its second argument, i.e., it is given by a linear representation of G on the vector space F. In this case one may ask whether the linear structure of F is preserved and carries over to the fibers of the constructed associated bundle. We show that this is indeed the case:

Theorem 20.4.1. *If F is a vector space of dimension k and ρ a linear representation of G, then the associated bundle $P \times_\rho F$ is a vector bundle of rank k.*

Proof. First, we need to check that there is a vector space structure on each fiber $P_x \times_\rho F$ with $x \in M$, which is inherited from the vector space structure on F. Consider two elements $[p, v], [p, w] \in P_x \times_\rho F$, where we used the fact that we can use the same element p to select representatives of the equivalence classes $[p, v]$ and $[p, w]$, since they belong to the same fiber. One can then define their sum and scalar multiple with $\lambda, \mu \in \mathbb{R}$ as

$$\lambda[p, v] + \mu[p, w] = [p, \lambda v + \mu w] , \tag{20.4.1}$$

where on the right hand side we use the vector space structure on F. Note that this is independent of the choice of the representative, since for any other representative $p \cdot g$ with $g \in G$ we have

$$\begin{aligned} \lambda[p, v] + \mu[p, w] &= \lambda[p \cdot g, \rho(g^{-1}, v)] + \mu[p \cdot g, \rho(g^{-1}, w)] \\ &= [p \cdot g, \lambda\rho(g^{-1}, v) + \mu\rho(g^{-1}, w)] \\ &= [p \cdot g, \rho(g^{-1}, \lambda v + \mu w)] \\ &= [p, \lambda v + \mu w] , \end{aligned} \tag{20.4.2}$$

using the fact that we chose ρ to be a linear representation of G on F. Further, we also need to check that the local trivializations of $P \times_\rho F$ restrict to vector space isomorphisms on every fiber. For this purpose we can use the fact that the restriction of the local trivialization $\phi_\rho : \pi_\rho^{-1}(U) \to U \times F$ we constructed earlier from a section $\sigma : U \to P$ on a set $U \subset M$ to a single fiber over $x \in U$ are simply given by the fiber diffeomorphisms $[p]^{-1} : P_x \times_\rho F \to F$ with $p = \sigma(x)$. These fiber diffeomorphisms are bijective by construction, and obviously linear, since

$$[p](\lambda v + \mu w) = [p, \lambda v + \mu w] = \lambda[p, v] + \mu[p, w] = \lambda[p](v) + \mu[p](w) \tag{20.4.3}$$

by definition of the vector space structure. It follows that $P \times_\rho F$ is indeed a vector bundle. ■

Recall from chapter 4 that a number of constructions which are known for vector spaces can also be applied to vector bundles, in particular the dual, direct sum and tensor product. It is known from representation theory that if the vector spaces to which these operations are applied carry linear representations of a group, then the vector spaces obtained from these operations also carry representations. Again one may ask whether this can be used also for associated vector bundles. We now show that this is indeed the case.

Theorem 20.4.2. *If F, F' are vector spaces and ρ, ρ' linear representations of G on F and F', respectively, then the associated bundles are related by*

$$(P \times_\rho F)^* \cong P \times_{\bar{\rho}} F^* , \tag{20.4.4}$$

$$(P \times_\rho F) \oplus (P \times_{\rho'} F') \cong P \times_{\rho \oplus \rho'} (F \oplus F') \tag{20.4.5}$$

and

$$(P \times_\rho F) \otimes (P \times_{\rho'} F') \cong P \times_{\rho \otimes \rho'} (F \otimes F') , \tag{20.4.6}$$

where $\bar{\rho}$ is the dual representation on the dual vector space F^* and $\rho \oplus \rho'$ and $\rho \otimes \rho'$ are the direct sum and tensor product of the representations ρ and ρ', and $\cong$ denotes that there exist canonical vector bundle isomorphisms.

Proof. This proof proceeds in two steps. First, we have to show that the fibers of the respective bundles agree with each other, and then ▶. . . ◀ ■

20.5 Associated affine bundles

Similarly to the case of linear representations on a vector space, one may also consider affine representations on an affine space. One may expect that in this case the associated fiber bundle will be an affine bundle. We now show that this is indeed the case.

Theorem 20.5.1. *If F is an affine space of dimension k modeled over the vector space $\vec{F}$ and ρ an affine representation of G with linear derivative $\vec{\rho}$ acting on $\vec{F}$, then the associated bundle $P \times_\rho F$ is a affine bundle of rank k modeled over the vector bundle $P \times_{\vec{\rho}} \vec{F}$.*

Proof. The proof is similar to that of theorem 20.4.1. We first show that there is an affine structure on each fiber, $P_x \times_\rho F$, which is inherited from the affine structure on F. Let $p \in P$ with $\pi(p) = x \in M$, which we use to express to elements $[p, a] \in P_x \times_\rho F$ and $[p, v] \in P_x \times_{\vec{\rho}} \vec{F}$ by representatives $a \in F$ and $v \in \vec{F}$. We define their sum as

$$[p, a] + [p, v] = [p, a + v] , \tag{20.5.1}$$

using the fact that F is an affine space modeled over $\vec{F}$. This definition is independent of the choice of the representative, since for any other representative $p \cdot g$ we find

$$\begin{aligned} [p, a] + [p, v] &= [p \cdot g, \rho(g^{-1}, a)] + [p \cdot g, \vec{\rho}(g^{-1}, v)] \\ &= [p \cdot g, \rho(g^{-1}, a) + \vec{\rho}(g^{-1}, v)] \\ &= [p \cdot g, \rho(g^{-1}, a + v)] \\ &= [p, a + v] , \end{aligned} \tag{20.5.2}$$

using the fact that ρ is an affine representation with linear derivative $\vec{\rho}$. Hence, $P_x \times_\rho F$ carries the structure of an affine space modeled over the vector space $P_x \times_{\vec{\rho}} \vec{F}$. Further, we check that the local trivializations of $P \times_\rho F$ restrict to affine space isomorphisms on every fiber. We make again use of the fact that the restriction of the local trivialization $\phi_\rho : \pi_\rho^{-1}(U) \to U \times F$ constructed from a section $\sigma : U \to P$ on a set $U \subset M$ to a single fiber over $x \in U$ is given by the fiber diffeomorphisms $[p]^{-1} : P_x \times_\rho F \to F$ with $p = \sigma(x)$. These fiber diffeomorphisms are bijective by construction and satisfy

$$[p](a + v) = [p, a + v] = [p, a] + [p, v] = [p](a) + [p](v) \tag{20.5.3}$$

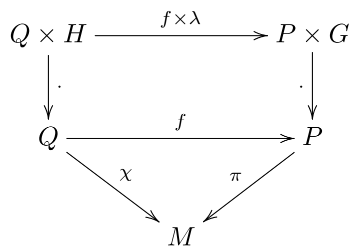
by definition of the affine space structure. It follows that $P \times_\rho F$ is indeed an affine bundle modeled over the vector bundle $P \times_{\vec{\rho}} \vec{F}$. ■

20.6 Reduction of the structure group

In the previous sections we have mostly discussed maps between bundles which have the same structure group. We now come to a different class of maps, which can be used to relate bundles with different structure groups. These different structure groups still need to be related to each other, and this relation will be given by a group homomorphism. We define the following notion, following [Bau14, sec. 2.5].

Definition 20.6.1 (Reduction of a principal bundle). Let $\pi : P \to M$ be a principal bundle with structure group G and $\lambda : H \to G$ a Lie group homomorphism. A λ-*reduction* of this bundle consists of a principal H-bundle $\chi : Q \to M$ and a map $f : Q \to P$ such that $\pi \circ f = \chi$ and $f(q \cdot h) = f(q) \cdot \lambda(h)$ for all $q \in Q$ and $h \in H$.

To see the similarity between this definition and the definition 20.2.1 of a principal bundle morphism, observe that we may summarize the maps appearing in definition 20.6.1 in the following commutative diagram:



(20.6.1)

By comparison with the diagram (20.2.1) defining a principal bundle morphisms, we see that the map f in definition 20.6.1 constitutes a principal bundle morphism covering the identity on the base manifold M.

Despite being called a "reduction", the group H does not have to be "smaller", i.e., a subgroup of G. There are important examples in physics, where H is, e.g., a double cover of G, which leads to spin groups and spin bundles, as we discuss in chapter 45. However, there are also numerous examples with H being a (closed) subgroup of G, and λ being the canonical inclusion. In this case we can find another possibility to express a reduction of the structure group, namely as a section of a particular associated bundle. This follows from the following theorem.

Theorem 20.6.1. *Let $\pi : P \to M$ be a principal bundle with structure group G and $H \subset G$ a closed subgroup. Then there exists a one-to-one correspondence between reductions of P with respect to the canonical inclusion $H \hookrightarrow G$ and global sections of the associated bundle $P \times_\rho G/H$, where $\rho : G \times G/H \to G/H$ is the action of G on the coset space via left multiplication.*

Proof. Here we make use of theorem 20.3.3 which states that we may equivalently express sections $\sigma \in \Gamma(P \times_\rho G/H)$ by equivariant maps $\phi \in C^\infty_G(P, G/H)$. Hence, we will show that there exists a one-to-one correspondence between principal bundle reductions and such equivariant maps.

Consider first an equivariant map $\phi \in C^\infty_G(P, G/H)$. Then we define the set

$$Q = \{p \in P, \phi(p) = eH\} \subset P \tag{20.6.2}$$

as those elements of P which are mapped to the coset eH belonging to the unit element $e \in G$, together with the canonical inclusion $Q \hookrightarrow P$ and projection $\chi = \pi|_Q : Q \to M$. To see that H acts from the right on Q, we use the equivariance of ϕ and the group action on P to find

$$\phi(q \cdot h) = h^{-1}\phi(q) = h^{-1}eH = eH\,, \tag{20.6.3}$$

so that $q \cdot h \in Q$ for all $q \in Q$ and $h \in H$. To see that this action is free and transitive on the fibers, let $q, q' \in Q$ with $\chi(q) = \chi(q')$. Since P is a principal bundle, there exists a unique group element $g \in G$ such that $q' = q \cdot g$. Using the equivariance of ϕ we have

$$eH = \phi(q') = \phi(q \cdot g) = g^{-1}\phi(q) = g^{-1}eH = g^{-1}H\,, \tag{20.6.4}$$

which means that e and g^{-1} define the same coset, and so $g \in H$. Hence, for every $q, q' \in Q$ with $\chi(q) = \chi(q')$, there exists a unique $h \in H$ such that $q' = q \cdot h$. This shows that $\chi : Q \to M$ is indeed a principal fiber bundle with structure group H. One now easily checks that the inclusion $Q \hookrightarrow P$ defines a reduction satisfying the properties given in definition 20.6.1.

To check the converse direction, let $\chi : Q \to M$ be a principal H-bundle, where $H \subset G$ is a closed subgroup, and $f : Q \to P$ a reduction of the principal bundle $\pi : P \to M$. For $(q, p) \in Q \times_M P$, i.e., $\chi(q) = \pi(p)$, define $\hat{\phi}(q, p) \in G$ as the unique element $g \in G$ such that $p \cdot g = f(q)$. Given another element $q' = q \cdot h \in Q$ within the same fiber, we have

$$p \cdot \hat{\phi}(q', p) = f(q') = f(q \cdot h) = f(q) \cdot h = p \cdot gh\,, \tag{20.6.5}$$

and so

$$\hat{\phi}(q \cdot h, p) = \hat{\phi}(q, p)h \tag{20.6.6}$$

with $h \in H$. Hence, $\hat{\phi}(q', p)$ and $\hat{\phi}(q, p)$ define the same coset, $\hat{\phi}(q', p)H = \hat{\phi}(q, p)H$. Since the coset is independent of q, we thus obtain a map $\phi : P \to G/H$. To see that this map is equivariant, let $p' = p \cdot g \in P$ with $g \in G$. Then we have

$$p \cdot g\hat{\phi}(q, p') = p' \cdot \hat{\phi}(q, p') = f(q) = p \cdot \hat{\phi}(q, p)\,, \tag{20.6.7}$$

and so

$$\phi(p \cdot q) = \phi(p') = \hat{\phi}(q, p')H = g^{-1}\hat{\phi}(q, p)H = g^{-1}\phi(p)\,, \tag{20.6.8}$$

which shows that ϕ is indeed equivariant, $\phi \in C^\infty_G(P, G/H)$.

One still needs to show that, starting from an equivariant map, constructing a reduction and again an equivariant map, one re-obtains the same equivariant map one has started from; and similarly, that starting from a reduction, one re-obtains the same reduction (up to isomorphism). We will not prove this here, but the proof is straightforward. ■

Recall that not every fiber bundle admits global sections. This statement also holds for the associated fiber bundle $P \times_\rho G/H$ we encountered in theorem 20.6.1. Hence, we conclude that, given a closed subgroup $H \subset G$, not every principal G-bundle $\pi : P \to M$ may be reduced to H. More generally, one finds that also for other homomorphisms $\lambda : H \to G$, bundle reductions may not exist in general, and may impose additional conditions on $\pi : P \to M$, sometimes related to the topology of M, which play an important role in physics.

We finally mention a few examples of reductions of principal bundles which we will encounter later and study in detail. These include unit frame bundles in section 24.1, oriented frame bundles in section 24.2, normalized frame bundles in section 24.3, orthonormal frame bundles in section 31.4, symplectic frame bundles in section 35.6, complex frame bundles in section 43.4 and spin bundles in section 45.5.

20.7 Extension of the structure group

We have seen in the previous section how a reduction of the structure group, given by a Lie group homomorphism, defines a relation between principal bundles, in form of a principal bundle homomorphism. We now come to a construction which is essentially the inverse of the principal bundle reduction, and we mostly follow [Bau14, sec. 2.5]. This is based on the following definition.

Definition 20.7.1 (Extension of a principle bundle)**.** Let $\chi : Q \to M$ be a principal fiber bundle with structure group H and $\lambda : H \to G$ a Lie group homomorphism. The λ-*extension* of Q is the bundle $\pi : P \to M$ defined by

$$P = Q \times_\rho G\,, \tag{20.7.1}$$

where $\rho : H \times G \to G$ is the left action defined by

$$\begin{array}{rcccc} \rho & : & H \times G & \to & G \\ & & (h, g) & \mapsto & \lambda(h)g \end{array}\,. \tag{20.7.2}$$

The extension of principal bundles has several important properties, which allow us to relate it to the reduction shown in the previous section. The most elementary is the following.

Theorem 20.7.1. *The λ-extension $\pi : P \to M$ constructed in definition 20.7.1 is a principal G-bundle.*

Proof. We have to show that G acts from the right on P, and that this action preserves the fibers and is free and transitive on them. For $q \in Q$ and $g, g' \in G$ we define this right action by

$$[q, g] \cdot g' = [q, gg']\,. \tag{20.7.3}$$

Note that this is independent of the choice of the representative q, since for $q' = q \cdot h$ with $h \in H$ we have

$$[q', \rho(h^{-1}, g)] \cdot g' = [q \cdot h, \lambda(h)^{-1}g] \cdot g' = [q \cdot h, \lambda(h)^{-1}gg'] = [q', \rho(h^{-1}, gg')]\,. \tag{20.7.4}$$

Further, it obviously defines a right action. This action is transitive, since for any $[q, g]$ and $[q, g']$ we have

$$[q, g] = [q, g'] \cdot g'^{-1}g\,, \tag{20.7.5}$$

and free, since $g'^{-1}g$ in the equation above is the unique element of G which satisfies this equation. Hence, it defines P as a principal G-bundle. ■

Given the λ-extension of a principal bundle, one may wonder whether and how it is possible to obtain a λ-reduction, and how this relates to the origin bundle Q. We see that this can be defined as follows.

Theorem 20.7.2. *Let $\pi : P \to M$ be the λ-extension of $\chi : Q \to M$ and*

$$\begin{array}{rcccc} f & : & Q & \to & P \\ & & q & \mapsto & [q, e] \end{array}\,, \tag{20.7.6}$$

where $e \in G$ is the unit element. Then f defines a λ-reduction of P.

Proof. We first check that f is a bundle map covering the identity on M. This follows from the relation

$$(\pi \circ f)(q) = \pi([q, e]) = \chi(q) \tag{20.7.7}$$

for all $q \in Q$. Further, it is a principal bundle homomorphism, since for any $h \in H$ we have

$$f(q \cdot h) = [q \cdot h, e] = [q, \lambda(h)e] = [q, e] \cdot \lambda(h) = f(q) \cdot \lambda(h) \,. \tag{20.7.8}$$

■

The statement above hence allows us to obtain a λ-reduction from the λ-extension. Now we pose the question whether also the converse is true, i.e., whether every λ-reduction can be related to or obtained from a λ-extension. This is indeed the case, up to isomorphism, which we show as follows.

Theorem 20.7.3. *Let $\pi : P \to M$ be a principal G-bundle and $f : Q \to P$ a λ-reduction of P. Then P is isomorphic to the λ-extension of Q.*

Proof. We sketch the proof by explicitly constructing the bundle isomorphism and showing that it is bijective. We define

$$\begin{array}{cccc} \psi & : & Q \times_\rho G & \to & P \\ & & [q, g] & \mapsto & f(q) \cdot g \end{array} \,. \tag{20.7.9}$$

To see that this is well-defined, let $h \in H$ and consider

$$\psi([q \cdot h, \lambda(h)^{-1}g]) = f(q \cdot h) \cdot \lambda(h)^{-1}g = f(q) \cdot \lambda(hh^{-1})g = f(q) \cdot g = \psi([q, g]) \,, \tag{20.7.10}$$

so that it is independent of the representative q. The map ψ is injective, since for any $g' \in G$ with $g \neq g'$ we have

$$\psi([q, g']) = f(q) \cdot g' \neq f(q) \cdot g = \psi([q, g]) \,, \tag{20.7.11}$$

since the action of G on the fibers of P is free. Finally, ψ is surjective, since for any $p \in P$ we can choose $q \in Q_{\pi(p)}$ and define $g \in G$ as the unique element satisfying $p = f(q) \cdot g$. This element exists, since G acts transitively on the fibers of P, and so we have $p = \psi([q, g])$. ■

Chapter 21

Jet manifolds and jet bundles

21.1 Contact and jets

In the previous lectures we have learned that tangent and cotangent vectors generalize the notions of the derivative of a function. Tangent vectors naturally appear as derivatives of curves $\gamma \in C^\infty(\mathbb{R}, M)$, while cotangent vectors appear as total derivatives of real functions $f \in C^\infty(M, \mathbb{R})$. We have also seen that the differential φ_* of a map $\varphi \in C^\infty(M, N)$ further generalizes this notion to maps between arbitrary manifolds. We now wish to generalize this notion to higher derivatives. In other words, we will generalize the notion of Taylor polynomials. These generalizations are called *jets*. In the most simple case of functions $f \in C^\infty(\mathbb{R}, \mathbb{R})$ they turn out to be exactly the Taylor polynomials. To arrive at this result, we start with the following definition.

Definition 21.1.1 (Contact on $C^\infty(\mathbb{R}, \mathbb{R})$). Two local functions $(U, f), (V, g) \in C_t^\infty(\mathbb{R}, \mathbb{R})$ with $t \in \mathbb{R}$ are said to have *r-contact* at t for $r \in \mathbb{N}$, written as $(U, f) \overset{t,r}{\sim} (V, g)$, if and only if

$$f^{(k)}(t) = g^{(k)}(t) \text{ for all } k \in \{0, \ldots, r\}\,. \tag{21.1.1}$$

Note that we have used *local* functions as discussed in section 1.4 here, as it will turn out to be more convenient later. This is sufficient for our purpose, since the value and derivatives of a function f at t only depend on the behavior of f on an arbitrarily small neighborhood of t, and we don't even need f to be defined anywhere outside this neighborhood. In the following, we want to take equivalence classes with respect to having contact. For this to be possible, the relation we defined must be an equivalence relation. We claim:

Theorem 21.1.1. *The relation $\overset{t,r}{\sim}$ of definition 21.1.1 is an equivalence relation.*

Proof. This is obvious from the definition. ∎

Hence, the following definition is valid.

Definition 21.1.2 (Jets of $C^\infty(\mathbb{R}, \mathbb{R})$). Let $(U, f) \in C_t^\infty(\mathbb{R}, \mathbb{R})$ be a real local function of one variable with $t \in \mathbb{R}$. For $r \in \mathbb{N}$, we define the *r-jet* $j_t^r f$ of f at t as the equivalence class

$$j_t^r f = \left\{ (V, g) \in C_t^\infty(\mathbb{R}, \mathbb{R}) \,\middle|\, (U, f) \overset{t,r}{\sim} (V, g) \right\} \tag{21.1.2}$$

of local functions $(V, g) \in C_t^\infty(\mathbb{R}, \mathbb{R})$ whose Taylor polynomials at t agree with the Taylor polynomial of f up to order r. The space of all r-jets is denoted

$$J^r(\mathbb{R}, \mathbb{R}) = \{j_t^r f, t \in \mathbb{R}, f \in C_t^\infty(\mathbb{R}, \mathbb{R})\}\,. \tag{21.1.3}$$

Note that we do not include the domain U in the notation $j_p^r f$, since it is not essential for the definition of the jet, and we will also omit it for the remaining notions of jets to be defined in this chapter; this will become more explicit in section 21.2. We will also frequently write just f instead of (U, f), whenever the domain is not important. In the following example, we illustrate the notions of contact and jets.

Example 21.1.1. Consider the three functions $f, g, h \in C^\infty(\mathbb{R}, \mathbb{R})$ defined by

$$f(t) = t - \frac{3\pi}{2}\,, \quad g(t) = 1 - \frac{t^2}{2}\,, \quad h(t) = \cos t\,, \tag{21.1.4}$$

and illustrated in figure 21.1. For the functions f and g, we have $f(t) = g(t)$, and hence 0-contact, for

$$t_{1,2} = -1 \pm \sqrt{3 + 3\pi}\,, \tag{21.1.5}$$

and so we write $f \overset{t_{1,2},0}{\sim} g$. They therefore define the same 0-jets $j_{t_{1,2}}^0 f = j_{t_{1,2}}^0 g$ at $t_{1,2}$. These are characterized by the common function value, or zeroth order Taylor coefficient,

$$f(t_{1,2}) = g(t_{1,2}) = -1 \pm \sqrt{3 + 3\pi} - \frac{3\pi}{2}\,, \tag{21.1.6}$$

which, together with the point $t_{1,2}$ itself where the functions are evaluated, uniquely determines the zeroth order jet. Note that the derivatives of f and g differ at these points, and so they do not have r-contact for any $r > 0$ and $t \in \mathbb{R}$. For all other $t \in \mathbb{R} \setminus \{t_1, t_2\}$, they have no 0-contact either.

The functions f and h satisfy $f(t) = h(t)$ and $f'(t) = h'(t)$ at $t = \frac{3\pi}{2}$. This means that at this point they have 1-contact, $f \overset{t,1}{\sim} h$, and hence also 0-contact, $f \overset{t,0}{\sim} h$. Hence, they have the same first-order jet $j_t^1 f = j_t^1 h$, which is characterized by

$$f(t) = h(t) = 0\,, \quad f'(t) = h'(t) = 1\,, \tag{21.1.7}$$

together with the value $t = \frac{3\pi}{2}$. Of course, they also have the same zeroth-order jet $j_t^0 f = j_t^0 h$ at the same value of t. All other jets, for $r > 1$ or $t \neq \frac{3\pi}{2}$, differ, $j_t^r f \neq j_t^r h$.

Finally, for g and h we find $g(0) = h(0)$, $g'(0) = h'(0)$ and $g''(0) = h''(0)$, and so $g \overset{0,2}{\sim} h$ (and hence also $g \overset{0,1}{\sim} h$ and $g \overset{0,0}{\sim} h$). Their highest order of contact at $t = 0$ is thus 2, and they have no contact at other values of t.

The example demonstrates that in the simple case of real functions, we can uniquely determine a jet $j_t^r f$ by a tuple

$$\left(t, f(t), f'(t), \ldots, f^{(r)}(t)\right)\,, \tag{21.1.8}$$

consisting of the point t and the values of the first $r + 1$ Taylor coefficients at this point. Understanding the elements of this tuple as coordinates, we thus have a mapping $J^r(\mathbb{R}, \mathbb{R}) \cong \mathbb{R}^{r+2}$. Another possibility, which is also found in the literature, is to identify the function value and derivatives with the finite Taylor polynomial

$$\tilde{t} \mapsto f(t) + f'(t)(\tilde{t} - t) + \ldots + \frac{f^{(r)}(t)}{r!}(\tilde{t} - t)^r\,, \tag{21.1.9}$$

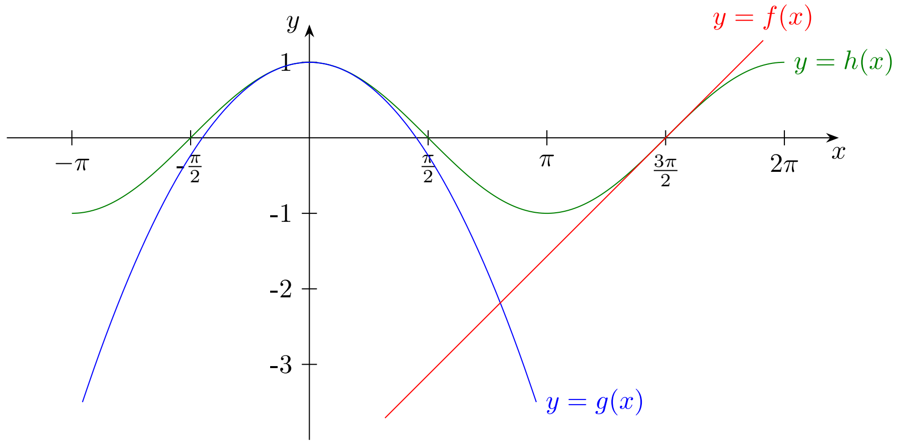


Figure 21.1: Contact of functions $f, g, h \in C^\infty(\mathbb{R}, \mathbb{R})$ given in example 21.1.1.

where also the point t must still be specified in order for the jet to be fully determined. This is obvious from the fact that a general polynomial of order r in $\tilde{t}$ has $r + 1$ coefficients, while we need an $(r + 2)$-tuple to specify a jet.

The importance of the base point t also becomes clear if we take a closer look at the definition of jets as equivalence classes with respect to the relation $\overset{t,r}{\sim}$. A local function (U, f) gives rise to a jet $j^r_t f$ for *every* $t \in U$. However, for distinct $t, t' \in U$ these are distinct, $j^r_t f \neq j^r_{t'} f$, even if all Taylor coefficients at these points agree, $f^{(k)}(t) = f^{(k)}(t')$ for all $k \in \mathbb{N}$. This is due to the fact that $j^r_t f$ is the equivalence class of all local functions (V, g) whose Taylor coefficients *at the point* t agree with those of f. The same local function (V, g) does not necessarily belong to the equivalence class $j^r_{t'} f$, and may not even be defined at t', if $t' \notin V$. Phrased differently, for each $t \in \mathbb{R}$ and $r \in \mathbb{N}$, the relation $\overset{t,r}{\sim}$ partitions the set $C^\infty_t(\mathbb{R}, \mathbb{R})$ of local functions into equivalence classes, namely the jets $j^r_t f$, and each local function (U, f) belongs to one equivalence class for each $t \in U$ and $r \in \mathbb{N}$; these equivalence classes can be distinguished by which other functions they contain, and this membership is uniquely linked to the point t at which all of their Taylor polynomials to order r agree.

Finally, one may pose the question why we take the technical burden of working with equivalence classes of local functions, instead of directly defining jets as tuples of Taylor coefficients or Taylor polynomials. The reason for this choice is that the more general notions of jets, which we will encounter in this chapter, will always be equivalence classes of maps and not possess the algebraic structure suggested by polynomials or the simple form of tuples; essentially, they will provide us with a coordinate-free notion of higher order derivatives of maps.. To see how this works, we will introduce the notions of contact and jets of curves.

Definition 21.1.3 (Contact on $C^\infty(\mathbb{R}, M)$)**.** Two local curves $(U, \gamma), (V, \beta) \in C^\infty_t(\mathbb{R}, M)$ on a manifold M with $t \in \mathbb{R}$ are said to have *r-contact* at t for $r \in \mathbb{N}$, written as $(U, \gamma) \overset{t,r}{\sim} (V, \beta)$, if and only if for all functions $f \in C^\infty(M, \mathbb{R})$ holds:

$$(U, f \circ \gamma) \overset{t,r}{\sim} (V, f \circ \beta)\,. \tag{21.1.10}$$

It should be clear that $(U, f \circ \gamma), (V, f \circ \beta) \in C^\infty_t(\mathbb{R}, \mathbb{R})$ are local functions, and so the relation $\overset{t,r}{\sim}$ here denotes the previously defined notion of contact. Again we wish to take equivalence classes, and so we need to make sure that this is an equivalence relation.

Theorem 21.1.2. *The relation* $\overset{t,r}{\sim}$ *of definition 21.1.3 is an equivalence relation.*

Proof. Reflexivity and symmetry are obvious from the definition. To check transitivity, consider $(U,\gamma),(V,\beta),(W,\zeta) \in C_t^\infty(\mathbb{R},M)$ with $(U,\gamma) \overset{t,r}{\sim} (V,\beta)$ and $(V,\beta) \overset{t,r}{\sim} (W,\zeta)$. Then for every $f \in C^\infty(M,\mathbb{R})$ holds

$$(U, f\circ\gamma) \overset{t,r}{\sim} (V, f\circ\beta)\,, \quad (V, f\circ\beta) \overset{t,r}{\sim} (W, f\circ\zeta)\,, \tag{21.1.11}$$

and thus also

$$(U, f\circ\gamma) \overset{t,r}{\sim} (W, f\circ\zeta)\,, \tag{21.1.12}$$

using the fact that $\overset{t,r}{\sim}$ is an equivalence relation for local functions. ■

Definition 21.1.4 (Jets of $C^\infty(\mathbb{R},M)$)**.** Let M be a manifold and $(U,\gamma) \in C_t^\infty(\mathbb{R},M)$ a local curve on M with $t \in \mathbb{R}$. For $r \in \mathbb{N}$, we define the *r-jet* $j_t^r\gamma$ of γ at t as the equivalence class

$$j_t^r\gamma = \left\{(V,\beta) \in C_t^\infty(\mathbb{R},M) \,\middle|\, (U,\gamma) \overset{t,r}{\sim} (V,\beta)\right\} \tag{21.1.13}$$

of local curves $(V,\beta) \in C_t^\infty(\mathbb{R},M)$ having r-contact. The space of all r-jets is denoted

$$J^r(\mathbb{R},M) = \{j_t^r\gamma, t \in \mathbb{R}, \gamma \in C_t^\infty(\mathbb{R},M)\}\,. \tag{21.1.14}$$

In this definition we have now made use of the previously introduced notion of contact for (local) curves, which we have shown to be an equivalence relation. The condition given in definition 21.1.3, which demands that $(U, f\circ\gamma),(V, f\circ\beta) \in C_t^\infty(\mathbb{R},\mathbb{R})$ for *all* $f \in C^\infty(M,\mathbb{R})$ seems strong at first sight. However, this intuition may be misleading. To illustrate this, we will explicitly construct the first order jets.

Example 21.1.2 (First order jets of $C^\infty(\mathbb{R},M)$). Let M be a manifold, $t \in \mathbb{R}$ and $\gamma \in C_t^\infty(\mathbb{R},M)$ a local curve on M. The 1-jet $j_t^1\gamma$ is the equivalence class of curves $\beta \in C_t^\infty(\mathbb{R},M)$ such that for all functions $f \in C^\infty(M,\mathbb{R})$ we have $(f\circ\gamma)(t) = (f\circ\beta)(t)$ and $(f\circ\gamma)'(t) = (f\circ\beta)'(t)$. We start with the first condition. Assume that $\gamma(t) \neq \beta(t)$. Then we could choose a function $f \in C^\infty(M,\mathbb{R})$ such that $f(\gamma(t)) \neq f(\beta(t))$, contradicting the first condition. Hence, it follows that $\gamma(t) = \beta(t)$. For the second condition, recall that

$$(f\circ\gamma)'(t) = \dot\gamma(t)(f)\,, \tag{21.1.15}$$

and analogously for β, by the definition 7.3.1 of the tangent vector of a curve. Assuming $\dot\gamma(t) \neq \dot\beta(t)$, one could find a function $f \in C^\infty(M,\mathbb{R})$ such that $\dot\gamma(t)(f) \neq \dot\beta(t)(f)$, contradicting the second condition. Hence, this condition implies $\dot\gamma(t) = \dot\beta(t)$. These conditions are both necessary and sufficient, so that each equivalence class $j_t^1\gamma$ is uniquely described by the point $\gamma(t) \in M$ and the tangent vector $\dot\gamma(t) \in T_{\gamma(t)}M$. Hence, there is a one-to-one correspondence between jets $j_t^1\gamma$ and pairs $(t,v) \in \mathbb{R}\times TM$, where $t \in \mathbb{R}$ denotes the value of the curve parameter where the jet is taken and $v \in TM$ fixes the tangent vector $\dot\gamma(t)$ at t (and thus also the point $\gamma(t) = \tau(\dot\gamma(t))$). Hence, we have a bijection $J^1(\mathbb{R},M) \cong \mathbb{R}\times TM$.

This shows that 1-jets are in this case simply tangent vectors, so that jets generalize the concept of tangent vectors. To see that we can also generalize cotangent vectors, we continue by constructing jets of real functions on manifolds. Again we start by defining a notion of contact.

Definition 21.1.5 (Contact on $C^\infty(M,\mathbb{R})$). Two local functions $(U,f),(V,g) \in C_p^\infty(M,\mathbb{R})$ on M with $p \in M$ are said to have *r-contact* at p for $r \in \mathbb{N}$, written as $(U,f) \overset{p,r}{\sim} (V,g)$, if and only if for all local curves $(W,\gamma) \in C_0^\infty(\mathbb{R}, U \cap V)$ with $\gamma(0) = p$ holds:

$$(W, f \circ \gamma) \overset{0,r}{\sim} (W, g \circ \gamma)\,. \tag{21.1.16}$$

The construction is very similar to the previous case of curves on M. We have simply reversed the order of composition in order to obtain a local function $f \circ \gamma \in C_0^\infty(\mathbb{R},\mathbb{R})$, and, without loss of generality, chosen the curve parameter at which the curve passes through p to be 0. However, a few technicalities arise from the fact that the local functions (U,f) and (V,g) will have different domains in general, which we have dealt with by restricting the image of the curve γ. These technicalities, which can also be overcome by working with germs as we see in section 21.2, are also relevant in showing that we have defined an equivalence relation.

Theorem 21.1.3. *The relation* $\overset{p,r}{\sim}$ *of definition 21.1.5 is an equivalence relation.*

Proof. Reflexivity and symmetry are obvious from the definition. To check transitivity, consider $(U,f),(V,g),(W,h) \in C_p^\infty(M,\mathbb{R})$ with $(U,f) \overset{t,r}{\sim} (V,g)$ and $(V,g) \overset{t,r}{\sim} (W,h)$. Further, let $(Y,\gamma) \in C_0^\infty(\mathbb{R}, U \cap W)$, and define

$$(Z,\tilde{\gamma}) = (\gamma^{-1}(U \cap V \cap W), \gamma|_{\gamma^{-1}(U\cap V\cap W)}) \in C_0^\infty(\mathbb{R}, U \cap V \cap W)\,. \tag{21.1.17}$$

Now we have the inclusion $C_0^\infty(\mathbb{R}, U\cap V\cap W) \subset C_0^\infty(\mathbb{R}, U \cap V)$, and from $(U,f) \overset{0,r}{\sim} (V,g)$ follows

$$(Z, f \circ \tilde{\gamma}) \overset{0,r}{\sim} (Z, g \circ \tilde{\gamma})\,, \tag{21.1.18}$$

where $\overset{0,r}{\sim}$ here denotes contact of a real function of a real variable. Similarly, one shows

$$(Z, g \circ \tilde{\gamma}) \overset{0,r}{\sim} (Z, h \circ \tilde{\gamma})\,, \tag{21.1.19}$$

and from theorem 21.1.1 then also follows

$$(Z, f \circ \tilde{\gamma}) \overset{0,r}{\sim} (Z, h \circ \tilde{\gamma})\,. \tag{21.1.20}$$

Further, γ and $\tilde{\gamma}$ agree on an open set Z containing 0, and so they yield the same Taylor series when composed with a function. Hence,

$$(Y, f \circ \gamma) \overset{0,r}{\sim} (Z, f \circ \tilde{\gamma})\,, \quad (Y, h \circ \gamma) \overset{0,r}{\sim} (Z, h \circ \tilde{\gamma})\,, \tag{21.1.21}$$

and once more employing theorem 21.1.1, finally

$$(Y, f \circ \gamma) \overset{0,r}{\sim} (Y, h \circ \gamma)\,. \tag{21.1.22}$$

Since this construction is valid for all $(Y,\gamma) \in C_0^\infty(\mathbb{R}, U\cap W)$, we conclude that $(U,f) \overset{p,r}{\sim} (W,h)$, showing that $\overset{p,r}{\sim}$ is transitive, and thus an equivalence relation. ■

With the preceding result at hand, we can now define jets of functions on a manifold as equivalence classes.

Definition 21.1.6 (Jets of $C^\infty(M,\mathbb{R})$)**.** Let M be a manifold and $(U,f) \in C^\infty_p(M,\mathbb{R})$ a local function on a manifold M with $p \in M$. For $r \in \mathbb{N}$, we define the *r-jet* $j^r_p f$ of f at p as the equivalence class

$$j^r_p f = \left\{(V,g) \in C^\infty_p(M,\mathbb{R}) \,\middle|\, (U,f) \overset{p,r}{\sim} (V,g)\right\} \tag{21.1.23}$$

of local functions $(V,g) \in C^\infty_p(M,\mathbb{R})$ having r-contact. The space of all r-jets is denoted

$$J^r(M,\mathbb{R}) = \{j^r_p f, p \in M, f \in C^\infty_p(M,\mathbb{R})\}\,. \tag{21.1.24}$$

As an example, we construct the first order jets.

Example 21.1.3 (First order jets of $C^\infty(M,\mathbb{R})$). Let M be a manifold, $p \in M$ and $f \in C^\infty_p(M,\mathbb{R})$ a real local function on M. The 1-jet $j^1_p f$ is the equivalence class of local functions $g \in C^\infty_p(M,\mathbb{R})$ such that for all local curves $\gamma \in C^\infty_0(\mathbb{R},M)$ with $\gamma(0) = p$ we have $(f\circ\gamma)(0) = (g\circ\gamma)(0)$ and $(f\circ\gamma)'(0) = (g\circ\gamma)'(0)$. The first condition obviously states that

$$f(p) = f(\gamma(0)) = g(\gamma(0)) = g(p)\,, \tag{21.1.25}$$

since $\gamma(0) = p$ is fixed by allowing only such curves γ which satisfy this condition. Further, for any $v \in T_pM$ one can find such a curve γ with $\dot\gamma(0) = v$. Therefore, the second condition is satisfied if and only if $\mathrm{d}f(p) = \mathrm{d}g(p)$. In other words, each equivalence class is uniquely described by the point $p \in M$, the function value $f(p) \in \mathbb{R}$ and the value $\mathrm{d}f(p) \in T^*_pM$ of its total derivative at p. Note that the latter is a covector $\mathrm{d}f(p) \in T^*M$ with $\bar\tau(\mathrm{d}f(p)) = p$, so that there is a one-to-one correspondence between jets $j^1_p f$ and pairs $(\alpha,u) \in T^*M \times \mathbb{R}$, where $\alpha \in T^*M$ fixes the covector and $u \in \mathbb{R}$ the function value. Hence, we have a bijection $J^1(M,\mathbb{R}) \cong T^*M \times \mathbb{R}$.

This shows that jets also generalize the concept of cotangent vectors. But the most powerful property of jets is the fact that we can also extend the definition to jets of maps between arbitrary manifolds. Now we can define contact as follows.

Definition 21.1.7 (Contact on $C^\infty(M,N)$)**.** For manifolds M,N, two local maps $(U,\varphi),(V,\vartheta) \in C^\infty_p(M,N)$ with $p \in M$ are said to have *r-contact* at p for $r \in \mathbb{N}$, written as $(U,\varphi) \overset{p,r}{\sim} (V,\vartheta)$, if and only if for all local curves $(W,\gamma) \in C^\infty_0(\mathbb{R},U\cap V)$ with $\gamma(0) = p$ and all functions $f \in C^\infty(N,\mathbb{R})$ holds:

$$(W, f\circ\varphi\circ\gamma) \overset{0,r}{\sim} (W, f\circ\vartheta\circ\gamma)\,. \tag{21.1.26}$$

Note that we have simply combined definitions 21.1.3 and 21.1.5, by composing with both a curve and a real function. As one expects, this also yields an equivalence relation.

Theorem 21.1.4. *The relation* $\overset{p,r}{\sim}$ *of definition 21.1.7 is an equivalence relation.*

Proof. The proof combines two previous proofs. One follows the same steps as in the proof of theorem 21.1.3, but with f and g replaced by $f\circ\varphi$ and $f\circ\vartheta$, respectively, as done in the proof of theorem 21.1.2. ∎

Now the final definition is straightforward.

Definition 21.1.8 (Jets of $C^\infty(M,N)$). Let M,N be manifolds and $(U,\varphi) \in C_p^\infty(M,N)$ a local map with $p \in M$. For $r \in \mathbb{N}$, we define the *r-jet* $j_p^r\varphi$ of φ at p as the equivalence class

$$j_p^r\varphi = \left\{(V,\vartheta) \in C_p^\infty(M,N) \,\middle|\, (U,\varphi) \overset{p,r}{\sim} (V,\vartheta)\right\} \tag{21.1.27}$$

of local functions $(V,\vartheta) \in C_p^\infty(M,N)$ having r-contact. The space of all r-jets is denoted

$$J^r(M,N) = \{j_p^r\varphi, p \in M, \varphi \in C_p^\infty(M,N)\}\,. \tag{21.1.28}$$

We have seen in the previous examples that if M or N is given by the real line $\mathbb{R}$, one can identify the first order jet spaces $J^1(M,N)$ with either the tangent or cotangent bundle, up to a direct factor $\mathbb{R}$. For the general case, the geometric picture is less obvious, since it involves properties of both manifolds M and N. We will shed more light on the geometry of these jet spaces in section 21.3.

21.2 Contact, jets and germs

The use of local maps and the fact that we used only derivatives of such maps in a single point already suggests that the jet $j_p^r\varphi$ of a local map φ remains unchanged if we replace φ by another local map which belongs to the same germ. We now show that this is indeed the case.

Theorem 21.2.1. *Let M,N be manifolds and $p \in M$. If two local maps $(U,\varphi),(V,\vartheta) \in C_p^\infty(M,N)$ have the same germ, $[U,\varphi] = [V,\vartheta]$, then $(U,\varphi) \overset{p,r}{\sim} (V,\vartheta)$ for all $r \in \mathbb{N}$.*

Proof. Let $(Y,\gamma) \in C_0^\infty(\mathbb{R}, U \cap V)$ be a local curve and $f \in C^\infty(N,\mathbb{R})$ a function. Since (U,φ) and (V,ϑ) have the same germ at p, there exists an open set $W \subseteq U \cap V$ containing p, such that $\varphi|_W = \vartheta|_W$. Now we can define another local curve $(Z,\tilde{\gamma})$ by

$$(Z,\tilde{\gamma}) = (\gamma^{-1}(W), \gamma|_{\gamma^{-1}(W)}) \in C_0^\infty(\mathbb{R}, W)\,. \tag{21.2.1}$$

Now it is clear that

$$\varphi \circ \tilde{\gamma} = \varphi|_W \circ \tilde{\gamma} = \vartheta|_W \circ \tilde{\gamma} = \vartheta \circ \tilde{\gamma}\,, \tag{21.2.2}$$

where the first and last equality hold since the image of $\tilde{\gamma}$ lies within W by construction, while the central equality holds since $\varphi|_W = \vartheta|_W$. Finally, we have

$$(Y, f \circ \varphi \circ \gamma) \overset{0,r}{\sim} (Z, f \circ \varphi \circ \tilde{\gamma})\,, \quad (Y, f \circ \vartheta \circ \gamma) \overset{0,r}{\sim} (Z, f \circ \vartheta \circ \tilde{\gamma})\,, \tag{21.2.3}$$

since γ and $\tilde{\gamma}$ agree on an open set Z containing 0 by definition. Combining these relations, and using the fact that $\overset{0,r}{\sim}$ is an equivalence relation, we find

$$(Y, f \circ \varphi \circ \gamma) \overset{0,r}{\sim} (Y, f \circ \vartheta \circ \gamma)\,. \tag{21.2.4}$$

Since this holds for all $(Y,\gamma) \in C_0^\infty(\mathbb{R}, U \cap V)$ and $f \in C^\infty(N,\mathbb{R})$, it follows that $(U,\varphi) \overset{p,r}{\sim} (V,\vartheta)$. ■

This result is very useful, as it allows us to replace a local map with any other local map with the same germ, if it only appears as a representative of a jet. In particular, note that a local map (U,φ) has the same germ as any restriction $(\tilde{U}, \varphi|_{\tilde{U}})$ to an open subset $\tilde{U} \subset U$, and vice versa. This means that we may choose the domain of the representative at our convenience, since this choice has no influence on the jet. We will make use of this freedom in the remainder of this chapter.

One may pose the question whether also the converse holds, i.e., whether a germ is uniquely determined by a (possibly infinite) sequence of jets of a local function. However, this is not the case, which can be seen as follows. Let $M = N = \mathbb{R}$ and $p = 0$, and consider

$$f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \begin{cases} 0 & \text{for } x = 0\,, \\ \exp(-x^{-2}) & \text{otherwise,} \end{cases} \tag{21.2.5}$$

as well as

$$g : \mathbb{R} \to \mathbb{R}, \quad x \mapsto 0 \tag{21.2.6}$$

Clearly, we have $f^{(r)}(0) = g^{(r)}(0) = 0$ for all $r \in \mathbb{N}$, and thus $j_0^r f = j_0^r g$. However, f and g do not have the same germ, since $f(x) \neq 0 = g(x)$ for all $x \neq 0$, and so there is no open set on which they agree.

21.3 Jet manifolds

In the examples discussed in section 21.1 it appeared that the jet spaces would be vector spaces, which may seem logical, since the Taylor polynomials we compared them to form vector spaces. However, for general jet spaces this is *not* the case. This false intuition comes from the fact that functions $\varphi \in C^\infty(\mathbb{R}^m, \mathbb{R}^n)$ form a vector space, whose structure comes from the vector space structure of $\mathbb{R}^n$. For maps between general manifolds there is no such structure. Nevertheless, the jet spaces carry a number of other structures, which we will explore in this section, mostly following [KSM93]. The first structure is the following.

Theorem 21.3.1. *Let M, N be manifolds of dimensions $\dim M = m, \dim N = n$ and $r \in \mathbb{N}$. Then the jet space $J^r(M, N)$ is a manifold of dimension $m + n\binom{m+r}{r}$.*

Proof. In order to equip the mentioned spaces with the structure of a manifold, we need to construct atlases. For this purpose, we make use of the fact that M and N are manifolds, and thus equipped with atlases. Let $p \in M$ and consider a jet $j_p^r\varphi$. This jet uniquely determines a point $q = \varphi(p) \in N$, which is independent of the representative φ of the jet. Let (U, ϕ) be a chart of M containing p and (V, χ) be a chart of N containing q. Then we can express a curve $\gamma : \mathbb{R} \to M$ with $\gamma(0) = p$ and a function $f : N \to \mathbb{R}$ locally as

$$\gamma_\phi = \phi \circ \gamma|_{\gamma^{-1}(U)}\,, \quad f_\chi = f \circ \chi^{-1}\,. \tag{21.3.1}$$

In order to describe a jet $j_p^r\varphi$, it will be sufficient to consider these local descriptions, since the jet depends only on function values and derivatives at the point p, which are given by

$$j_0^r(f \circ \varphi \circ \gamma) = j_0^r(f_\chi \circ \chi \circ \varphi \circ \phi^{-1} \circ \gamma_\phi)\,. \tag{21.3.2}$$

Here $\chi \circ \varphi \circ \phi^{-1}$ is a map between open subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively. Recall that the jet given above is defined via the derivatives up to order r. Using the chain rule, it follows that these are fully determined by the partial derivatives of γ_ϕ, f_χ and $\chi \circ \varphi \circ \phi^{-1}$ up to order r at the respective points. Hence, any functions $\varphi, \vartheta : M \to N$ for which $\chi \circ \varphi \circ \phi^{-1}$ and $\chi \circ \vartheta \circ \phi^{-1}$ have the same partial derivatives at $\phi(p)$ will define the same jet. Conversely, if any of the aforementioned derivatives differs, we can find some function f and curve γ such that

$$j_0^r(f \circ \varphi \circ \gamma) \neq j_0^r(f \circ \vartheta \circ \gamma)\,. \tag{21.3.3}$$

Hence, we conclude that the jet $j_p^r\varphi$ is uniquely determined by the coordinates $\phi(p)$ of the point p and the partial derivatives of $\chi \circ \varphi \circ \phi^{-1}$ up to order r at $\phi(p)$, and vice versa $j_p^r\varphi$ uniquely determines these values. We can therefore use the aforementioned data as coordinates for a chart (W, ψ), where

$$W = \{j_p^r\varphi \in J^r(M, N) | p \in U \wedge \varphi(p) \in V\} \tag{21.3.4}$$

and $\psi : W \to \mathbb{R}^k$ assigns to $j^r_p\varphi$ the values of the aforementioned derivatives. We still need to determine the value k, which will become the dimension of the manifold. First note that $\phi(p) \in \mathbb{R}^m$, giving m coordinates. We are thus left with the partial derivatives of $\chi \circ \varphi \circ \phi^{-1}$. Here we can consider each of the n coordinates assigned by χ on its own. Recalling that we consider only smooth maps, their partial derivatives commute, and so there are $\binom{m+r'-1}{r'}$ partial derivatives with respect to m variables, such that their total order equals r'. For the jet $j^r_p\varphi$, we need the partial derivatives from zeroth order (which specify the point $\varphi(p)$) up to order r. For each coordinate on N, we thus have

$$\sum_{r'=0}^{r} \binom{m+r'-1}{r'} = \binom{m+r}{r} \tag{21.3.5}$$

partial derivatives to consider, and so the total number of coordinates on $J^r(M, N)$ is

$$k = m + n\binom{m+r}{r}. \tag{21.3.6}$$

Keeping in mind that M and N are covered by charts, we can perform the construction detailed above for any jet $j^r_p\varphi$, and thus construct charts which will cover $J^r(M, N)$. Finally, for any other pair (U', ϕ') and (V', χ') of charts, which are compatible with (U, ϕ) and (V, χ), respectively, so that their transition functions are smooth bijections, since we consider only smooth manifolds, the partial derivatives of $\chi \circ \varphi \circ \phi^{-1}$ and $\chi' \circ \varphi \circ \phi'^{-1}$ are bijectively and smoothly related to each other via the derivatives of the aforementioned transition function. Hence, the charts we obtain via the construction detailed above are compatible. We have thus constructed an atlas on $J^r(M, N)$, and so equipped it with the structure of a smooth manifold. ■

We can illustrate the construction given in the proof above by explicitly constructing the coordinates on the given jet manifolds. Let (x^α) be coordinates on M and (y^a) coordinates on N, with Greek indices in the range $1, \ldots, \dim M$ and Latin indices in the range $1, \ldots, \dim N$. In these coordinates a map $\varphi : M \to N$ can be expressed by the coordinate functions $y(x)$. The r-jet of φ is then given by those maps $\vartheta : M \to N$ which have the same Taylor polynomial

$$\sum_{\lambda_1+\ldots+\lambda_m \leq r} \frac{(x^1 - x^1_0)^{\lambda_1} \cdot \ldots \cdot (x^m - x^m_0)^{\lambda_m}}{\lambda_1! \cdot \ldots \cdot \lambda_m!} \frac{\partial^{\lambda_1+\ldots+\lambda_m}}{(\partial x^1)^{\lambda_1} \cdots (\partial x^m)^{\lambda_m}} y^a(x_0) \tag{21.3.7}$$

up to order r around a chosen point p with coordinates x^α_0. A r-jet $j^r_p\varphi$ is thus uniquely determined by the coordinates (x^α) of the base point, the values of the coordinate functions $y^a(x_0)$ and their derivatives of order at most r at x_0. We will use these values as coordinates on $J^r(M, N)$. To simplify the notation, we define a *multiindex* $\Lambda = (\lambda_1, \ldots, \lambda_m)$ to be an m-tuple of natural numbers $\lambda_\alpha \in \mathbb{N}$ and denote their sum by $|\Lambda|$. For the $|\Lambda|$'th order derivative appearing in the Taylor polynomial we simply write $\partial_\Lambda y^a(x_0)$. In this notation, a r-jet $j^r_p\varphi$ is uniquely determined by (x^α) and the values $\partial_\Lambda y^a(x_0)$ for $0 \leq |\Lambda| \leq r$. This allows us to use them as coordinates $(x^\alpha, y^a_\Lambda = \partial_\Lambda y^a(x_0))$ on $J^r(M, N)$.

Note that some authors also use the notation $(x^\alpha, y^a, y^a_\Lambda)$ for coordinates on $J^r(M, N)$ instead, where $1 \leq |\Lambda| \leq r$. In other words, the coordinates $y^a_{(0,\ldots,0)}$ are instead denoted y^a. This is of course equivalent to our choice of coordinates.

Once again it should be noted that despite their nice coordinate form, which maps jets into a vector space of polynomials, there is no vector space structure on jets, i.e., there is no way to treat them as vectors. It is only their coordinate representation we used here that has this structure, but it is not defined on the jets themselves without using coordinates, and is not independent of the choice of coordinates.

Example 21.3.1. Let $\dim M = 2$ and $\dim N = 1$. We use coordinates (x^1, x^2) on M and the coordinate y on N in order to construct coordinates on $J^3(M, N)$. Here we need to consider the multiindices

$$\Lambda \in \{(0,0), (1,0), (0,1), (2,0), (1,1), (0,2), (3,0), (2,1), (1,2), (0,3)\}. \tag{21.3.8}$$

On $J^3(M,N)$, we thus have coordinates (x^1, x^2, y_Λ) with Λ taking the values above, so that $\dim J^3(M,N) = 12$. This agrees with the dimension formula above.

We can then continue by exploring the general structure of jet manifolds for a given pair M, N of manifolds, as well as the relations between them. We start with the most basic jet space, which is of zeroth order. Its structure is fairly simple.

Theorem 21.3.2. *For manifolds M, N, the zeroth-order jet space $J^0(M,N)$ is canonically diffeomorphic to $M \times N$.*

Proof. Let $p \in M$ and $\varphi \in C^\infty_p(M,N)$. Its zeroth-order jet j^0_p is the equivalence class

$$j^0_p\varphi = \left\{\vartheta \in C^\infty_p(M,N) \,\middle|\, \varphi \overset{p;0}{\sim} \vartheta\right\} . \tag{21.3.9}$$

As discussed in the proof of theorem 21.3.1, this jet is uniquely defined by $p \in M$ and $\varphi(p) \in N$, hence by $(p, \varphi(p)) \in M \times N$ and vice versa. Further, given charts (U, ϕ) of M and (V, χ) of N with $p \in U$ and $\varphi(p) \in V$, one finds that the chart (W, ψ) of $J^0(M,N)$ constructed in the proof of theorem 21.3.1 agrees with that of the product manifold $M \times N$, showing that these manifolds are diffeomorphic. ■

Next, we aim to study higher order jet spaces. To approach their structure, we introduce the following definition.

Definition 21.3.1 (Jet projection)**.** Let M, N be manifolds and $k, r \in \mathbb{N}$ with $r > k$. Then we call $\pi_{r,k} : J^r(M,N) \to J^k(M,N), j^r_p\varphi \mapsto j^k_p\varphi$ the *jet projection* from $J^r(M,N)$ to $J^k(M,N)$.

It should be clear that the jet projections $\pi_{r,k}$ are well-defined, since a r-jet uniquely determines a k-jet for $r > k$, which is obtained by "forgetting" the terms in the Taylor polynomial whose order is larger than k. Since both $J^r(M,N)$ and $J^k(M,N)$ are manifolds, one may expect that they are smooth maps. We will show this next, among some of its properties.

Theorem 21.3.3. *The jet projections $\pi_{r,k}$ are surjective submersions.*

Proof. Let $p \in M$, $\varphi \in C^\infty_p(M,N)$ and consider charts (U, ϕ) of M and (V, χ) of N with $p \in U$ and $\varphi(p) \in V$, as well as the induced charts (W_r, ψ_r) of $J^r(M,N)$ and (W_k, ψ_k) of $J^k(M,N)$, with $j^r_p\varphi \in W_r$ and $j^k_p\varphi = \pi_{r,k}(j^r_p\varphi) \in W_k$. In these charts, the projection $\pi_{r,k}$ is represented by a map $W_r \to W_k$ which omits those coordinates which correspond to derivatives of order higher than k. This map is smooth, and hence also $\pi_{r,k}$ is smooth. Further, it is a submersion, and so $\pi_{r,k}$ is a submersion. Finally, for all jets $j^k_p\varphi \in J^k(M,N)$ there exists at least one representative $\varphi \in C^\infty_p(M,N)$, and thus a jet $j^r_p\varphi \in J^r(M,N)$ with $j^k_p\varphi = \pi_{r,k}(j^r_p\varphi)$, showing that $\pi_{r,k}$ is surjective. ■

One may thus already suspect that these maps define the projections of fiber bundles. Before we can show this, we need to study the local and global structure of jet manifolds in more detail. For this purpose, we first make use of the fact that $J^0(M,N) \cong M \times N$, in order to define the following maps.

Definition 21.3.2 (Source and target projections)**.** For every jet manifold $J^r(M,N)$ we call

$$\begin{array}{rcccc} \mathrm{pr}_1 \circ \pi_{r,0} & : & J^r(M,N) & \to & M \\ & & j^r_p\varphi & \mapsto & p \end{array} \tag{21.3.10}$$

the *source projection* and

$$\begin{array}{rcccc} \mathrm{pr}_2 \circ \pi_{r,0} & : & J^r(M,N) & \to & N \\ & & j^r_p\varphi & \mapsto & \varphi(p) \end{array} \tag{21.3.11}$$

the *target projection.*

Hence, we have the following "tower" of surjective submersions:

$$\begin{array}{ccccc} & & J^r(M,N) & & \\ & & \downarrow{\scriptstyle \pi_{r,r-1}} & & \\ & & J^{r-1}(M,N) & & \\ & & \vdots & & \\ & & J^1(M,N) & & \\ & & \downarrow{\scriptstyle \pi_{1,0}} & & \\ & & J^0(M,N) \cong M \times N & & \\ & \swarrow{\scriptstyle \mathrm{pr}_1} & & \searrow{\scriptstyle \mathrm{pr}_2} & \\ M & & & & N \end{array} \tag{21.3.12}$$

These allow us to define the following spaces, where we make use of the relation $J^0(M,N) \cong M \times N$ to canonically identify elements of these manifolds.

Definition 21.3.3 (Jet spaces)**.** For every jet manifold $J^r(M,N)$, $p \in M$ and $q \in N$ we define the *jet spaces*:

$$J^r_p(M,N) = \pi^{-1}_{r,0}(\mathrm{pr}_1^{-1}(p))\,, \tag{21.3.13a}$$
$$J^r(M,N)_q = \pi^{-1}_{r,0}(\mathrm{pr}_2^{-1}(q))\,, \tag{21.3.13b}$$
$$J^r_p(M,N)_q = \pi^{-1}_{r,0}(p,q) = J^r_p(M,N) \cap J^r(M,N)_q\,. \tag{21.3.13c}$$

In order for these spaces to become the fibers of a fiber bundle, we need to show that they are diffeomorphic to a typical fiber manifold. To show this, let us first state the following property.

Theorem 21.3.4. *The jet spaces $J^r_p(M,N)$, $J^r(M,N)_q$ and $J^r_p(M,N)_q$ are embedded submanifolds of $J^r(M,N)$.*

Proof. This follows from theorem 13.2.2 and the fact that the maps $\pi_{r,0}$, $\mathrm{pr}_1 \circ \pi_{r,0}$ and $\mathrm{pr}_2 \circ \pi_{r,0}$ are submersions. ■

Next, we will show that these manifolds are "similar", i.e., diffeomorphic, for different choices of the points p and q. As we have already seen in the proof of theorem 21.3.1, a jet $j^r_p\varphi$ is fully determined by the point p as well as the value and finitely many partial derivatives of the coordinate expression for φ at the point p, which is determined by the behavior of the function φ in an arbitrarily small neighborhood of the point p, where it takes values in some neighborhood of $\varphi(p) = q$, which can also be chosen arbitrarily small. One may therefore conclude that the manifold $J_p^r(M, N)_q$ is independent of the global structure of M and N, while for $J_p^r(M, N)$ and $J^r(M, N)_q$ the global structure of N enters through the freedom to choose the point $\varphi(p)$ in the former case, and that of M in the freedom to choose the base point in the latter case. We now show this explicitly.

Theorem 21.3.5. *Let M, N be manifolds, $U \subseteq M$ and $V \subseteq N$ open sets. Then the following manifolds are diffeomorphic:*

1. $J_p^r(U, N) \cong J_p^r(M, N)$ *for all* $p \in U$,
2. $J^r(M, V)_q \cong J^r(M, N)_q$ *for all* $q \in V$,
3. $J_p^r(U, V)_q \cong J_p^r(M, N)_q$ *for all* $p \in U$ *and* $q \in V$.

Proof.

1. First note that every local map $(U_\varphi, \varphi) \in C_p^\infty(U, N)$ is also an element of $C_p^\infty(M, N)$, since $U_\varphi \subseteq U \subseteq M$. Hence, there exists a canonical inclusion $C_p^\infty(U, N) \hookrightarrow C_p^\infty(M, N)$. Conversely, for every local map $(M_\varphi, \varphi) \in C_p^\infty(M, N)$ there exists a local map $(M_\varphi \cap U, \varphi|_{M_\varphi \cap U}) \in C_p^\infty(U, N)$, since U is open; restricting the domain to U thus defines an assignment $C_p^\infty(M, N) \to C_p^\infty(U, N)$. Recall from theorem 21.2.1 that the jet depends only on the germ. Clearly, (M_φ, φ) and $(M_\varphi \cap U, \varphi|_{M_\varphi \cap U})$ have the same germ at p, since they agree on the open set $M_\varphi \cap U$, and so they define the same jet. Hence, there exists a bijection between $J_p^r(U, N)$ and $J_p^r(M, N)$.

   To see that this bijection is a diffeomorphism, pick a chart of U containing p, and construct a chart of $J_p^r(U, N)$ following the procedure outlined in the proof of theorem 21.3.1. Keeping in mind that a chart of U is also a chart of $M \supseteq U$, one can construct a chart of $J_p^r(M, N)$ from the same chart. Then the bijection given above relates these two charts, and is thus a diffeomorphism.

2. Let $p \in M$ and consider a local map $(U_\varphi, \varphi) \in C_p^\infty(M, V)$ with $\varphi(p) = q$. Since $V \subset N$, this is also an element of $C_p^\infty(M, N)$, and so there is a canonical inclusion $C_p^\infty(M, V) \hookrightarrow C_p^\infty(M, N)$. Conversely, every element $(U_\varphi, \varphi) \in C_p^\infty(M, N)$ can be restricted to $(U_\varphi \cap \varphi^{-1}(V), \varphi|_{U_\varphi \cap \varphi^{-1}(V)}) \in C_p^\infty(M, V)$, since V and hence $\varphi^{-1}(V)$ is open. One then proceeds as in the proof of the first proposition, using the fact that restrictions have the same germ and therefore define the same jet.

3. The proof proceeds again as for the previous cases. Let $(U_\varphi, \varphi) \in C_p^\infty(U, V)$ with $\varphi(p) = q$ be a local map. This map is also en element of $C_p^\infty(M, N)$. Conversely, every element $(U_\varphi, \varphi) \in C_p^\infty(M, N)$ can be restricted to $(U_\varphi \cap U \cap \varphi^{-1}(V), \varphi|_{U_\varphi \cap U \cap \varphi^{-1}(V)}) \in C_p^\infty(U, V)$. The remainder of the proof is analogous as before. ■

Recall that locally every manifold is diffeomorphic to an open subset of Euclidean space. Since the jets depends only on this local structure, one may thus further conclude that once we fix a point $p \in M$ or $q \in N$, or both, the corresponding jet space does not depend on the containing manifold at all, except for its dimension, or on the particular choice of the point, up to diffeomorphism. This will be shown next.

Theorem 21.3.6. *Let M, N be manifolds with $m = \dim M$ and $n = \dim N$. Then the following manifolds are diffeomorphic:*

1. $J_p^r(M, N) \cong J_0^r(\mathbb{R}^m, N)$ *for all* $p \in M$,
2. $J^r(M, N)_q \cong J^r(M, \mathbb{R}^n)_0$ *for all* $q \in N$,
3. $J_p^r(M, N)_q \cong J_0^r(\mathbb{R}^m, \mathbb{R}^n)_0$ *for all* $p \in M$ *and* $q \in N$.

Proof. 1. Let (U,ϕ) be a chart of M with $p \in U$ and $\phi(p) = 0 \in \mathbb{R}^m$. By theorem 21.3.5, we have $J^r_p(M,N) \cong J^r_p(U,N)$, and so we will work with the latter instead of the former. For any local map $(U_\varphi,\varphi) \in C^\infty_p(U,N)$ there exists a local map $(\phi(U_\varphi),\varphi \circ \phi^{-1}) \in C^\infty_0(\phi(U),N)$, and vice versa. This establishes a bijection between $C^\infty_p(U,N)$ and $C^\infty_0(\phi(U),N)$. Now recall that two local maps $(U_\varphi,\varphi),(U_\vartheta,\vartheta) \in C^\infty_p(U,N)$ define the same r-jet at p if and only if for all local curves $(W,\gamma) \in C^\infty_0(\mathbb{R},U)$ with $\gamma(0) = p$ and all functions $f \in C^\infty(N,\mathbb{R})$ holds

$$(W, f \circ \varphi \circ \gamma) \overset{0,r}{\sim} (W, f \circ \vartheta \circ \gamma)\,. \tag{21.3.14}$$

Using the fact that for each such local curve there exists a local curve $(W,\phi \circ \gamma) \in C^\infty_0(\mathbb{R},\phi(U))$ and vice versa, as well as

$$f \circ \varphi \circ \gamma = f \circ \varphi \circ \phi^{-1} \circ \phi \circ \gamma\,, \quad f \circ \vartheta \circ \gamma = f \circ \vartheta \circ \phi^{-1} \circ \phi \circ \gamma\,, \tag{21.3.15}$$

we see that

$$(U_\varphi,\varphi) \overset{p,r}{\sim} (U_\vartheta,\vartheta) \tag{21.3.16}$$

if and only if

$$(\phi(U_\varphi),\varphi \circ \phi^{-1}) \overset{p,r}{\sim} (\phi(U_\vartheta),\vartheta \circ \phi^{-1})\,. \tag{21.3.17}$$

Hence, we have a bijection between $J^r_p(U,N)$ and $J^r_0(\phi(U),N)$. This bijection is a diffeomorphism, since it relates the charts constructed from (U,ϕ) and $(\phi(U),\mathrm{id}_{\mathbb{R}^m})$ on the respective jet bundles. Finally, using again theorem 21.3.5 we have $J^r_0(\phi(U),N) \cong J^r_0(\mathbb{R}^m,N)$, and thus $J^r_p(M,N) \cong J^r_0(\mathbb{R}^m,N)$.

2. Let (V,χ) be a chart of N with $q \in V$ and $\chi(q) = 0 \in \mathbb{R}^n$. By theorem 21.3.5, we have $J^r(M,N)_q \cong J^r_q(M,V)$, and so we will work with the latter instead of the former. For any local map $(U_\varphi,\varphi) \in C^\infty_p(M,V)$, where $p \in M$ and $\varphi(p) = q$, there exists a local map $(U_\varphi,\chi \circ \varphi) \in C^\infty_p(M,\chi(V))$, and vice versa. This establishes a bijection between $C^\infty_p(M,V)$ and $C^\infty_p(M,\chi(V))$. Now recall that two local maps $(U_\varphi,\varphi),(U_\vartheta,\vartheta) \in C^\infty_p(M,V)$ with $\varphi(p) = \vartheta(p) = q$ define the same r-jet at p if and only if for all local curves $(W,\gamma) \in C^\infty_0(\mathbb{R},M)$ with $\gamma(0) = p$ and all functions $f \in C^\infty(V,\mathbb{R})$ holds

$$(W, f \circ \varphi \circ \gamma) \overset{0,r}{\sim} (W, f \circ \vartheta \circ \gamma)\,. \tag{21.3.18}$$

Using the fact that for each such function there exists a function $f \circ \chi^{-1} \in C^\infty(\chi(V),\mathbb{R})$ and vice versa, as well as

$$f \circ \varphi \circ \gamma = f \circ \chi^{-1} \circ \chi \circ \varphi \circ \gamma\,, \quad f \circ \vartheta \circ \gamma = f \circ \chi^{-1} \circ \chi \circ \vartheta \circ \gamma\,, \tag{21.3.19}$$

we see that

$$(U_\varphi,\varphi) \overset{p,r}{\sim} (U_\vartheta,\vartheta) \tag{21.3.20}$$

if and only if

$$(U_\varphi,\chi \circ \varphi) \overset{p,r}{\sim} (U_\vartheta,\chi \circ \vartheta)\,. \tag{21.3.21}$$

Hence, we have a bijection between $J^r(M,V)_q$ and $J^r(M,\chi(V))_0$. This bijection is a diffeomorphism, since it relates the charts constructed from (V,χ) and $(\chi(V),\mathrm{id}_{\mathbb{R}^n})$ on the respective jet bundles. Finally, using again theorem 21.3.5 we have $J^r(M,\chi(V))_0 \cong J^r(M,\mathbb{R}^n)_0$, and thus $J^r(M,N)_q \cong J^r(M,\mathbb{R}^n)_q$.

3. For the third proposition, we can combine the proofs of the first two propositions. Let (U,ϕ) be a chart of M with $p \in U$ and $\phi(p) = 0 \in \mathbb{R}^m$, as well as (V,χ) a chart of N with $q \in V$ and $\chi(q) = 0 \in \mathbb{R}^n$. By theorem 21.3.5, we have $J^r_p(M,N)_q \cong J^r_p(U,V)_q$, and so we will work with the latter instead of the former. For any local map $(U_\varphi,\varphi) \in C^\infty_p(U,V)$ there exists a local map $(\phi(U_\varphi),\chi \circ \varphi \circ \phi^{-1}) \in C^\infty_0(\phi(U),\chi(V))$, and vice versa. This establishes a bijection between $C^\infty_p(U,V)$ and $C^\infty_0(\phi(U),\chi(V))$. Now recall that two local

maps $(U_\varphi, \varphi), (U_\vartheta, \vartheta) \in C^\infty_p(U, V)$ define the same r-jet at p if and only if for all local curves $(W, \gamma) \in C^\infty_0(\mathbb{R}, U)$ with $\gamma(0) = p$ and all functions $f \in C^\infty(V, \mathbb{R})$ holds

$$(W, f \circ \varphi \circ \gamma) \overset{0;r}{\sim} (W, f \circ \vartheta \circ \gamma)\,. \tag{21.3.22}$$

Using the fact that for each such local curve there exists a local curve $(W, \phi \circ \gamma) \in C^\infty_0(\mathbb{R}, \phi(U))$ and vice versa, a function $f \circ \chi^{-1} \in C^\infty(\chi(V), \mathbb{R})$ and vice versa, as well as

$$f \circ \varphi \circ \gamma = f \circ \chi^{-1} \circ \chi \circ \varphi \circ \phi^{-1} \circ \phi \circ \gamma\,, \quad f \circ \vartheta \circ \gamma = f \circ \chi^{-1} \circ \chi \circ \vartheta \circ \phi^{-1} \circ \phi \circ \gamma\,, \tag{21.3.23}$$

we see that

$$(U_\varphi, \varphi) \overset{p;r}{\sim} (U_\vartheta, \vartheta) \tag{21.3.24}$$

if and only if

$$(\phi(U_\varphi), \chi \circ \varphi \circ \phi^{-1}) \overset{p;r}{\sim} (\phi(U_\vartheta), \chi \circ \vartheta \circ \phi^{-1})\,. \tag{21.3.25}$$

Hence, we have a bijection between $J^r_p(U, V)_q$ and $J^r_0(\phi(U), \chi(V))_0$. This bijection is a diffeomorphism, since it relates the charts constructed from (U, ϕ) and (V, χ) as well as $(\phi(U), \mathrm{id}_{\mathbb{R}^m})$ and $(\chi(V), \mathrm{id}_{\mathbb{R}^n})$ on the respective jet bundles. Finally, using again theorem 21.3.5 we have $J^r_0(\phi(U), \chi(V))_0 \cong J^r_0(\mathbb{R}^m, \mathbb{R}^n)_0$, and thus $J^r_p(M, N)_0 \cong J^r_0(\mathbb{R}^m, \mathbb{R})_0$. ■

Note that the diffeomorphisms we found above depend on the choice of a chart of M, and so there is no canonical diffeomorphism. This is reminiscent of the structure of a fiber bundle, where each fiber is diffeomorphic to a given manifold, but the choice of the diffeomorphism is related to the local trivializations. We now come to the crucial result, showing that the aforementioned spaces are indeed fibers of a fiber bundle with total space $J^r(M, N)$.

Theorem 21.3.7. *Let M, N be manifolds with $m = \dim M$ and $n = \dim N$. Then the following tuples are fiber bundles:*

1. $(J^r(M, N), M, \overleftarrow{\pi}_r, J^r_0(\mathbb{R}^m, N))$, *where* $\overleftarrow{\pi}_r = \mathrm{pr}_1 \circ \pi_{r,0} : J^r(M, N) \to M, j^r_p\varphi \mapsto p$.
2. $(J^r(M, N), N, \overrightarrow{\pi}_r, J^r(M, \mathbb{R}^n)_0)$, *where* $\overrightarrow{\pi}_r = \mathrm{pr}_2 \circ \pi_{r,0} : J^r(M, N) \to N, j^r_p\varphi \mapsto \varphi(p)$.
3. $(J^r(M, N), M \times N, \overleftrightarrow{\pi}_r, J^r_0(\mathbb{R}^m, \mathbb{R}^n)_0)$, *where* $\overleftrightarrow{\pi}_r = \pi_{r,0} : J^r(M, N) \to J^0(M, N) \cong M \times N, j^r_p\varphi \mapsto (p, \varphi(p))$.

Proof. ▶. . . ◀ ■

It is important to note that these fiber bundles are in general not trivial. Further, also the following holds.

Theorem 21.3.8. *The triple $(J^r(M, N), J^k(M, N), \pi_{r,k})$, where M, N are manifolds and $k, r \in \mathbb{N}$ with $r > k$, is a smooth fiber bundle.*

Proof. ▶. . . ◀ ■

21.4 Pullback and pushforward of jets

From their definition, we know that jets are equivalence classes of (local) maps. Further, we know that the composition of (smooth) maps is again a (smooth) maps. One may thus wonder whether the operation of map composition is compatible with the structure of equivalence classes defined by jets. We start by showing the following.

Theorem 21.4.1. *Let M, M', N, N' be manifolds, $p \in M$, $q \in N$, $p' \in M'$, $q' \in N'$, $\varphi : M' \to M$ and $\vartheta : N \to N'$ maps such that $\varphi(p') = p$ and $\vartheta(q) = q'$ and $r \in \mathbb{N}$. If $(U, \psi), (U, \tilde{\psi}) \in C_p^\infty(M, N)$ are local maps with $(U, \psi) \overset{p,r}{\sim} (U, \tilde{\psi})$, then the following hold:*

1. $(U, \vartheta \circ \psi) \overset{p,r}{\sim} (U, \vartheta \circ \tilde{\psi})$,
2. $(\varphi^{-1}(U), \psi \circ \varphi) \overset{p',r}{\sim} (\varphi^{-1}(U), \tilde{\psi} \circ \varphi)$,
3. $(\varphi^{-1}(U), \vartheta \circ \psi \circ \varphi) \overset{p',r}{\sim} (\varphi^{-1}(U), \vartheta \circ \tilde{\psi} \circ \varphi)$.

Proof. ▶ . . . ◀ ■

We thus see that composition with maps preserves the equivalence class structure defined by jets. Recalling that jets form the elements of a manifold, one may thus construct the following.

Theorem 21.4.2. *Let M, M', N, N' be manifolds, $p \in M$, $p' \in M'$,, $\varphi : M' \to M$ and $\vartheta : N \to N'$ maps such that $\varphi(p') = p$ and $r \in \mathbb{N}$. Then there exist smooth maps defined by the pushforward*

$$\begin{array}{cccc} \vartheta_* & : & J^r(M, N) & \to & J^r(M, N') \\ & & j_p^r \psi & \mapsto & j_p^r(\vartheta \circ \psi) \end{array} \tag{21.4.1}$$

and the pullback

$$\begin{array}{cccc} \varphi^* & : & J_p^r(M, N) & \to & J_{p'}^r(M', N) \\ & & j_p^r \psi & \mapsto & j_{p'}^r(\psi \circ \varphi) \end{array} \tag{21.4.2}$$

such that the diagram

$$\begin{array}{ccc} J_p^r(M, N) & \xrightarrow{\vartheta_*} & J_p^r(M, N') \\ \downarrow{\scriptstyle \varphi^*} & & \downarrow{\scriptstyle \varphi^*} \\ J_{p'}^r(M', N) & \xrightarrow[\vartheta_*]{} & J_{p'}^r(M', N') \end{array} \tag{21.4.3}$$

commutes.

Proof. ▶ . . . ◀ ■

It is important to note that we have fixed the source p, p' of the jets appearing in the manifolds above for the pullback φ^*. This is due to the fact that φ may not be surjective, so that one can only pull back jets at points $p \in M$ which lie in the image of φ and not all of $J^r(M, N)$. Further, φ may not be injective, and so one needs to specify the point $p' \in M'$ as well. This restriction does not apply for the pushforward ϑ_*, since the map ϑ possesses a unique image for each $q \in N$, so that every jet in $J^r(M, N)$ can uniquely be pushed to $J^r(M, N)$.

Note that the jet manifolds given above form the total spaces of several fiber bundles, as we have seen in theorems 21.3.7 and 21.3.8. Having a map ϑ_* between them, the question arises whether this map preserves the fiber bundle structure, and thus defines fiber bundle morphisms, and if this is the case, which maps they cover. This can be shown as follows.

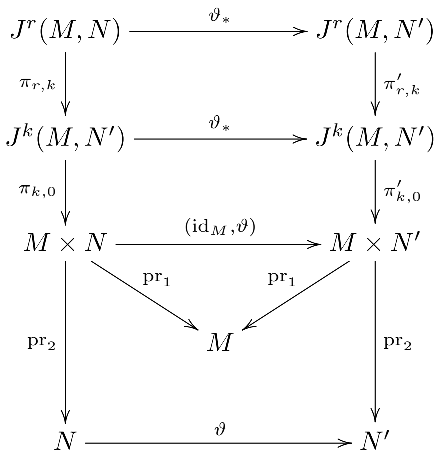
Theorem 21.4.3. *For each of the bundles given in theorems 21.3.7 and 21.3.8, the pushforward $\vartheta_* : J^r(M, N) \to J^r(M, N')$ along a map $\vartheta : N \to N'$ is a bundle morphism, covering the following maps:*

1. *from $(J^r(M, N), J^k(M, N), \pi_{r,k})$ to $(J^r(M, N'), J^k(M, N'), \pi'_{r,k})$ covering $\vartheta_* : J^k(M, N) \to J^k(M, N')$,*
2. *from $(J^r(M, N), M, \overleftarrow{\pi}_r)$ to $(J^r(M, N'), M, \overleftarrow{\pi}'_r)$ covering $\mathrm{id}_M : M \to M$,*

3. *from* $(J^r(M,N), N, \overrightarrow{\pi}_r)$ *to* $(J^r(M,N'), N', \overrightarrow{\pi}'_r)$ *covering* $\vartheta : N \to N'$,

4. *from* $(J^r(M,N), M \times N, \overleftrightarrow{\pi}_r)$ *to* $(J^r(M,N'), M \times N', \overleftrightarrow{\pi}'_r)$ *covering* $(\mathrm{id}_M, \vartheta) : M \times N \to M \times N'$.

Proof. Due to the fact that all bundle maps of the fiber bundles above are defined through the bundle maps $\pi_{r,k}$ and $\pi'_{r,k}$ as well as the projections of $M \times N$ and $M \times N'$ onto their respective factors, it is sufficient to show that the following diagram commutes:



(21.4.4)

▶...◀ ■

21.5 Jet groups

In the following section we will show how the notion of jets leads to an interesting class of Lie groups, which play an important role in the construction of certain bundles, which allow lifts of diffeomorphisms, and which generalize the notion of pullback we discussed in section 11.3. We will encounter such bundles in section 22.5. Here we mostly follow [KSM93].

In order to construct a group of jets, we first need to define a composition of jets. The crucial observation that allows this construction is the following.

Theorem 21.5.1. *Let* M, N, O *be manifolds,* $p \in M$, $q \in N$ *and* $r \in \mathbb{N}$. *If two pairs* $(U, \varphi), (\tilde{U}, \tilde{\varphi}) \in C^\infty_p(M,N)$ *and* $(V, \vartheta), (\tilde{V}, \tilde{\vartheta}) \in C^\infty_q(N,O)$ *with* $q = \varphi(p) = \tilde{\varphi}(p)$ *satisfy*

$$(U, \varphi) \overset{p,r}{\sim} (\tilde{U}, \tilde{\varphi}), \quad (V, \vartheta) \overset{q,r}{\sim} (\tilde{V}, \tilde{\vartheta}), \tag{21.5.1}$$

then

$$(U \cap \varphi^{-1}(V), \vartheta \circ \varphi|_{U \cap \varphi^{-1}(V)}) \overset{p,r}{\sim} (\tilde{U} \cap \tilde{\varphi}^{-1}(\tilde{V}), \tilde{\vartheta} \circ \tilde{\varphi}_{\tilde{U} \cap \tilde{\varphi}^{-1}(\tilde{V})}). \tag{21.5.2}$$

Proof. Let (W, χ) be a chart of N with $q \in W$. Since φ and $\tilde{\varphi}$ are smooth and thus continuous, $\varphi^{-1}(W)$ and $\tilde{\varphi}^{-1}$ are open, and we can consider the local maps $(U \cap \varphi^{-1}(W), \varphi|_{U \cap \varphi^{-1}(W)}) \in C^\infty_p(M,N)$ and $(\tilde{U} \cap \tilde{\varphi}^{-1}(W), \tilde{\varphi}|_{\tilde{U} \cap \tilde{\varphi}^{-1}(W)}) \in C^\infty_p(M,N)$. Since the latter are defined by restriction, they have the same germ as the former, so that

$$(U, \varphi) \overset{p,r}{\sim} (U \cap \varphi^{-1}(W), \varphi|_{U \cap \varphi^{-1}(W)}), \quad (\tilde{U}, \tilde{\varphi}) \overset{p,r}{\sim} (\tilde{U} \cap \tilde{\varphi}^{-1}(W), \tilde{\varphi}|_{\tilde{U} \cap \tilde{\varphi}^{-1}(W)}) \tag{21.5.3}$$

by theorem 21.2.1, and thus also

$$(U \cap \varphi^{-1}(W), \varphi|_{U \cap \varphi^{-1}(W)}) \overset{p,r}{\sim} (\tilde{U} \cap \tilde{\varphi}^{-1}(W), \tilde{\varphi}|_{\tilde{U} \cap \tilde{\varphi}^{-1}(W)}). \tag{21.5.4}$$

Similarly, for $(V \cap W, \vartheta|_{V \cap W}) \in C^\infty_q(N, O)$ and $(\tilde{V} \cap W, \tilde{\vartheta}_{\tilde{V} \cap W}) \in C^\infty_q(N, O)$ we have

$$(V, \vartheta) \overset{p,r}{\sim} (V \cap W, \vartheta|_{V \cap W}), \quad (\tilde{V}, \tilde{\vartheta}) \overset{p,r}{\sim} (\tilde{V} \cap W, \tilde{\vartheta}_{\tilde{V} \cap W}) \tag{21.5.5}$$

by theorem 21.2.1, and thus also

$$(V \cap W, \vartheta|_{V \cap W}) \overset{p,r}{\sim} (\tilde{V} \cap W, \tilde{\vartheta}_{\tilde{V} \cap W}). \tag{21.5.6}$$

Now consider a local curve

$$\gamma \in C^\infty_0(\mathbb{R}, U \cap \tilde{U} \cap \varphi^{-1}(V) \cap \tilde{\varphi}^{-1}(\tilde{V}) \cap \varphi^{-1}(W) \cap \tilde{\varphi}^{-1}(W)), \tag{21.5.7}$$

as well as a function $f \in C^\infty(O)$. Then we have

$$\frac{\mathrm{d}^r}{\mathrm{d}t^r}(f \circ \vartheta \circ \varphi \circ \gamma)(t)\bigg|_{t=0} = \frac{\mathrm{d}^r}{\mathrm{d}t^r}(f \circ \vartheta \circ \chi^{-1} \circ \chi \circ \varphi \circ \gamma)(t)\bigg|_{t=0}, \tag{21.5.8a}$$

$$\frac{\mathrm{d}^r}{\mathrm{d}t^r}(f \circ \tilde{\vartheta} \circ \tilde{\varphi} \circ \gamma)(t)\bigg|_{t=0} = \frac{\mathrm{d}^r}{\mathrm{d}t^r}(f \circ \tilde{\vartheta} \circ \chi^{-1} \circ \chi \circ \tilde{\varphi} \circ \gamma)(t)\bigg|_{t=0}, \tag{21.5.8b}$$

where we have omitted the restrictions of the domains of the appearing functions, since these are only necessary for the definition of γ to guarantee that it lies entirely in the domain of all appearing functions, but do not change the value of the derivative given above. We see that on the right hand side appear compositions of local functions $g \in C^\infty_0(\mathbb{R}, \mathbb{R}^n)$ and $h \in C^\infty_{\chi(q)}(\mathbb{R}^n, \mathbb{R})$, where $n = \dim N$, and for these we can write the derivative as

$$\frac{\mathrm{d}^r}{\mathrm{d}t^r}(h \circ g)(t)\bigg|_{t=0} = \blacktriangleright \ldots \blacktriangleleft \tag{21.5.9}$$

Note that these depend only on derivatives of g and h up to order r, and thus their r-jets. As shown in theorem 21.3.6, from the assumption that φ and $\tilde{\varphi}$ define the same r-jet at p, the same follows also for $\chi \circ \varphi$ and $\chi \circ \tilde{\varphi}$. Analogously, one concludes from ϑ and $\tilde{\vartheta}$ to $\vartheta \circ \chi^{-1}$ and $\tilde{\vartheta} \circ \chi^{-1}$. Hence, also their compositions define the same r-jet. ∎

We see that if φ and $\tilde{\varphi}$, as well as ϑ and $\tilde{\vartheta}$, have r-contact, then also their compositions $\vartheta \circ \varphi$ and $\tilde{\vartheta} \circ \tilde{\varphi}$ have r-contact. In other words, the jets $j^r_p(\vartheta \circ \varphi)$ and $j^r_p(\tilde{\vartheta} \circ \tilde{\varphi})$ agree if $j^r_p\varphi = j^r_p\tilde{\varphi}$ and $j^r_q\vartheta = j^r_q\tilde{\vartheta}$. Hence, $j^r_p(\vartheta \circ \varphi)$ depends only on the jets $j^r_p\varphi$ and $j^r_q\vartheta$, but not on the particular choice of representatives φ and ϑ. Therefore, the following definition is valid.

Definition 21.5.1 (Composition of jets)**.** Let M, N, O be manifolds, $u \in M$, $v \in N$, $w \in O$ and $r \in \mathbb{N}$. For jets $j^r_u\varphi \in J^r_u(M, N)_v$ and $j^r_v\vartheta \in J^r_v(N, O)_w$ we define the *composition* $j^r_v\vartheta \circ j^r_u\varphi = j^r_u(\vartheta \circ \varphi)$.

In order to define a group structure, we need to check that this operation is associative, which we will do next.

Theorem 21.5.2. *For $r \in \mathbb{N}$, manifolds M, N, O, P, points $u \in M$, $v \in N$, $w \in O$, $x \in P$ and jets $j^r_u\varphi \in J^r_u(M, N)_v$, $j^r_v\vartheta \in J^r_v(N, O)_w$ and $j^r_w\psi \in J^r_w(O, P)_x$ holds:*

$$(j^r_w\psi \circ j^r_v\vartheta) \circ j^r_u\varphi = j^r_w\psi \circ (j^r_v\vartheta \circ j^r_u\varphi). \tag{21.5.10}$$

Proof. By definition, one has

$$\begin{aligned}(j^r_w\psi \circ j^r_v\vartheta) \circ j^r_u\varphi &= j^r_v(\psi \circ \vartheta) \circ j^r_u\varphi \\ &= j^r_u((\psi \circ \vartheta) \circ \varphi) \\ &= j^r_u(\psi \circ (\vartheta \circ \varphi)) \\ &= j^r_w\psi \circ j^r_u(\vartheta \circ \varphi) \\ &= j^r_w\psi \circ (j^r_v\vartheta \circ j^r_u\varphi)\end{aligned} \tag{21.5.11}$$

using the associativity of map composition. ∎

Further, we need a neutral element. Since jets are defined via map composition, and the neutral element of map composition is the identity map, the following result is straightforward.

Theorem 21.5.3. *Let $r \in \mathbb{N}$ and M, N be manifolds with $p \in M$ and $q \in N$. For any jet $j_p^r\varphi \in J_p^r(M, N)_q$ holds*

$$j_q^r\mathrm{id}_N \circ j_p^r\varphi = j_p^r\varphi = j_p^r\varphi \circ j_p^r\mathrm{id}_M\,. \tag{21.5.12}$$

Proof. From the definition immediately follows

$$j_q^r\mathrm{id}_N \circ j_p^r\varphi = j_p^r(\mathrm{id}_N \circ \varphi) = j_p^r\varphi \tag{21.5.13}$$

and

$$j_p^r\varphi \circ j_p^r\mathrm{id}_M = j_p^r(\varphi \circ \mathrm{id}_M) = j_p^r\varphi\,. \tag{21.5.14}$$

■

Finally, we also need the notion of an inverse. It turns out that for the composition of jets, not every element possesses an inverse, and so we will have to restrict the jets we consider to those which possess an inverse. We define them as follows.

Definition 21.5.2 (Invertible jet). Let M, N be manifolds, $p \in M$, $q \in N$ and $r \in \mathbb{N}$. A jet $j_p^r\varphi \in J_p^r(M, N)_q$ is called *invertible* if there exists a jet $j_q^r\tilde{\varphi} \in J_q^r(N, M)_p$ such that $j_q^r\tilde{\varphi} \circ j_p^r\varphi = j_p^r(\mathrm{id}_M)$ and $j_p^r\varphi \circ j_q^r\tilde{\varphi} = j_q^r(\mathrm{id}_N)$. The space of all invertible elements in $J^r(M, N)$ is denoted $\tilde{J}^r(M, N)$.

In the following, we will determine a few criteria for jets to be invertible. The most simple case is trivial.

Theorem 21.5.4. *For any manifolds M, N, every zeroth order jet $j_p^0\varphi \in J^0(M, N)$ is invertible.*

Proof. Recall that a zeroth-order jet $j_p^0\varphi \in J^0(M, N) \cong M \times N$ is simply given by the pair $(p, q) = (p, \varphi(p)) \in M \times N$. Let $\tilde{\varphi} : N \to M, x \mapsto p$ be the constant map. Then $\tilde{\varphi}(\varphi(p)) = p = \mathrm{id}_M(p)$, and thus $j_q^0\tilde{\varphi} \circ j_p^0\varphi = j_p^0\mathrm{id}_M$, as well as $\varphi(\tilde{\varphi}(q)) = q = \mathrm{id}_N(q)$, and so $j_p^0\varphi \circ j_q^0\tilde{\varphi} = j_q^0\mathrm{id}_N$. ■

We now come to more interesting cases. Recalling that jets essentially generalize partial derivatives, we can use the properties of the derivatives of inverse functions to show the following.

Theorem 21.5.5. *For any manifolds M, N, every first order jet $j_p^1\varphi \in J^1(M, N)$ is invertible if and only if the rank of φ at p equals $\dim M = \dim N$, i.e., if and only if $\varphi_*|_{T_pM} : T_pM \to T_{\varphi(p)}N$ is bijective.*

Proof. Let $q = \varphi(p) \in N$. If $j_p^1\varphi$ is invertible, then there exists $\tilde{\varphi} \in C_q^\infty(N, M)$ such that $j_q^1\tilde{\varphi} \circ j_p^1\varphi = j_p^1(\mathrm{id}_M)$ and $j_p^1\varphi \circ j_q^1\tilde{\varphi} = j_q^1(\mathrm{id}_N)$. This is equivalent to stating that $\tilde{\varphi}(q) = p$ and that for all local curves $\gamma \in C_0^\infty(\mathbb{R}, M)$ with $\gamma(0) = p$ and functions $f \in C^\infty(N, \mathbb{R})$ holds

$$\begin{aligned}(\mathrm{d}f \circ \dot{\gamma})(0) &= (f \circ \gamma)'(0)\\ &= (f \circ \mathrm{id}_M \circ \gamma)'(0)\\ &= (f \circ \tilde{\varphi} \circ \varphi \circ \gamma)'(0)\\ &= (\mathrm{d}f \circ \tilde{\varphi}_* \circ \varphi_* \circ \dot{\gamma})(0)\,,\end{aligned} \tag{21.5.15}$$

while for all local curves $\tilde{\gamma} \in C_0^\infty(\mathbb{R}, N)$ with $\tilde{\gamma}(0) = q$ and functions $\tilde{f} \in C^\infty(M, \mathbb{R})$ holds

$$\begin{aligned}(\mathrm{d}\tilde{f} \circ \dot{\tilde{\gamma}})(0) &= (\tilde{f} \circ \tilde{\gamma})'(0) \\ &= (\tilde{f} \circ \mathrm{id}_N \circ \tilde{\gamma})'(0) \\ &= (\tilde{f} \circ \tilde{\varphi} \circ \varphi \circ \tilde{\gamma})'(0) \\ &= (\mathrm{d}\tilde{f} \circ \tilde{\varphi}_* \circ \varphi_* \circ \dot{\tilde{\gamma}})(0) .\end{aligned} \tag{21.5.16}$$

Since the tangent vectors $\dot{\gamma}(0)$ of all such curves span T_pM, while the cotangent vectors $\mathrm{d}f(p)$ span T_p^*M, the former is the case if and only if $\tilde{\varphi}_* \circ \varphi_*|_{T_pM} = \mathrm{id}_{T_pM}$. Likewise, the latter is the case if and only if $\varphi_* \circ \tilde{\varphi}_*|_{T_pN} = \mathrm{id}_{T_pN}$. Hence, $\tilde{\varphi}_*|_{T_pN}$ is the inverse of $\varphi_*|_{T_pM}$, and so the latter is invertible.

To show the converse, assume that $\varphi_*|_{T_pM}$ is invertible. To construct $\tilde{\varphi}$, consider charts (U, ϕ) of M with $p \in U$ and (V, χ) of N with $q \in V$. From the assumption follows that the Jacobian $J = \mathrm{D}(\chi \circ \varphi \circ \phi^{-1})(\phi(p))$ is an invertible matrix. Consider a local map $\tilde{\varphi}$ defined by

$$\tilde{\varphi}(q') = \phi^{-1}(\phi(p) + J^{-1}(\chi(q') - \chi(q))) , \tag{21.5.17}$$

defined for all $q' \in V$ with $\phi(p) + J^{-1}(\chi(q') - \chi(q)) \in \phi(U)$. Clearly, its domain contains q, and $\tilde{\varphi}(q) = p$. Further, for all $v \in T_pM$ holds

$$\begin{aligned}(\tilde{\varphi}_* \circ \varphi_*)(v) &= (\tilde{\varphi} \circ \varphi)_*(v) \\ &= (\phi^{-1} \circ \phi \circ \tilde{\varphi} \circ \chi^{-1} \circ \chi \circ \varphi \circ \phi^{-1} \circ \phi)_*(v) \\ &= \phi_*^{-1}(\mathrm{D}(\phi \circ \tilde{\varphi} \circ \chi^{-1}) \cdot \mathrm{D}(\chi \circ \varphi \circ \phi^{-1}) \cdot \phi_*(v)) \\ &= \phi_*^{-1}(J^{-1} \cdot J \cdot \phi_*(v)) \\ &= (\phi_*^{-1} \circ \phi_*)(v) \\ &= (\phi^{-1} \circ \phi)_*(v) \\ &= v ,\end{aligned} \tag{21.5.18}$$

where we used the fact that the linear map $\phi \circ \tilde{\varphi} \circ \chi^{-1}$ satisfies

$$(\phi \circ \tilde{\varphi} \circ \chi^{-1})(y) = \phi(p) + J^{-1}(y - \chi(q)) , \tag{21.5.19}$$

and so its Jacobian is J^{-1}. Analogously, one shows that $(\varphi_* \circ \tilde{\varphi}_*)(w) = w$ for all $w \in T_qN$. It thus follows that $j_q^1\tilde{\varphi} \circ j_p^1\varphi = j_p^1(\mathrm{id}_M)$ and $j_p^1\varphi \circ j_q^1\tilde{\varphi} = j_q^1(\mathrm{id}_N)$, and so $j_p^r\varphi$ is invertible. ∎

The appearance of the Jacobian signals an important result.

Theorem 21.5.6. *For any manifolds M, N, every jet $j_p^r\varphi \in J^r(M, N)$ with $r > 1$ is invertible if and only if $j_p^1\varphi \in J^1(M, N)$ is invertible.*

Proof. Consider the same charts (U, ϕ) of M with $p \in U$ and (V, χ) of N with $q = \varphi(p) \in V$ as in the proof of theorem 21.5.5. If $j_p^1\varphi$ is invertible, then the Jacobian $\mathrm{D}(\chi \circ \varphi \circ \phi^{-1})(\phi(p))$ is non-degenerate. By the inverse function theorem, then there exists an open set $W \subset \chi(V)$, with open preimages under χ, φ and ϕ^{-1}, such that $\chi \circ \varphi \circ \phi^{-1}$ possesses an inverse defined on W. This can be written as $\phi \circ \tilde{\varphi} \circ \chi^{-1}$ by suitable composing with the chart functions, because ϕ and χ are bijections on the relevant subsets generated by W. Then $\tilde{\varphi}$ is a local inverse of φ, with

$$\tilde{\varphi} \circ \varphi = \mathrm{id}_M|_{\varphi^{-1}(\chi^{-1}(W))} , \quad \varphi \circ \tilde{\varphi} = \mathrm{id}_N|_{\chi^{-1}(W)} . \tag{21.5.20}$$

Hence, its jet $j_q^r\tilde{\varphi}$ is an inverse of $j_p^r\varphi$.

Conversely, if we know that an inverse $j_q^r\tilde{\varphi}$ exists, we can use the jet projection $\pi_{r,1}$ to obtain an inverse $j_q^1\tilde{\varphi}$ of $j_p^1\varphi$. ∎

So far, we have only been discussing the existence of an inverse jet, but we have not yet excluded the possibility that there exists several inverses. This will be discussed next.

Theorem 21.5.7. *If a jet $j^r_p\varphi \in J^r(M,N)$ is invertible, then its inverse is unique.*

Proof. If both $j^r_{\varphi(p)}\tilde{\varphi}$ and $j^r_{\varphi(p)}\tilde{\varphi}'$ are inverses of $j^r_p\varphi$, then we find

$$\begin{aligned} j^r_{\varphi(p)}\tilde{\varphi} &= j^r_{\varphi(p)}(\tilde{\varphi} \circ \mathrm{id}_N) \\ &= j^r_{\varphi(p)}\tilde{\varphi} \circ j^r_{\varphi(p)}\mathrm{id}_N \\ &= j^r_{\varphi(p)}\tilde{\varphi} \circ j^r_p\varphi \circ j^r_{\varphi(p)}\tilde{\varphi}' \\ &= j^r_p\mathrm{id}_M \circ j^r_{\varphi(p)}\tilde{\varphi}' \\ &= j^r_{\varphi(p)}(\mathrm{id}_M \circ \tilde{\varphi}') \\ &= j^r_{\varphi(p)}\tilde{\varphi}' \,. \end{aligned} \tag{21.5.21}$$ ■

With these results at hand, we can finally come to the central definition of this section. In order to form a group G, we must make sure that *any* two elements of G can be composed. Since jet composition is defined only if the source of one jet agrees with the target of another, we must therefore restrict ourselves to the invertible elements of a jet space $\tilde{J}^r_p(M,M)_p$, where all jets have the same source and target $p \in M$. From theorem 21.3.6 we know that, up to diffeomorphism, this space neither depends on the choice of the manifold M, nor the point p, except for the dimension of the former. Hence, we can make a canonical choice and define the following.

Definition 21.5.3 (Jet group). For $r, n \in \mathbb{N}$, the *jet group* is defined as $\mathrm{J}^r(n) = \tilde{J}^r_0(\mathbb{R}^n,\mathbb{R}^n)_0$.

From its construction, it becomes clear that $\mathrm{J}^r(n)$ is a group. Further, recall that $J^r_0(\mathbb{R}^n,\mathbb{R}^n)_0$ is a manifold, and one may expect the same to be true for $\tilde{J}^r_0(\mathbb{R}^n,\mathbb{R}^n)_0$. Given these two structures, the question arises whether they are compatible. This will be answered next.

Theorem 21.5.8. *The jet group $\mathrm{J}^r(n)$ is a Lie group of dimension $\dim \mathrm{J}^r(n) = n\left(\binom{n+r}{n} - 1\right)$.*

Proof. We show the smoothness of the group operations of multiplication and inverse by constructing a suitable chart. Note that $\mathbb{R}^n$ is canonically equipped with Cartesian coordinates (x^a). In these coordinate, an invertible r-jet $g \in \mathrm{J}^r(n)$ can uniquely be expressed through the Taylor polynomial

$$g(x) = \sum_{k=1}^{r} g^a{}_{b_1\cdots b_k} x^{b_1} \cdots x^{b_k} \,, \tag{21.5.22}$$

where the coefficients satisfy

$$g^a{}_{b_1\cdots b_k} = g^a{}_{(b_1\cdots b_k)} \,, \tag{21.5.23}$$

and $g^a{}_b$ are the components of an invertible matrix. Note that this polynomial, being an element of $C^\infty(\mathbb{R}^n,\mathbb{R}^n$, is also a canonical representative of the jet. The jet composition, which constitutes the group structure, is thus given by the composition

$$g(h(x)) = \sum_{k=1}^{r} \sum_{l_1,\ldots,l_k} g^a{}_{b_1\cdots b_k} h^{b_1}{}_{c_{1,1}\cdots c_{1,l_1}} \cdots h^{b_k}{}_{c_{k,1}\cdots c_{k,l_k}} x^{c_{1,1}} \cdots x^{c_{k,l_k}} \,, \tag{21.5.24}$$

where the sum runs over such values that the total degree of the polynomial is at most r. This is obviously smooth in the components of g and h. For the inverse, one has

$$g^{-1}(x) = \blacktriangleright \ldots \blacktriangleleft \,, \tag{21.5.25}$$

where $\tilde{g}^a{}_b$ is the inverse of $g^a{}_b$. This is also smooth, since the inverse is smooth in $\mathrm{GL}(n,\mathbb{R})$. Hence, $\mathrm{J}^r(n)$ is a Lie group. ■

It is instructive to take a look at the most simple example.

Theorem 21.5.9. *The jet group* $\mathrm{J}^1(n)$ *is isomorphic to the general linear group* $\mathrm{GL}(n, \mathbb{R})$.

Proof. This follows from the fact that an invertible 1-jet, following theorem 21.5.8, is uniquely determined by an invertible linear function $g : \mathbb{R}^n \to \mathbb{R}^n$, and thus an element of $\mathrm{GL}(n, \mathbb{R})$. ■

21.6 Jet bundles

We have already learned that a particularly useful class of maps are sections of fiber bundles, and that many useful objects such as vector or tensor fields fall into this category. We will now study the jets of these maps, using the conventions in [Sau89]. Since jets only depend on the *local* behavior of a map, i.e., its germ as discussed in section 21.2, we make use of *local* sections here, as defined in definition 2.3.2 - similarly to the local maps used in definition 21.1.8. Another reason, which makes it even necessary to use local instead of global sections, is the fact that there are fiber bundles which do not have any global sections, but are still interesting objects for constructing jet bundles. Common examples, which we will encounter later, are principal fiber bundles, and in particular frame bundles. Hence, we define jets of local sections as follows.

Definition 21.6.1 (Jets of local sections)**.** Let $\pi : E \to M$ be a fiber bundle, $p \in M$ and $\Gamma_p(E)$ the space of all local sections whose domain contains p. For $r \in \mathbb{N}$ and a local section $\sigma \in \Gamma_p(E)$ with domain U_σ we define the *r-jet* $j^r_p\sigma$ of σ at p as the equivalence class

$$j^r_p\sigma = \left\{\tau \in \Gamma_p(E) \,\middle|\, (U_\sigma, \sigma) \overset{p,r}{\sim} (U_\tau, \tau)\right\} \tag{21.6.1}$$

of local sections $\tau \in \Gamma_p(E)$ with domain U_τ having r-contact at p. The space of all r-jets at p is denoted $J^r_p(E)$, while the space of all r-jets is denoted $J^r(E)$.

Note that local sections are in particular local maps, and so the notion of r-contact given in definition 21.1.7 can be applied. The main difference between this definition and the definition 21.1.8 is that we do not consider arbitrary maps from M to E in the construction of the equivalence classes, but only sections. This restriction also reduces the number of dimensions of the jet space, which we can state as follows.

Theorem 21.6.1. *Let* $\pi : E \to M$ *be a fiber bundle with fiber* F *and dimensions* $\dim M = m, \dim F = n$ *and* $r \in \mathbb{N}$. *For each* $p \in M$ *the space* $J^r_p(E)$ *is a manifold of dimension* $n\binom{m+r}{r}$, *while* $J^r(E)$ *is a manifold of dimension* $m + n\binom{m+r}{r}$.

Proof. ▶. . . ◀ ■

We see that instead of the dimension $\dim E$ of the target manifold we only have the dimension $\dim F$ which enters the formula of the dimension. To see why this is the case, we can construct coordinates in the same way as we did for the jet manifolds of arbitrary maps. By definition, every fiber bundle is locally trivial, i.e., for every $p \in M$ there exists an open set $U \subset M$ containing p such that $U \times F \cong \pi^{-1}(U) \subset E$. Given coordinates (x^α) on U and (y^a) on F we can thus use coordinates (x^α, y^a) on $\pi^{-1}(U)$. Let now $\sigma : U \to \pi^{-1}(U)$ be a local section, whose domain we also assume to be U. (If it had a different domain $U' \ni p$ instead, we could simply replace U by $U \cap U'$ in the remainder of this construction.) This section is described by assigning coordinates (x^α, y^a) of the target space to coordinates (x^α) of the domain. However, the first part of these target coordinates is already fixed by the condition that σ is a section, and thus $\pi \circ \sigma = \mathrm{id}_U$. Hence, σ is uniquely determined by the coordinate functions $y^a(x)$. In other words, a section σ looks *locally* just like a map from U to F. Using the coordinate functions $y^a(x)$ we can use the same construction as in the previous section to construct coordinates (y^a_Λ) on $J^r_p(E)$ and (x^α, y^a_Λ) on $J^r(E)$.

Now it is also easy to see the following.

Theorem 21.6.2. *Let $\pi : E \to M$ be a fiber bundle and $p \in M$. Then $J^0_p(E) \cong \pi^{-1}(p) \cong F$ and $J^0(E) \cong E$.*

Proof. Recall that a 0-jet $j^0_p\sigma$ of a local section σ is uniquely determined by the value $\sigma(p) \in \pi^{-1}(p) \cong F$, which proves the first statement. The second statement follows from the fact that $J^0(E)$ is simply the union of $J^r_p(E)$ for all $p \in M$, while E is the union of all $\pi^{-1}(p)$. One can easily show that the maps $J^0_p(E) \to F$ and $J^0(E) \to E$ derived from these identifications are diffeomorphisms. ■

Given now a number of jet manifolds, we may consider maps between them. A very useful class of maps is defined as follows.

Definition 21.6.2 (Jet projection)**.** Let $\pi : E \to M$ be a fiber bundle and $0 \leq k \leq r$. The *k-jet projection* is the map $\pi_{r,k} : J^r(E) \to J^k(E)$ which assigns to the r-jet $j^r_p\sigma$ of any local section σ its k-jet $j^k_p\sigma$ for every $p \in M$. The map $\pi_{r,0} : J^r(E) \to E$ is also called the *target projection*, while $\pi_r = \pi \circ \pi_{r,0} : J^r(E) \to M$ is called the *source projection*.

Of course we must check that the projections given above are indeed well-defined. This is the case, since any two local sections σ, τ which have the same r-jet also have the same k-jet for $k \leq r$, which follows immediately from the definition of jets. Therefore, the k-jet $j^k_p\sigma$ of a local section σ is uniquely determined by its r-jet $j^r_p\sigma$, as we presumed in the definition above. We will not prove here that the jet projections are smooth maps - the proof is lengthy, but simple. In coordinates (x^α, y^a_Λ) on $J^r(E)$ one can easily see that the projection $\pi_{r,k}$ simply discards all coordinates y^a_Λ with $|\Lambda| > k$ and keeps only the coordinates on $J^k(E)$. These maps have even more nice properties.

Theorem 21.6.3. *The triples $(J^r(E), J^k(E), \pi_{r,k})$ with $r > k$, $(J^r(E), E, \pi_{r,0})$ and $(J^r(E), M, \pi_r)$ are fiber bundles.*

Proof. Note that $E = J^0(E)$ is just a special case with $k = 0$, and so we will not discuss it separately here. For $p \in M$, consider a local trivialization (U, ϕ) with $p \in U \subset M$ and $\phi : \pi^{-1}(U) \to U \times F$. Now there is a one-to-one correspondence between local sections $\sigma \in \Gamma_{|U}(E)$ and maps $\sigma_\phi : U \to F$ given by

$$\sigma_\phi = \mathrm{pr}_2 \circ \phi \circ \sigma \quad \Leftrightarrow \quad \sigma = \phi^{-1} \circ (\mathrm{id}_U, \sigma_\phi)\,. \tag{21.6.2}$$

►...◄ ■

Some of the aforementioned bundles carry some additional structure, which deserves further discussion. We will start with the following bundle, which turns out to have the structure of an affine bundle.

Theorem 21.6.4. *For $r \in \mathbb{N}$, the bundle $(J^r(E), J^{r-1}(E), \pi_{r,r-1})$ is an affine bundle modeled over the vector bundle $\pi^*_{r-1,0}(VE) \otimes \pi^*_{r-1}(\mathrm{Sym}^r T^*M)$.*

Proof. ►...◄ ■

The last bundle of the list in theorem 21.6.3 will also be of particular interest for us, and has its own name.

Definition 21.6.3 (Jet bundle)**.** Let $\pi : E \to M$ be a fiber bundle and $r \in \mathbb{N}$. The bundle $(J^r(E), M, \pi_r)$ is called the *r'th jet bundle* of E.

We first study how these bundles are related amongst each other. It is not very surprising that the maps between them turn out to be bundle morphisms.

Theorem 21.6.5. *For $0 \leq k \leq r$, the map $\pi_{r,k} : J^r(E) \to J^k(E)$ is a bundle morphism from $(J^r(E), M, \pi_r)$ to $(J^k(E), M, \pi_k)$ covering the identity id_M on M.*

Proof. ▶. . . ◀ ■

$$\begin{array}{ccccccc} J^r(E) & \xrightarrow[\pi_{r,r-1}]{} & J^{r-1}(E) & \dashrightarrow & J^1(E) & \xrightarrow[\pi_{1,0}]{} & J^0(E) \cong E \\ \downarrow{\scriptstyle \pi_r} & & \downarrow{\scriptstyle \pi_{r-1}} & & \downarrow{\scriptstyle \pi_1} & & \downarrow{\scriptstyle \pi} \\ M & \xrightarrow{\mathrm{id}_M} & M & \dashrightarrow & M & \xrightarrow{\mathrm{id}_M} & M \end{array} \tag{21.6.3}$$

Further, one may study what happens if the original bundle (E, M, π) carries additional structure, and whether any of this addition structure can be carried to jet bundles. We start with the case of vector bundles.

Theorem 21.6.6. *Let (E, M, π) be a vector bundle. Then also the jet bundles $(J^r(E), M, \pi_r)$ are vector bundles and the maps $\pi_{r,k}$ are vector bundle homomorphisms.*

Proof. To show that $(J^r(E), M, \pi_r)$ is a vector bundle, we have to check that for $p \in M$ the fiber $J^r_p(E)$ carries the structure of a vector space, and that the local trivializations restrict to vector space isomorphisms. ▶Thus, also the space of sections is a vector space, where the addition and scalar multiplication are defined pointwise. From this follows that also the jet spaces $J^r_x(E)$ for $x \in M$ are vector spaces, since for any local sections σ, τ around x and $\mu, \nu \in \mathbb{R}$ the definition

$$\mu j^r_x \sigma + \nu j^r_x \tau = j^r_x(\mu\sigma + \nu\tau) \tag{21.6.4}$$

yields a vector space structure. Thus, $\pi_r : J^r(E) \to M$ is a vector bundle. (Note, however, that the bundles $\pi_{r,k} : J^r(E) \to J^k(E)$, and thus in particular $\pi_{r,0} : J^r(E) \to E$, are *not* vector bundles, since the fibers of these bundles are not vector (sub)spaces, but affine spaces.)◀ ■

Having established this relation, it is a small step to generalize the statement to affine bundles in place of vector bundles.

Theorem 21.6.7. *Let (E, M, π) be an affine bundle modeled over the vector bundle $(\vec{E}, M, \vec{\pi})$. Then also the jet bundles $(J^r(E), M, \pi_r)$ are affine bundles modeled over the vector bundles $(J^r(\vec{E}), M, \vec{\pi}_r)$ and the maps $\pi_{r,k}$ are affine bundle morphisms.*

Proof. ▶. . . ◀ ■

Another interesting structure we have encountered in particular in the context of principal bundles is that of a fiber preserving group action. Here we consider a left action; of course, the same statement also holds for right actions.

Theorem 21.6.8. *Let (E, M, π) be a fiber bundle carrying a left action $\cdot : G \times E \to E, (g, e) \mapsto g \cdot e$ by a Lie group G that preserves the fibers, $\pi(g \cdot e) = \pi(e)$ for all $(g, e) \in G \times E$. Then also the jet bundles $(J^r(E), M, \pi_r)$ carry induced left actions, and the maps $\pi_{r,k}$ are equivariant.*

Proof. ►...◄ ■

So far, we have restricted ourselves to considering jet bundles over a single fiber bundle (E, M, π). We now consider how to relate jet bundles over different fiber bundles over a common base manifold. We start with the case of a fibered product.

Theorem 21.6.9. *Let (E, M, π) and (E', M, π') be two fiber bundles over a common base manifold M and $(E \times_M E', M, \pi \times_M \pi')$ their fibered product. Then their jet bundles are related by*

$$J^r(E \times_M E') \cong J^r(E) \times_M J^r(E') \,. \tag{21.6.5}$$

Proof. ►...◄ ■

Recall that if we have vector bundles, also their fibered product can be given the structure of a vector bundle, known as the direct sum. It is straightforward to conclude that also this structure can be carried over to the jet bundles.

Theorem 21.6.10. *Let (E, M, π) and (E', M, π') be two vector bundles over a common base manifold M and $(E \oplus E', M, \pi \oplus \pi')$ their direct sum. Then their jet bundles are related by*

$$J^r(E \oplus E') \cong J^r(E) \oplus J^r(E') \,. \tag{21.6.6}$$

Proof. ►...◄ ■

One may wonder whether the aforementioned statement holds also for tensor product bundles. However, this is not the case! This can most easily seen by studying the ranks of the respective bundles. Let E and E' be vector bundles of rank n and n', respectively, over a manifold M of dimension m. Then their tensor product bundle is of rank nn'. The rank of the jet bundles $J^r(E)$, $J^r(E')$ and $J^r(E \otimes E')$ is then given by

$$n\binom{m+r}{r}, \quad n'\binom{m+r}{r}, \quad nn'\binom{m+r}{r}. \tag{21.6.7}$$

However, the rank of $J^rE \otimes J^rE'$ is

$$nn'\binom{m+r}{r}^2 \neq nn'\binom{m+r}{r}. \tag{21.6.8}$$

21.7 Prolongation of bundle morphisms

Another important relation between fiber bundles can be established by a bundle morphism. We now show that this induces also a bundle morphism between the corresponding jet bundles. We start by showing the following.

Theorem 21.7.1. *Let (E, M, π) and (E', M, π') be two fiber bundles over a common base manifold M, $\varphi : E \to E'$ a bundle morphism covering the identity, $p \in M$, $r \in \mathbb{N}$ and $\sigma, \tau \in \Gamma_p(E)$ two local sections such that $\sigma \overset{p,r}{\sim} \tau$. Then $\varphi \circ \sigma \overset{p,r}{\sim} \varphi \circ \tau$.*

Proof. This follows directly from theorem 21.5.1, using the fact that $\sigma, \tau : M \to E$ and $\varphi : E \to E'$ are maps. ■

Recall from theorem 2.7.3 that $\varphi \circ \sigma : M \to E'$ and $\varphi \circ \tau : M \to E'$ are local sections of the bundle $\pi' : E' \to M$. From the theorem above follows that if σ and τ define the same r-jet at p, i.e., $j^r_p\sigma = j^r_p\tau$, then also $j^r_p(\varphi \circ \sigma) = j^r_p(\varphi \circ \tau)$. While the former is an element of $J^r_p(E)$, the latter belongs to $J^r_p(E')$. This allows to define the following map.

Definition 21.7.1 (Jet prolongation of bundle morphisms)**.** Let (E, M, π) and (E', M, π') be two fiber bundles over a common base manifold M and $\varphi : E \to E'$ a bundle morphism covering the identity. For $r \in \mathbb{N}$, the map $j^r\varphi : J^r(E) \to J^r(E')$ defined by

$$j^r\varphi(j^r_p\sigma) = j^r_p(\varphi \circ \sigma) \tag{21.7.1}$$

for each $p \in M$ and $\sigma \in \Gamma_p(E)$ is called the *r-jet prolongation* of φ.

One may ask for a coordinate representation of this construction. Let (x^μ) be coordinates on M and (x^μ, y^a) as well as (x^μ, z^A) induced coordinates on E and E', respectively. A section $\sigma : M \to E$ assigns to a point in M with coordinates $x = (x^\mu)$ the coordinate values $y^a(x)$, while the map $\varphi : E \to E'$ is fully determined by assigning fiber coordinates $z^A(x, y)$, since the coordinates x^μ are already determined by the fact that φ covers the identity. The composition $\varphi \circ \sigma$ is then represented by assigning to x the coordinates $z^A(x, y(x))$. The jet $j^r_p(\varphi \circ \sigma)$ is then determined by the derivatives with respect to x^μ up to order r, and hence by the chain rule depends only on derivatives of $y^a(x)$ up to order r, which determine the jet $j^r_p\sigma$.

We can now recall that $J^r(E)$ forms the total space of a number of bundles, in particular $\pi_r : J^r(E) \to M$ and $\pi_{r,k} : J^r(E) \to J^k(E)$, including the special case $J^0(E) \cong E$, and analogously for E' in place of E. One may therefore ask whether this bundle structure is preserved by $j^r\varphi$. We now show that this is indeed the case.

Theorem 21.7.2. *Let (E, M, π) and (E', M, π') be two fiber bundles over a common base manifold M and $\varphi : E \to E'$ a bundle morphism covering the identity. Then for $r \in \mathbb{N}$ the r-jet prolongation $j^r\varphi : J^r(E) \to J^r(E')$ is a bundle morphism from $\pi_r : J^r(E) \to M$ to $\pi'_r : J^r(E') \to M$ covering the identity on M, as well as a bundle morphisms from $\pi_{r,k} : J^r(E) \to J^k(E)$ to $\pi'_{r,k} : J^r(E') \to J^k(E')$ covering $j^k\varphi : J^k(E) \to J^k(E')$ for all $k = 0, \ldots, r$.*

Proof. Let $p \in M$ and $\sigma \in \Gamma_p(E)$ a local section. Then we have

$$\begin{aligned}
(\pi'_{r,k} \circ j^r\varphi)(j^r_p\sigma) &= \pi'_{r,k}(j^r_p(\varphi \circ \sigma)) \\
&= j^k_p(\varphi \circ \sigma) \\
&= j^k\varphi(j^k_p\sigma) \\
&= (j^k\varphi \circ \pi_{r,k})(j^r_p\sigma)\,,
\end{aligned} \tag{21.7.2}$$

as well as

$$(\pi'_r \circ j^r\varphi)(j^r_p\sigma) = \pi'_r(j^r_p(\varphi \circ \sigma)) = p = r(j^r_p\sigma)\,. \tag{21.7.3}$$

■

The preceding statement can also be stated by saying that the diagram

$$
\begin{array}{ccc}
J^r(E) & \xrightarrow{j^r\varphi} & J^r(E') \\
\downarrow{\scriptstyle \pi_{r,k}} & & \downarrow{\scriptstyle \pi'_{r,k}} \\
J^k(E) & \xrightarrow{j^k\varphi} & J^k(E') \\
\downarrow{\scriptstyle \pi_{k,0}} & & \downarrow{\scriptstyle \pi'_{k,0}} \\
E & \xrightarrow{\varphi} & E' \\
& {\scriptstyle \pi} \searrow \quad \swarrow {\scriptstyle \pi'} & \\
& M &
\end{array}
\tag{21.7.4}
$$

commutes. Recall from theorem 21.6.4 that there is a special case included above for $r = k+1$, in which case one finds the structure of an affine bundle. Naturally one may ask whether this affine bundle structure is preserved. This is indeed the case, as we will show next.

Theorem 21.7.3. *Let (E, M, π) and (E', M, π') be two fiber bundles over a common base manifold M and $\varphi : E \to E'$ a bundle morphism covering the identity. Then for $r \geq 1$ the map $j^r\varphi : J^r(E) \to J^r(E')$ is an affine bundle morphism from $\pi_{r,r-1} : J^r(E) \to J^{r-1}(E)$ to $\pi'_{r,k} : J^r(E') \to J^{r-1}(E')$ covering $j^{r-1}\varphi : J^{r-1}(E) \to J^{r-1}(E')$.*

Proof. ►. . . ◄ ■

The aforementioned statements hold for jet bundles over general fiber bundles $\pi : E \to M$. As we have seen in section 21.6, in the case that $\pi : E \to M$ is equipped with additional structure, this also carries to the jet bundles $\pi_r : J^r(E) \to M$. Hence, one may also ask whether this additional structure on the jet bundle is preserved by $j^r\varphi$, provided that φ preserves the structure on $\pi : E \to M$. We first show this for vector bundles.

Theorem 21.7.4. *Let (E, M, π) and (E', M, π') be two vector bundles over a common base manifold M and $\varphi : E \to E'$ a vector bundle morphism covering the identity. Then the maps $j^r\varphi : J^r(E) \to J^r(E')$ are vector bundle morphisms from $\pi_r : J^r(E) \to M$ to $\pi'_r : J^r(E') \to M$ covering the identity on M.*

Proof. ►. . . ◄ ■

For affine bundles, the following holds.

Theorem 21.7.5. *Let (E, M, π) and (E', M, π') be two affine bundles over a common base manifold M and $\varphi : E \to E'$ an affine bundle morphism covering the identity. Then the maps $j^r\varphi : J^r(E) \to J^r(E')$ are affine bundle morphisms from $\pi_r : J^r(E) \to M$ to $\pi'_r : J^r(E') \to M$ covering the identity on M.*

Proof. ►. . . ◄ ■

21.8 Prolongation of sections

Once we have constructed a fiber bundle, we are of course interested in its sections. For the jet bundle of a fiber bundle E there is a particular way to construct sections of $J^r(E)$ from the sections of E, which we define as follows.

Definition 21.8.1 (Jet prolongation)**.** Let $\pi : E \to M$ be a fiber bundle and $\sigma \in \Gamma_{|U}(E)$ a local section with domain $U \subset M$. For $r \in \mathbb{N}$ the *r-jet prolongation* $j^r\sigma \in \Gamma_{|U}(J^r(E))$ of σ is the local section with domain U of the bundle $\pi_r : J^r(E) \to M$ such that $(j^r\sigma)(p) = j^r_p\sigma$ for all $p \in U$.

It is once again easy to check that this construction is well-defined and indeed yields a section of the jet bundle. We illustrate this construction using coordinates (x^α) on $U \subset M$, (y^a) on F and (x^α, y^a) on $\pi^{-1}(U)$, from which we derive coordinates (x^α, y^a_Λ) on $\pi_r^{-1}(U)$. In these coordinates a section σ is locally expressed by the coordinate functions $y^a(x)$. Its r-jet prolongation $j^r\sigma$ is then expressed by the coordinate functions $y^a_\Lambda(x) = \partial_\Lambda y^a(x)$.

Now we have constructed an important and helpful tool which we will apply to physics. We can now make precise what it means that some function "depends on the value and derivatives up to order r of some section at some point". Such a function will simply be a function on $J^r(E)$, and if we feed it with a jet prolongation of some section, it will have exactly the dependence we need.

Recall from definition 21.6.2 that we have certain projection maps between the total spaces of jet bundles, which "forget" higher jet orders. It is straightforward that these also relate different jet prolongations.

Theorem 21.8.1. *Let $\pi : E \to M$ be a fiber bundle and $r \geq k$. For a local section $\sigma \in \Gamma_{|U}(E)$ with domain $U \subset M$ the jet prolongations $j^r\sigma$ and $j^k\sigma$ are related by $j^k\sigma = \pi_{r,k} \circ j^r\sigma$.*

Proof. For every $p \in U$ one has

$$(j^k\sigma)(p) = j^k_p\sigma = \pi_{r,k}(j^r_p\sigma) = (\pi_{r,k} \circ j^r\sigma)(p) \tag{21.8.1}$$

by the definition 21.6.2 of $\pi_{r,k}$. ■

Naturally, one may ask which sections of a jet bundle $\pi_r : J^r(E) \to M$ are prolongations of sections of $\pi : E \to M$. This question is easily answered by the fact that for any prolongation we can recover the original section by applying the projection $\pi_{r,0}$, which follows from the previous statement for $k = 0$. This leads to the following statement.

Theorem 21.8.2. *Let $\pi : E \to M$ be a fiber bundle and $J^r(E)$ its r'th jet bundle. A local section $\Sigma : U \to J^r(E)$ on $U \subset M$ is the prolongation of a section, i.e., $\Sigma = j^r\sigma$ for some $\sigma : U \to E$, if and only if $\Sigma = j^r(\pi_{r,0} \circ \Sigma)$, and one has $\sigma = \pi_{r,0} \circ \Sigma$.*

Proof. First note that if $\Sigma = j^r\sigma$ for some local section σ, then $\sigma = j^0\sigma$ can be recovered from Σ using theorem 21.8.1 as $\sigma = \pi_{r,0} \circ \Sigma$, and it follows that

$$\Sigma = j^r\sigma = j^r(\pi_{r,0} \circ \Sigma)\,. \tag{21.8.2}$$

Conversely, if the equation above holds, Σ is obviously a prolongation of $\sigma = \pi_{r,0} \circ \Sigma$, showing the equivalence of both statements. ■

21.9 Differential forms on jet bundles

We have seen in the last lecture that for every fiber bundle $\pi : E \to M$ the jet spaces $J^r(E)$ for $r \in \mathbb{N}$ form an inverse sequence

$$M \xleftarrow{\pi} E \xleftarrow{\pi_{1,0}} J^1(E) \xleftarrow{\pi_{2,1}} J^2(E) \xleftarrow{\pi_{3,2}} \ldots\,, \tag{21.9.1}$$

where the maps $\pi_{r,k} : J^r(E) \to J^k(E)$ are the projections of fiber bundles. This equips the jet bundles with a very rich structure, some of which can be seen when we study differential forms on the jet spaces. There are particular types of differential forms which deserve special attention. The first such type is something we have encountered already in section 19.4; in the special case if jet bundles, it is defined as follows.

Definition 21.9.1 (Horizontal form on a jet bundle)**.** Let $\pi : E \to M$ be a fiber bundle and $r \in \mathbb{N}$. A k-form $\omega \in \Omega^k(J^r(E))$ is called *horizontal* if it is horizontal with respect to $\pi_r : J^r(E) \to M$, i.e., if it vanishes on the kernel $\ker \pi_{r*}$ of $\pi_{r*} : TJ^r(E) \to TM$.

Hence, we define horizontality with respect to the fiber bundle $\pi_r : J^r(E) \to M$. To illustrate this definition, consider a trivializing subset $U \subset M$ equipped with coordinates (x^a), as well as adapted coordinates (x^a, y^μ) on $\pi^{-1}(U) \subset E$, from which we construct induced coordinates (x^a, y^μ_Λ) with $0 \leq |\Lambda| \leq r$ on $\pi_r^{-1}(U) \subset J^r(E)$. We further introduce the notation

$$\partial_a = \frac{\partial}{\partial x^a}\,, \quad \partial^\Lambda_\mu = \frac{\partial}{\partial y^\mu_\Lambda} \tag{21.9.2}$$

for the corresponding coordinate basis of $TJ^r(E)$. A tangent vector $\xi \in TJ^r(E)$ can thus be written in the form $\xi = u^a\partial_a + v^\mu_\Lambda \partial^\Lambda_\mu$. For the pushforward one then finds $\pi_{r*}(\xi) = u^a\partial_a$, where ∂_a now denotes the coordinate basis of TM. Hence, a vertical tangent vector satisfies $u^a = 0$, and is thus of the form $\xi = v^\mu_\Lambda \partial^\Lambda_\mu$. Writing the coordinate basis of $T^*J^r(E)$ as $\mathrm{d}x^a, \mathrm{d}y^\mu_\Lambda$ we thus find that $\omega \in \Omega^k(J^r(E))$ is horizontal if and only if it is of the form

$$\omega = \omega_{a_1\cdots a_k} \mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k}\,. \tag{21.9.3}$$

While horizontal forms can be defined on any fiber bundle, the structure of prolongations on jet bundles allow also the definition of a complimentary notion, which we define as follows.

Definition 21.9.2 (Contact form)**.** Let $\pi : E \to M$ be a fiber bundle and $r \in \mathbb{N}$. A k-form $\omega \in \Omega^k(J^r(E))$ is called a *contact form* if its pullback $(j^r\sigma)^*(\omega) \in \Omega^k(M)$ vanishes for every local section σ of $\pi : E \to M$.

In order to illustrate the space of contact forms, it is helpful to first restrict the discussion to contact one-forms. Let $\omega = \alpha_a \mathrm{d}x^a + \beta^\Lambda_\mu \mathrm{d}y^\mu_\Lambda \in \Omega^1(J^r(E))$ be a one-form and $\sigma : M \to E$ a section. Its r-jet prolongation $j^r\sigma : M \to J^r(E)$ takes the form

$$(x^a) \mapsto (x^a, y^\mu_\Lambda) = (x^a, \partial_\Lambda\sigma^\mu)\,. \tag{21.9.4}$$

For the pullback of ω one therefore has

$$(j^r\sigma)^*(\omega) = (\alpha_a + \beta^\Lambda_\mu \partial_a\partial_\Lambda\sigma^\mu)\mathrm{d}x^a \in \Omega^1(M)\,. \tag{21.9.5}$$

Hence, ω is a contact form if and only if the components α_a satisfy

$$\alpha_a = -\beta^\Lambda_\mu \partial_a\partial_\Lambda\sigma^\mu \tag{21.9.6}$$

for all $a = 1, \ldots, \dim M$ and all local sections σ. To understand the right hand side, recall that for $|\Lambda| < r$, $\partial_a\partial_\Lambda\sigma^\mu$ is simply the value of the coordinate $y^\mu_{\Lambda a}$, which is to be interpreted as

$$y^\mu_{\Lambda a} = y^\mu_{(\lambda_1,\ldots,\lambda_a+1,\ldots,\lambda_n)}\,. \tag{21.9.7}$$

Now it is useful to split the sum as

$$\alpha_a = -\sum_{|\Lambda|<r} \beta_\mu^\Lambda \partial_a \partial_\Lambda \sigma^\mu - \sum_{|\Lambda|=r} \beta_\mu^\Lambda \partial_a \partial_\Lambda \sigma^\mu = -\sum_{|\Lambda|<r} \beta_\mu^\Lambda y_{\Lambda a}^\mu - \sum_{|\Lambda|=r} \beta_\mu^\Lambda \partial_a \partial_\Lambda \sigma^\mu \,. \tag{21.9.8}$$

Recalling that α_a and β_μ^Λ are component functions depending on the coordinates (x^a, y_Λ^μ), we see that both the left hand side and the first sum are now fully expressed as functions of these coordinates, while the last sum still explicitly depends on the particular choice of σ. Since this equation must hold for all possible choices of σ, the coefficients in this second sum must vanish, hence $\beta_\mu^\Lambda = 0$ for $|\Lambda| = r$. Therefore, any contact one-form is uniquely determined by the choice of the component functions β_μ^Λ. Introducing a suitable basis, we can therefore write ω in the form $\omega = \beta_\mu^\Lambda \theta_\Lambda^\mu$, where we make use of the following definition.

Definition 21.9.3 (Basic contact one-form)**.** Let $\pi : E \to M$ be a fiber bundle and $r \in \mathbb{N}$, $U \subset M$ and (x^a), (x^a, y^μ) and (x^a, y_Λ^μ) with $0 \leq |\Lambda| \leq r$ local coordinates on U, $\pi^{-1}(U)$ and $\pi_r^{-1}(U)$, respectively. The *basic contact one-forms* with respect to these coordinates are the one-forms

$$\theta_\Lambda^\mu = \mathrm{d}y_\Lambda^\mu - y_{(\lambda_1+1,\lambda_2,\ldots,\lambda_n)}^\mu \mathrm{d}x^1 - y_{(\lambda_1,\lambda_2+1,\ldots,\lambda_n)}^\mu \mathrm{d}x^2 - \ldots - y_{(\lambda_1,\lambda_2,\ldots,\lambda_n+1)}^\mu \mathrm{d}x^n \,, \tag{21.9.9}$$

where $0 \leq |\Lambda| \leq r - 1$.

One easily checks that these are indeed contact forms. Further, one finds that every contact one-form must be of the form $\theta = \beta_\mu^\Lambda \theta_\Lambda^\mu$, so that the basic contact one-forms constitute a basis of all contact one-forms. This basis almost complements the basis $\mathrm{d}x^a$ of the horizontal one-forms, but not completely, since the space spanned by $\mathrm{d}y_\Lambda^\mu$ with $|\Lambda| = r$ is missing. Nevertheless, one may define a basis as follows.

Definition 21.9.4 (Contact basis)**.** Let $\pi : E \to M$ be a fiber bundle and $r \in \mathbb{N}$, $U \subset M$ and (x^a), (x^a, y^μ) and (x^a, y_Λ^μ) with $0 \leq |\Lambda| \leq r$ local coordinates on U, $\pi^{-1}(U)$ and $\pi_r^{-1}(U)$, respectively. The *contact basis* with respect to these coordinates is the basis of $\Omega^1(J^r(E))$ given by

$$(\mathrm{d}x^a, \theta_\Lambda^\mu, \mathrm{d}y_{\tilde\Lambda}^\mu) \tag{21.9.10}$$

with $0 \leq |\Lambda| \leq r - 1$ and $|\tilde\Lambda| = r$.

Using the exterior product, we may now construct higher order differential forms from these basis elements. Before we do so explicitly, it is helpful to state a few general properties of the exterior product of forms on jet bundles. We first remark the following property.

Theorem 21.9.1. *The exterior product of two horizontal forms is again horizontal.*

Proof. Let $\omega \in \Omega^k(J^r(E))$ and $\tau \in \Omega^l(J^r(E))$ be horizontal and $X \in \mathrm{Vect}(J^r(E))$ a vertical vector field, i.e., $\pi_{r*} \circ X = 0$. Then we have

$$\iota_X(\omega \wedge \tau) = \iota_X \omega \wedge \tau + (-1)^k \omega \wedge \iota_X \tau = 0 \,, \tag{21.9.11}$$

so that also $\omega \wedge \tau$ is horizontal. ∎

An even stronger statement holds for contact forms.

Theorem 21.9.2. *The contact forms form an ideal (the* contact ideal*) of the exterior algebra, i.e., the exterior product of an arbitrary form and a contact form is again a contact form.*

Proof. Let $\omega \in \Omega^k(J^r(E))$ be a contact form, $\tau \in \Omega^l(J^r(E))$ an arbitrary l-form and $\sigma : U \to E$ a local section on $U \subset M$. Then we have the pullback

$$(j^r\sigma)^*(\omega \wedge \tau) = (j^r\sigma)^*(\omega) \wedge (j^r\sigma)^*(\tau) = 0\,, \tag{21.9.12}$$

so that also $\omega \wedge \tau$ is a contact form. ■

With these statements in place we now proceed with the following definition.

Definition 21.9.5 (l-contact form)**.** Let $\pi : E \to M$ be a fiber bundle and $r \in \mathbb{N}$. A $(k+l)$-form $\omega \in \Omega^{k+l}(J^r(E))$ is called *l-contact* if it is a linear combination of exterior products of k horizontal one-forms and l contact one-forms. The space of all such forms is denoted $\Omega^{k,l}(J^r(E))$.

Obviously one has $\Omega^{k,l}(J^r(E)) \subset \Omega^{k+l}(J^r(E))$. However, one cannot completely decompose $\Omega^{k+l}(J^r(E))$ into such subspaces, since the forms $\mathrm{d}y^\mu_\Lambda$ with $|\Lambda| = r$ do not belong to any such space. However, this can easily be fixed by lifting forms to the next higher jet bundle, which leads to the following theorem.

Theorem 21.9.3. *Let $\pi : E \to M$ be a fiber bundle and $r \in \mathbb{N}$. The pullback of every k-form $\omega \in \Omega^k(J^r(E))$ along $\pi_{r+1,r}$ uniquely decomposes into*

$$\pi^*_{r+1,r}(\omega) = \sum_{i=0}^{k} \mathrm{p}_i\omega \in \Omega^k(J^{r+1}(E))\,, \tag{21.9.13}$$

where $\mathrm{p}_i\omega \in \Omega^{k-i,i}(J^{r+1}(E))$.

Proof. ►...◄ ■

Note in particular that the component $\mathrm{p}_0\omega$ is horizontal, and hence also the notation $h\omega$ can be found in the literature [Kru15]. We then proceed with the following statement.

Theorem 21.9.4. *Let $\omega \in \Omega^{k,l}(J^r(E))$ and $\chi \in \Omega^{k',l'}(J^r(E))$. Then the following hold:*

1. $\omega \wedge \chi \in \Omega^{k+k',l+l'}(J^r(E))$,
2. $\pi^*_{r+1,r}(\mathrm{d}\omega) \in \Omega^{k+1,l}(J^{r+1}(E)) \oplus \Omega^{k,l+1}(J^{r+1}(E))$.

Proof. ►...◄ ■

This property of the exterior product is immediately clear. For the exterior derivative it means that $\pi^*_{r+1,r}(\mathrm{d}\omega)$ can be uniquely written as the sum of two terms, one of them belonging to $\Omega^{k+1,l}(J^{r+1}(E))$, the other one to $\Omega^{k,l+1}(J^{r+1}(E))$:

$$\pi^*_{r+1,r}(\mathrm{d}\omega) = \mathrm{p}_l\mathrm{d}\omega + \mathrm{p}_{l+1}\mathrm{d}\omega\,. \tag{21.9.14}$$

This allows us to decompose the exterior derivative in the following way.

Definition 21.9.6 (Horizontal and vertical differentials)**.** Let $\pi : E \to M$ be a fiber bundle. For $k, l \in \mathbb{N}$ the *horizontal* (or total) differential

$$\begin{array}{rcccl} \mathrm{d}_H & : & \Omega^{k,l}(J^r(E)) & \to & \Omega^{k+1,l}(J^{r+1}(E)) \\ & & \omega & \mapsto & \mathrm{p}_l \mathrm{d}\omega \end{array} \tag{21.9.15}$$

and *vertical* differential

$$\begin{array}{rcccl} \mathrm{d}_V & : & \Omega^{k,l}(J^r(E)) & \to & \Omega^{k,l+1}(J^{r+1}(E)) \\ & & \omega & \mapsto & \mathrm{p}_{l+1} \mathrm{d}\omega \end{array} \tag{21.9.16}$$

are the unique functions such that $\mathrm{d}_H\omega + \mathrm{d}_V\omega = \pi^*_{r+1,r}(\mathrm{d}\omega)$ for all $\omega \in \Omega^{k,l}(J^r(E))$.

In order to work with these differentials, we first state a few of their properties, which will then allow us to write them using coordinates.

Theorem 21.9.5. *For each $\omega \in \Omega^{k,l}(J^r(E))$ and $\chi \in \Omega^{k',l'}(J^r(E))$ the horizontal and vertical differentials d_H and d_V satisfy:*

1. *d_H and d_V are antiderivations:*

$$\mathrm{d}_H(\omega \wedge \chi) = \mathrm{d}_H\omega \wedge \pi^*_{r+1,r}(\chi) + (-1)^{k+l}\pi^*_{r+1,r}(\omega) \wedge \mathrm{d}_H\chi\,, \tag{21.9.17a}$$

$$\mathrm{d}_V(\omega \wedge \chi) = \mathrm{d}_V\omega \wedge \pi^*_{r+1,r}(\chi) + (-1)^{k+l}\pi^*_{r+1,r}(\omega) \wedge \mathrm{d}_V\chi\,. \tag{21.9.17b}$$

2. *$\mathrm{d}_H^2 = 0$, $\mathrm{d}_V^2 = 0$ and $\mathrm{d}_H\mathrm{d}_V = -\mathrm{d}_V\mathrm{d}_H$.*

Proof. ▶ . . . ◀ ■

With these properties we can now construct the coordinate expressions for d_H and d_V by applying them to functions (zero-forms) and one-forms, since all differential forms can be constructed from these simplest forms. For $f \in \Omega^{0,0}(J^r(E))$, the exterior derivative can be written in the contact basis as

$$\begin{aligned} \pi^*_{r+1,r}(\mathrm{d}f) &= \frac{\partial f}{\partial x^a}\mathrm{d}x^a + \frac{\partial f}{\partial y^\mu_\Lambda}\mathrm{d}y^\mu_\Lambda \\ &= \frac{\partial f}{\partial x^a}\mathrm{d}x^a + \frac{\partial f}{\partial y^\mu_\Lambda}(\theta^\mu_\Lambda + y^\mu_{\Lambda a}\mathrm{d}x^a) \\ &= \left(\frac{\partial f}{\partial x^a} + y^\mu_{\Lambda a}\frac{\partial f}{\partial y^\mu_\Lambda}\right)\mathrm{d}x^a + \frac{\partial f}{\partial y^\mu_\Lambda}\theta^\mu_\Lambda\,. \end{aligned} \tag{21.9.18}$$

Hence, we have the vertical differential given by

$$\mathrm{d}_V f = \frac{\partial f}{\partial y^\mu_\Lambda}\theta^\mu_\Lambda \in \Omega^{0,1}(J^{r+1}(E))\,, \tag{21.9.19}$$

where the summation over Λ goes over $0 \leq |\Lambda| \leq r$, and hence covers the contact basis of $\Omega^{0,1}(J^{r+1}(E))$. For the horizontal differential then follows

$$\mathrm{d}_H f = \pi^*_{r+1,r}(\mathrm{d}f) - \mathrm{d}_V f = \mathrm{D}_a f\,\mathrm{d}x^a \in \Omega^{1,0}(J^{r+1}(E))\,, \tag{21.9.20}$$

where we introduced the total derivative

$$\mathrm{D}_a f = \frac{\partial f}{\partial x^a} + y^\mu_{\Lambda a}\frac{\partial f}{\partial y^\mu_\Lambda} = \frac{\partial f}{\partial x^a} + \sum_{|\Lambda|=0}^{r} y^\mu_{(\lambda_1,\ldots,\lambda_\alpha+1,\ldots,\lambda_n)}\frac{\partial f}{\partial y^\mu_\Lambda}\,. \tag{21.9.21}$$

For the horizontal and vertical coordinate differentials we have

$$\mathrm{d}_H(\mathrm{d}x^a) = 0\,, \quad \mathrm{d}_V(\mathrm{d}x^a) = 0\,, \quad \mathrm{d}_H(\mathrm{d}y^\mu_\Lambda) = 0\,, \quad \mathrm{d}_V(\mathrm{d}y^\mu_\Lambda) = 0\,, \tag{21.9.22}$$

which follows from the fact that they are closed, i.e., their exterior derivatives vanish. Finally, the basic contact forms satisfy

$$\begin{aligned}\pi^*_{r+1,r}(\mathrm{d}\theta^\mu_\Lambda) &= \pi^*_{r+1,r}[\mathrm{d}(\mathrm{d}y^\mu_\Lambda - y^\mu_{\Lambda a}\mathrm{d}x^a)]\\ &= -\mathrm{d}y^\mu_{\Lambda a} \wedge \mathrm{d}x^a\\ &= \mathrm{d}x^a \wedge (\theta^\mu_{\Lambda a} + y^\mu_{\Lambda ab}\mathrm{d}x^b)\\ &= \mathrm{d}x^a \wedge \theta^\mu_{\Lambda a}\,,\end{aligned} \tag{21.9.23}$$

where the second term in the last step vanishes since the coordinates $y^\mu_{\Lambda ab}$ (which are, in fact, coordinates on $J^{r+2}(E)$) are symmetric, while $\mathrm{d}x^a \wedge \mathrm{d}x^b$ is antisymmetric. It follows that the vertical differential vanishes,

$$\mathrm{d}_V\theta^\mu_\Lambda = 0\,, \tag{21.9.24}$$

while the horizontal differential is given by

$$\begin{aligned}\mathrm{d}_H\theta^\mu_\Lambda &= \mathrm{d}x^a \wedge \theta^\mu_{\Lambda a}\\ &= \mathrm{d}x^1 \wedge \theta^\mu_{(\lambda_1+1,\lambda_2,\ldots,\lambda_n)} + \mathrm{d}x^2 \wedge \theta^\mu_{(\lambda_1,\lambda_2+1,\ldots,\lambda_n)} + \ldots + \mathrm{d}x^n \wedge \theta^\mu_{(\lambda_1,\lambda_2,\ldots,\lambda_n+1)}\end{aligned} \tag{21.9.25}$$

Since any differential form on $J^r(E)$ can be constructed as a linear combination of wedge products of the forms above, we can thus explicitly calculate the vertical and horizontal differentials for all differential forms.

Chapter 22

Frame bundles

22.1 Frame bundles over vector bundles

An important class of fiber bundles, which can be constructed from any vector bundles, is given by *frame bundles*. There exist different types of frame bundles, most of which require an additional structure to be defined on the fibers of the vector bundle under consideration. However, the most simple type does not require any such structure. It can be defined as follows.

Definition 22.1.1 (Frame bundle)**.** Let $\pi : E \to M$ be a vector bundle of rank k, with total space E, fiber $\mathbb{R}^k$ and base manifold M. A *frame* at $x \in M$ is a bijective linear function $p : \mathbb{R}^k \to E_x \equiv \pi^{-1}(x) \subset E$. The set of all frames constitutes the (general linear) *frame bundle* $F(E)$ with projection $\varpi : F(E) \to M$ mapping $p : \mathbb{R}^k \to E_x$ to $x \in M$.

This definition deserves a few clarifications. Recall that for a vector bundle of rank k, for every base point $x \in M$ the fiber $E_x \subset E$ over x is diffeomorphic to $\mathbb{R}^k$ and carries the same vector space structure. Hence, there is a well-defined notion of bijective, linear maps $p : \mathbb{R}^k \to E_x$. All these maps over a given point x, which we call frames, constitute the fiber $F_x(E) \subset F(E)$ of the frame bundle $F(E)$ over x. Since every frame maps all of $\mathbb{R}^k$ into the same fiber E_x, there exists a well defined projection map $\varpi : F(E) \to M$, which assigns to $p \in F(E)$ the base point x. This projection map can also explicitly written as follows. We take the zero element of $\mathbb{R}^k$, which is mapped by a frame to the zero element $p(0) \in E_x$ of some fiber E_x. The base point x of this fiber is recovered as $x = \pi(p(0))$. Hence, we may write the projection ϖ as $\varpi(p) = \pi(p(0))$. In fact, instead of the zero element we could have chosen any other element of $\mathbb{R}^k$ here, since all are mapped into the same fiber E_x. Choosing the zero element is just convenient, because it is canonically distinguished.

Another and more intuitive picture of a frame is that of an ordered basis of the vector space E_x. This picture is related to the definition above as follows. Denoting the canonical basis of $\mathbb{R}^k$ by $(\mathbb{e}_i, i = 1, \ldots, k)$, a frame defines an ordered basis $(p_i, i = 1, \ldots, k)$ of E_x as $p_i = p(\mathbb{e}_i)$. Conversely, the frame $p : \mathbb{R}^k \to E_x$ can be constructed from an ordered basis (p_i) as the map $p : v^i\mathbb{e}_i \mapsto v^i p_i$, which is linear by definition.

In order to complete the fiber bundle structure of $F(E)$, one also needs to provide its local trivializations. We will do so alongside the following statement. Using the fact that the group $\mathrm{GL}(k, \mathbb{R})$ acts on the bases of a real, k-dimensional vector space, one deduces the following important property of frame bundles.

Theorem 22.1.1. *The frame bundle $F(E)$ over a vector bundle E of rank k is a principal* *fiber bundle with structure group* $\mathrm{GL}(k,\mathbb{R})$*, where the right action is defined as* $p \cdot g = p \circ g$ *for* $p \in F(E)$ *and* $g \in \mathrm{GL}(k,\mathbb{R})$*.*

Proof. Let $x \in M$ and $p, p' \in F_x(E)$. Then there exists a unique $g \in \mathrm{GL}(k,\mathbb{R})$, such that $p' = p \circ g$, which is given by $g = p^{-1} \circ p'$. Hence, $g \in \mathrm{GL}(k,\mathbb{R})$ acts freely and transitively on the fibers of $F(E)$. Further, since E is a vector bundle, there exists a local basis (ϵ_i) on $U \subset M$ with $x \in U$, which defines a local section $\epsilon : U \to F(E)$. Setting $\phi^{-1}(x,g) = \epsilon(x) \circ g$ defines a local trivialization ϕ of $F(E)$, which equips it with the structure of a smooth fiber bundle. ■

Using either of the interpretations given above, we may use local coordinates on E in order to construct local coordinates on $F(E)$. For this purpose, let $U \subset M$ be a trivializing subset of M with coordinates $(x^\mu, \mu = 1, \ldots, n)$, and consider induced vector bundle coordinates (x^μ, y^a) on $\pi^{-1}(U)$, where $a = 1, \ldots, k$. By definition, these coordinates are chosen such that they correspond to a basis ϵ_a of E_x for every $x \in U$. A frame $p : \mathbb{R}^k \to E_x$ is then uniquely defined by the coordinates x^μ of the base point x, together with the components $p^a{}_i$ of the images $p(\mathbb{e}_i) = p^a{}_i \epsilon_a$ of the basis vectors $\mathbb{e}_i$. Hence, coordinates on $\varpi^{-1}(U) \subset F(E)$ are given by $(x^\mu, p^a{}_i)$, where $p^a{}_i$ are the components of an invertible matrix.

22.2 Vector bundles as associated bundles

Note that the frame bundle is canonically constructed from any vector bundle. One may ask whether also the converse construction is possible, i.e., whether one can obtain a vector bundle from a principal $\mathrm{GL}(k,\mathbb{R})$-bundle, since the latter is reminiscent of a frame bundle. One may already guess that this is possible by using the construction of associated fiber bundles detailed in the previous section, where the action is given by the natural left action $\rho : \mathrm{GL}(k,\mathbb{R}) \times \mathbb{R}^k \to \mathbb{R}^k$ given by matrix multiplication. Hence, one has two constructions: Given a vector bundle $\pi : E \to M$, one finds the frame bundle $F(E)$, which is a principal $\mathrm{GL}(k,\mathbb{R})$-bundle, while from a principal $\mathrm{GL}(k,\mathbb{R})$-bundle P one finds the associated bundle $P \times_\rho \mathbb{R}^k$, which is a vector bundle of rank k. We now check that each of these constructions is indeed the inverse of the other.

Theorem 22.2.1. *Let* $\varpi : P \to M$ *be a principal* $\mathrm{GL}(k,\mathbb{R})$*-bundle and* $\rho : \mathrm{GL}(k,\mathbb{R}) \times \mathbb{R}^k \to \mathbb{R}^k$ *the left action given by matrix multiplication. Then the associated bundle* $P \times_\rho \mathbb{R}^k$ *is a vector bundle, and its frame bundle* $F(P \times_\rho \mathbb{R}^k)$ *is canonically isomorphic to* P*.*

Proof. It follows from the fact that ρ is a linear representation of $\mathrm{GL}(k,\mathbb{R})$ and theorem 20.4.1 that $P \times_\rho \mathbb{R}^k$ is a vector bundle of rank k. To show that its frame bundle is canonically isomorphic to P, consider an element $p \in P$, and let $x = \varpi(p) \in M$. Recall that p gives rise to a fiber diffeomorphism $[p] : \mathbb{R}^k \to \varpi_\rho^{-1}(x)$, where ϖ_ρ denotes the bundle projection of $P \times_\rho \mathbb{R}^k$, which satisfies

$$[p](v) = [p, v]\,. \tag{22.2.1}$$

By construction of the associated bundle, this map is bijective. Further, it is linear, which can be seen from the construction of the vector bundle structure in theorem 20.4.1. Hence, it constitutes a frame at x, and so $[p] \in F(P \times_\rho F)$. We now show that the assignment $p \mapsto [p]$ is a principal bundle isomorphism from P to $F(P \times_\rho F)$. Let $q \in F(P_x \times_\rho F)$ be a frame at $x \in M$. Then there exists $p \in P_x$ with $[p] \in F(P \times_\rho F)$. Since both $[p]$ and q are frames, there exists a unique $g \in \mathrm{GL}(k,\mathbb{R})$ such that $q = [p] \circ g$. This yields another element $p \cdot g \in P_x$, for which holds

$$[p \cdot g](v) = [p](\rho(g,v)) = [p](gv) = ([p] \circ g)(v) = q(v) \tag{22.2.2}$$

for all $v \in \mathbb{R}^k$, and hence $[p \cdot g] = q$. This shows that $p \mapsto [p]$ is surjective, since for every $q \in F(P \times_\rho F)$ we can find a preimage in P. To show that it is also injective, we must show that this preimage does not depend on the choice of the initial element p. If we had chosen another element p', we would have found $p' \cdot g'$ with $[p' \cdot g'] = q$ for another $g' \in \mathrm{GL}(k, \mathbb{R})$. Since both $p \cdot g$ and $p' \cdot g'$ are elements of P_x, there exists a unique $\tilde{g}$ such that

$$p \cdot g \cdot \tilde{g} = p' \cdot g' . \tag{22.2.3}$$

Then we have

$$q = [p' \cdot g'] = [p \cdot g \cdot \tilde{g}] = [p \cdot g] \circ \tilde{g} = q \cdot \tilde{g} , \tag{22.2.4}$$

and thus $\tilde{g} = \mathbb{1}$, by applying the inverse frame q^{-1} from the left, proving that $p \cdot g = p' \cdot g'$ is unique, independent of the choice of p, and hence $p \mapsto [p]$ is also injective, and therefore bijective. Further, it preserves the fibers, since $[p] \in F(P_x \times_\rho F)$ for $p \in P_x$, and compatible with the right action, as follows from the relation (22.2.2). ►Show smoothness.◄ ∎

Next, we discuss the opposite direction.

Theorem 22.2.2. *Let $\pi : E \to M$ be a vector bundle of rank k and $\rho : \mathrm{GL}(k, \mathbb{R}) \times \mathbb{R}^k \to \mathbb{R}^k, (g, v) \mapsto gv$ the natural left action given by matrix multiplication. Then $F(E) \times_\rho \mathbb{R}^k$ is canonically isomorphic to E, where $F(E)$ is the frame bundle of E.*

Proof. We have to check that there is a one-to-one correspondence between elements $e \in E$ and $[p, v] \in F(E) \times_\rho \mathbb{R}^k$, which constitutes a vector bundle isomorphism between these two bundles. Given $[p, v] \in F(E) \times_\rho \mathbb{R}^k$, one can define $e = p(v)$. This is independent of the choice of the representative, since for $[p, v] = [p', v']$ one has an element $g \in \mathrm{GL}(k, \mathbb{R})$ such that $p' = p \cdot g$ and $v' = g^{-1}v$, and hence

$$e' = p'(v') = (p \circ g \circ g^{-1})(v) = p(v) = e . \tag{22.2.5}$$

Hence, this defines a map $\varphi : F(E) \times_\rho \mathbb{R}^k \to E$. To show that this is bijective, we explicitly construct its inverse. For $e \in E$, let $x = \pi(e) \in M$ and pick a frame $p \in F_x(E)$. Setting $v = p^{-1}(e) \in \mathbb{R}^k$, one obtains an element $[p, v] \in F(E) \times_\rho \mathbb{R}^k$. Using the same argument as above, one finds that this is independent of the choice of p. Further, it is obvious that $p(v) = e$, and so these two maps indeed establish a one-to-one correspondence between $[p, v] \in F(E) \times_\rho \mathbb{R}^k$ and $e \in E$. Finally, this one-to-one correspondence covers the identity on M, since $e \in E_x$ and $p \in F_x(M)$ have the same base point $x \in M$.

To check the smoothness of this map, we show that it relates the local trivializations. Recall that a local trivialization $\tilde{\epsilon}$ of E on $U \subset M$ can be obtained from a local basis ϵ on U as the inverse of

$$\tilde{\epsilon}^{-1}(x, v) \mapsto \epsilon_i(x) v^i . \tag{22.2.6}$$

This is equivalent to a local section $\epsilon : U \to F(E)$, which then defines a local trivialization ϕ of $F(E)$ as the inverse of

$$\phi(x, g) = \epsilon(x) \cdot g , \tag{22.2.7}$$

as well as a local trivialization ϕ_ρ of $F(E) \times_\rho \mathbb{R}^k$ as the inverse of the map

$$\phi_\rho^{-1}(x, v) = [\epsilon(x), v] . \tag{22.2.8}$$

Now it is easy to see that

$$\varphi(\phi_\rho^{-1}(x, v)) = \varphi([\epsilon(x), v]) = \epsilon(x)(v) = \epsilon_i(x) v^i = \tilde{\epsilon}^{-1}(x, v) , \tag{22.2.9}$$

and so φ relates the local trivializations of E and $F(E) \times_\rho \mathbb{R}^k$. Finally, we can write

$$\varphi|_{\varpi_\rho^{-1}(U)} = \tilde{\epsilon}^{-1} \circ \phi_\rho , \tag{22.2.10}$$

showing that φ is a composition of smooth maps on $\varpi_\rho^{-1}(U)$, and thus smooth. Since the local trivializations cover $F(E) \times_\rho \mathbb{R}^k$, it is smooth everywhere, thus finally proving that it is indeed a vector bundle isomorphism. ∎

22.3 Coframes

As discussed in section 4.1, for every vector vector bundle $\pi : E \to M$, there exists another canonically defined vector bundle of the same rank, which is given by the dual bundle E^*. Of course, also E^* comes with a frame bundle $F(E^*)$, and it is straightforward to expect that the frame bundles of E and E^* are closely related. In this section we clarify this relation, and provide a few notions which are frequently used in the relevant literature. We start with the following definition.

Definition 22.3.1 (Coframe bundle). Let $\pi : E \to M$ be a vector bundle of rank k, with total space E, fiber $\mathbb{R}^k$ and base manifold M. A *coframe* at $x \in M$ is a bijective linear function $q : E_x \to \mathbb{R}^k$. The set of all coframes constitutes the (general linear) *coframe bundle* $F^*(E)$ with projection $\varpi : F^*(E) \to M$ mapping $q : E_x \to \mathbb{R}^k$ to $x \in M$.

Further, we specify a right action on $F^*(E)$:

Theorem 22.3.1. *The coframe bundle $F^*(E)$ over a vector bundle E of rank k is a principal fiber bundle with structure group $\mathrm{GL}(k, \mathbb{R})$, where the right action is defined as $q \cdot g = g^{-1} \circ q$ for $q \in F^*(E)$ and $g \in \mathrm{GL}(k, \mathbb{R})$.*

Proof. Let $x \in M$ and $q, q' \in F^*_x(E)$. Then there exists a unique $g \in \mathrm{GL}(k, \mathbb{R})$, such that $q' = g^{-1} \circ q$, which is given by $g = q \circ q'^{-1}$. Hence, $g \in \mathrm{GL}(k, \mathbb{R})$ acts freely and transitively on the fibers of $F^*(E)$. Further, since E is a vector bundle, there exists a local basis (ϵ_i) on $U \subset M$ with $x \in U$, which defines a local section $\epsilon^{-1} : U \to F^*(E)$. Setting $\phi^{-1}(x, g) = g^{-1} \circ \epsilon^{-1}(x)$ defines a local trivialization ϕ of $F^*(E)$, which equips it with the structure of a smooth fiber bundle. ∎

Note that in contrast to the frame bundle, we need to compose from the left, and to use the inverse group element in order to still obtain a right action. Now the following statement becomes obvious.

Theorem 22.3.2. *For any vector bundle E, the frame bundle $F(E)$ and coframe bundle $F^*(E)$ are canonically isomorphic, and the isomorphism is given by $\bullet^{-1} : F(E) \to F^*(E)$.*

Proof. Clearly, given $x \in M$, $p^{-1} \in F^*_x(E)$ is a coframe at x for each $p \in F_x(E)$ and vice versa, so that $\bullet^{-1}$ is a bijection which preserves the fibers. It commutes with the right multiplication, since

$$(p \cdot g)^{-1} = (p \circ g)^{-1} = g^{-1} \circ p^{-1} = p^{-1} \cdot g \tag{22.3.1}$$

for all $g \in \mathrm{GL}(k, \mathbb{R})$. Finally, if (U, ϕ) is a local trivialization of $F(E)$, then $(U, \phi \circ \bullet^{-1})$ is a local trivialization of $F^*(E)$ and vice versa, which follows from the construction of local trivializations given in theorems 22.1.1 and 22.3.1. Hence, $\bullet^{-1}$ is an isomorphism of principal $\mathrm{GL}(k, \mathbb{R})$-bundles. ∎

There is, of course, no practical difference between working with frames or coframes, since being elements of canonically isomorphic bundles, both contain exactly the same information. The only difference lies in their interpretation. Recall that we interpreted a frame $p : \mathbb{R}^k \to E_x$ at $x \in M$ as a basis (p_i) of E_x, with basis vectors $p_i = p(\mathbb{e}_i)$. Similarly, we can interpret a coframe $q : E_x \to \mathbb{R}^k$, which assigns to $y \in E_x$ an element $q(y) = y^i \mathbb{e}_i$, as a cobasis (q^i) with $q^i(y) = y^i$. The latter, of course, is nothing else than a basis of the dual vector space E^*_x, which is a fiber of the dual bundle E^*. This finally brings us to the following conclusion.

Theorem 22.3.3. *For any vector bundle $\pi : E \to M$ of rank k, its frame bundle $F(E)$ and frame bundle $F(E^*)$ of its dual E^* are canonically isomorphic.*

Proof. For $x \in M$, let $q : \mathbb{R}^k \to E_x^*$ be a frame of the dual bundle, and denote by $\tilde{q} : E_x \to \mathbb{R}^k, y \mapsto \langle q(\mathbb{e}^i), y \rangle \mathbb{e}_i$ the corresponding coframe $\tilde{q} \in F_x^*(E)$. This induces a bijective mapping between $F(E^*)$ and $F^*(E)$, which can easily be seen by constructing the inverse. For $\tilde{q} : E_x \to \mathbb{R}^k$, we define the corresponding dual bundle frame as the unique element $q \in F_x(E^*)$ which satisfies $\langle q(\mathbb{e}^i), y \rangle >= \langle \mathbb{e}^i, q(y) \rangle$ for all $y \in E_x$. Together with the isomorphism between $F(E)$ and $F^*(E)$, we thus have a bijective mapping between $F(E)$ and $F(E^*)$. ►Show compatibility with right action. . . ◄ ■

Keep in mind, however, although $F(E) \cong F(E^*)$ via the canonical isomorphism given above, these exists no such canonical isomorphism between E and E^*! These are distinct bundles, and an in order to construct an isomorphism between them one must supply *additional* information. The bundle E^* it self, however, is of course fully defined by E. Following our discussion in section 22.2, we can obtain E as an associated bundle to $F(E)$, and analogously E^* as associated bundle to $F(E^*)$. The fact that these two frame bundles are isomorphic suggests that we can also obtain E^* as associated bundle to $F(E)$. This is of course the case.

Theorem 22.3.4. *Let $\pi : E \to M$ be a vector bundle of rank k and $\bar{\rho} : \mathrm{GL}(k, \mathbb{R}) \times \mathbb{R}^k \to \mathbb{R}^k, (g, v) \mapsto (g^t)^{-1} v$ the dual of the natural left action given by matrix multiplication. Then $F(E) \times_{\bar{\rho}} \mathbb{R}^k$ is canonically isomorphic to the dual bundle E^*, where $F(E)$ is the frame bundle of E.*

Proof. ►. . . ◄ ■

22.4 Tensor bundles as associated bundles

22.5 Higher order frame bundles

Recall from definition 22.1.1 that we defined a frame at a point $x \in M$ of a vector bundle $\pi : E \to M$ of rank k as a bijective linear function $p : \mathbb{R}^k \to E_x$. We now introduce another perspective on frames, which allows us to generalize their notion. Recall that a linear function $p : \mathbb{R}^k \to E_x$ in particular maps $0 \in \mathbb{R}^k$ to $0 \in E_x$, and that the components of $p(v^i \mathbb{e}_i)$ with respect to some given basis of E_x are simply linear functions of the components v^i, i.e., homogeneous first-order polynomials. In other words, we can identify the linear function p by its first order jet $j_0^1 p \in J_0^1(\mathbb{R}^k, E_x)_0$. The additional requirement that p is bijective, and so possesses an inverse p^{-1}, then implies that this jet must be invertible, $j_0^1 p \in \tilde{J}_0^1(\mathbb{R}^k, E_x)_0$. We can thus canonically identify the fiber $F_x(E)$ of frames at $x \in M$ with the space $\tilde{J}_0^1(\mathbb{R}^k, E_x)_0$, and even write just p instead of $j_0^1 p$. This identification suggests a simple generalization, which we define as follows.

Definition 22.5.1 (Higher order frame bundle)**.** Let $\pi : E \to M$ be a vector bundle of rank k, with total space E, fiber $\mathbb{R}^k$ and base manifold M. A *r-th order frame* at $x \in M$ is an invertible jet $p \in \tilde{J}_0^r(\mathbb{R}^k, E_x)_0$. The set of all r-th order frames constitutes the *r-th order frame bundle* $F^r(E)$ with projection $\varpi_r : F^r(E) \to M$ mapping $p \in \tilde{J}_0^r(\mathbb{R}^k, E_x)_0$ to $x \in M$.

In section 22.1 we have see that the ordinary frame bundle is a principal bundle with structure group $\mathrm{GL}(k, \mathbb{R})$. One may therefore wonder whether a similar property holds also for the higher order frame bundles we defined above. We now show that this is indeed the case.

Theorem 22.5.1. *The r-th order frame bundle $F^r(E)$ over a vector bundle E of rank k is a principal fiber bundle with structure group $\mathrm{J}^r(k)$, where the right action is defined as $p \cdot g = p \circ g$ for $p \in F^r(E)$ and $g \in \mathrm{J}^r(k)$.*

Proof. Let $x \in M$. Given $p \in F^r_x(E) = \tilde{J}^r_0(\mathbb{R}^k, E_x)_0$ and $g \in \mathrm{J}^r(k) = \tilde{J}^r_0(\mathbb{R}^k, \mathbb{R}^k)_0$, their exists a well-defined composition $p \circ g \in \tilde{J}^r_0(\mathbb{R}^k, E_x)_0$. This defines an action of $\mathrm{J}^r(k)$ on $F^r(E)$ which preserves the fibers. This is a right action, since for $p \in F^r_x(E)$ and $g, h \in \mathrm{J}^r(k)$ holds

$$p \circ (gh) = p \circ (g \circ h) = (p \circ g) \circ h\,, \tag{22.5.1}$$

since jet composition is associative. This action is also smooth, since jet composition is smooth. From the fact that frames are invertible by definition further follows that for each pair $p, p' \in F^r_x(E)$ of frames at the same point $x \in M$ there exists a unique element $g = p^{-1} \circ p' \in \mathrm{J}^r(k)$ such that $p' = p \circ g$. Hence, the action is free and transitive. Finally, given a local basis (ϵ_i) of E on $U \subset M$, which we can also understand as a local section $\epsilon : U \to F(E)$, we can finally construct a local trivialization (U, ϕ) by defining

$$\phi(p) = (\varpi_r(p), j^r_0(\epsilon(x)^{-1}) \circ p) \in U \times \mathrm{J}^r(k) \tag{22.5.2}$$

for all $p \in \varpi_r^{-1}(U)$. This gives $F^r(E)$ the structure of a principal $\mathrm{J}^r(k)$-bundle. ■

As mentioned in the beginning of this section, higher order frame bundles are constructed as generalizations of ordinary frame bundles, and so we expect to recover the ordinary frame bundle as a special case. This easily follows from our construction.

Theorem 22.5.2. *The first-order frame bundle $F^1(E)$ over a vector bundle E is isomorphic to the frame bundle $F(E)$.*

Proof. The existence of a bijection between $F^1(E)$ and $F(E)$ follows immediately from the definition of higher order frame bundles and the fact that bijective, linear maps from $\mathbb{R}^k$ to E_x for $x \in M$ are canonically identified with invertible first-order jets in $\tilde{J}^r_0(\mathbb{R}^k, E_x)_0$ and vice versa. This identification commutes with the right action of $\mathrm{J}^r(k) \cong \mathrm{GL}(k, \mathbb{R})$ and preserves the local trivializations. Hence, it constitutes a principal bundle isomorphism. ■

▶Discuss associated bundles.◀

22.6 Tangent frame bundle

A prototypical example of a vector bundle which exists on every manifold and is often encountered in physics is, of course, the tangent bundle. By applying the procedure detailed in the preceding section to construct its frame bundle, we obtain a principal bundle which likewise is canonically defined over any manifold. The definition is straightforward.

Definition 22.6.1 (Tangent frame bundle). Let M be a manifold of dimension $\dim M = n$. Its *tangent frame bundle* (or simply its *frame bundle*) is the frame bundle of its tangent bundle $\tau : TM \to M$. It is usually denoted $\mathrm{GL}(M)$ or FM.

In the following, it will be useful to introduce coordinates on the tangent frame bundle, similarly to the coordinates we have introduced on frames bundles in general. For this purpose, recall that any coordinates (x^μ) on the base manifold M induce a coordinate basis (∂_μ) on the tangent bundle TM, and so we can write a tangent vector in the form $\bar{x}^\mu \partial_\mu$, giving rise to coordinates $(x^\mu, \bar{x}^\mu)$ on TM. (This will be discussed in more detail in section 29.1.) Writing a frame $p \in F_x M$ with $x \in M$ as

$$\begin{array}{rccc} p : & \mathbb{R}^n & \to & T_x M \\ & v^i \mathbb{e}_i & \mapsto & p^\mu{}_i v^i \partial_\mu \end{array}\,, \tag{22.6.1}$$

we have thus introduced coordinates $(x^\mu, p^\mu{}_i)$ on FM, where $p^\mu{}_i$ constitute the components of an invertible matrix. Naturally, these coordinates also define a basis of the tangent bundle TFM, which we denote as

$$\partial_\mu = \frac{\partial}{\partial x^\mu}\,, \quad \bar{\partial}_\mu{}^i = \frac{\partial}{\partial p^\mu{}_i}\,, \tag{22.6.2}$$

and the dual basis

$$\mathrm{d}x^\mu\,, \quad \mathrm{d}p^\mu{}_i \tag{22.6.3}$$

of the cotangent bundle T^*FM.

To illustrate the usefulness of these coordinates, we use them to express the right translation and fundamental vector fields. Denoting $g \in \mathrm{GL}(n, \mathbb{R})$ by the components $g^i{}_j$ of an invertible matrix, the right translation reads

$$\begin{array}{cccc} R_g & : & FM & \to & FM \\ & & (x^\mu, p^\mu{}_i) & \mapsto & (x^\mu, p^\mu{}_j g^j{}_i) \end{array} . \tag{22.6.4}$$

Further, using the fact that the elements of $T_e\mathrm{GL}(n, \mathbb{R}) \cong \mathfrak{gl}(n, \mathbb{R})$ are real $n \times n$ matrices, we can denote a Lie algebra element

$$a = a^i{}_j \mathcal{H}_i{}^j\,, \tag{22.6.5}$$

and so the fundamental vector field $\tilde{a} \in \mathrm{Vect}(FM)$ takes the form

$$\tilde{a} = a^i{}_j \tilde{\mathcal{H}}_i{}^j = p^\mu{}_j a^j{}_i \bar{\partial}_\mu{}^i\,. \tag{22.6.6}$$

This expression will turn out to be useful also in later sections.

by its matrix components $a^i{}_j$ with respect to the basis $\mathcal{H}_i{}^j$. A whole theory has been developed around the question whether the frame bundle of a given manifold is trivial or not. Those manifolds whose frame bundle is trivial deserve an own name.

Definition 22.6.2 (Parallelizable manifold)**.** A manifold M whose frame bundle $\mathrm{GL}(M)$ is trivial is called *parallelizable*.

Definition 22.6.3 (Parallelization)**.** A global section $\sigma \in \Gamma(\mathrm{GL}(M))$ of the frame bundle $\mathrm{GL}(M)$ of a parallelizable manifold M is called a *parallelization* of M.

A nice case for studying this property is given in the following example.

> *Example* 22.6.1. The only spheres S^n that are parallelizable are S^1, S^3 and S^7. A Cartesian product of at least two spheres is parallelizable if and only if at least one of them is odd.

We will not prove this here, since the proof is highly non-trivial. However, we can nicely prove the following.

Theorem 22.6.1. *Every Lie group G is parallelizable.*

Proof. Let G be a Lie group of dimension $\dim G = n$ and $(X_1, \ldots, X_n)$ a basis of the Lie algebra $\mathfrak{g}$. Every basis element X_i is a left invariant vector field on G. For all $g \in G$, the vectors $(X_1(g), \ldots, X_n(g))$ constitute a basis of T_gG, and thus define a frame at g. This yields a global section of the frame bundle. It thus follows that the frame bundle is trivial, so that G is parallelizable. ■

As we have seen in the previous sections, there is a close relationship between frame bundles, their underlying vector bundles, as well as their dual and tensor product bundles. For tensor bundles of the tangent and cotangent bundle, we have seen in section 12.1 that one can define the notion of pullback along diffeomorphisms for their sections. Since we construct elements of the tangent frame bundle with the help of vectors, one may therefore expect a similar notion to exist also for (local) sections of the tangent frame bundle. This can be defined as follows.

Definition 22.6.4 (Pullback of a frame field). Let M, N be manifolds of dimension $\dim M = \dim N = n$ and $\varphi : M \to N$ a diffeomorphism. For a local section $\varepsilon : U \to FN$ of the frame bundle FN of N defined on $U \subset N$, we define the *pullback* of ε as the local section $\varphi^*\varepsilon : \varphi^{-1}(U) \to FM$ such that for every $x \in \varphi^{-1}(U)$ and $v \in \mathbb{R}^n$ holds

$$(\varphi^*\varepsilon)(x)(v) = \varphi_*^{-1}(\varepsilon(\varphi(x))(v)) . \tag{22.6.7}$$

There is another possibility to understand this construction. Recall that a tangent frame $p \in F_xM$ over $x \in M$ with $\dim M = n$ is a bijective, linear function $p : \mathbb{R}^n \to T_xM$. Given a diffeomorphism $\varphi : M \to N$, it follows from theorem 10.1.4 that $\varphi_*|_{T_xM} : T_xM \to T_{\varphi(x)}N$ is a vector space isomorphism, i.e., a bijective, linear function. Hence, also $\varphi_* \circ p : \mathbb{R}^n \to T_{\varphi(x)}N$ is bijective and linear, and thus a frame, $\varphi_* \circ p \in F_{\varphi(x)}N$. One may thus define the following function.

Definition 22.6.5 (Frame bundle lift). Let M, N be manifolds of dimension $\dim M = \dim N = n$ and $\varphi : M \to N$ a diffeomorphism. Its *frame bundle lift* is the map

$$\begin{array}{cccc} \varphi_\circ & : & FM & \to & FN \\ & & p & \mapsto & \varphi_\circ(p) = \varphi_* \circ p \end{array} . \tag{22.6.8}$$

It does not come as a surprise that the frame bundle lift satisfies the following property.

Theorem 22.6.2. *The frame bundle lift $\varphi_\circ : FM \to FN$ of a diffeomorphism $\varphi : M \to N$ between manifolds of dimension n is a principal $\mathrm{GL}(n, \mathbb{R})$-bundle isomorphism covering φ.*

Proof. ▶. . . ◀ ■

Given a left action $\rho : \mathrm{GL}(n, \mathbb{R}) \times F \to F$ on a manifold F, one may then proceed by carrying this bundle map to the corresponding associated bundles. This can be done by defining another notion as follows.

Definition 22.6.6 (Associated bundle lift). Let M, N be manifolds of dimension $\dim M = \dim N = n$ and $\varphi : M \to N$ a diffeomorphism. For a left action $\rho : \mathrm{GL}(n, \mathbb{R}) \times F \to F$ on a manifold F, one defines the *associated bundle lift* as the map

$$\begin{array}{cccc} \varphi_\rho & : & FM \times_\rho F & \to & FN \times_\rho F \\ & & [p, f]_\rho & \mapsto & [\varphi_\circ(p), f]_\rho \end{array} . \tag{22.6.9}$$

One needs to check that this is well-defined, i.e., independent of the choice of the representative $(p, f) \in [p, f]_\rho$. We will do so by proving an even stronger statement.

Theorem 22.6.3. *The associated bundle lift $\varphi_\rho : FM \times_\rho F \to FN \times_\rho F$ is a fiber bundle isomorphism covering φ, and a vector bundle isomorphism if $\rho : \mathrm{GL}(n, \mathbb{R}) \times F \to F$ is a linear representation on a vector space F.*

Proof. ▶...◀ ■

Among the most important actions of the group $\mathrm{GL}(n, \mathbb{R})$ we have encountered so far are the natural representation on $\mathbb{R}^n$, its dual and their tensor products and powers. We have seen in the previous sections that the associated bundles obtained from these actions are the tangent, cotangent and tensor bundles. Recall from section 12.1 that we have defined the pullback of sections of these bundles. One may expect that these notions are related; it turns out that this is the case, and we prove the following statement.

Theorem 22.6.4. *Given a diffeomorphism $\varphi : M \to N$ and a tensor field $T \in \Gamma(T^r_sN)$, the pullback $\varphi^*(T)$ and the lift $\varphi^r_s : T^r_sM \to T^r_sN$ are related by*

$$T \circ \varphi = \varphi^r_s \circ \varphi^*(T) . \tag{22.6.10}$$

Proof. ▶...◀ ■

Before we conclude this section, it is helpful to introduce another helpful notion, which exists on the tangent frame bundle. We use the following definition.

Definition 22.6.7 (Canonical one-form)**.** Let M be a manifold of dimension $\dim M = n$ and $\varpi : FM \to M$ the tangent frame bundle of M. The *canonical one-form* $\theta \in \Omega^1(FM, \mathbb{R}^n)$ is the $\mathbb{R}^n$-valued one-form on FM such that

$$\theta_p(\xi) = p^{-1}(\varpi_*(\xi)) \tag{22.6.11}$$

for all $p \in FM$ and $\xi \in T_pFM$.

This definition deserves a brief explanation. A $\mathbb{R}^n$-valued one-form on FM assigns to every tangent vector $\xi \in TFM$ an element of $\mathbb{R}^n$, and this assignment is linear on the fibers T_pFM for $p \in FM$. The canonical one-form does so by using the fact that a frame $p \in FM$ is a bijective, linear map from $\mathbb{R}^n$ to $T_{\varpi(p)}M$. Given $\xi \in T_pFM$, one can obtain a tangent vector $\varpi_*(\xi) \in T_{\varpi(p)}M$ to M. Then one has $p^{-1}(\varpi_*(\xi)) \in \mathbb{R}^n$. Since both p and ϖ_* are linear on every fibers, this also holds for their composition. It follows that θ as constructed above is indeed a $\mathbb{R}^n$-valued one-form on FM.

It is instructive to derive a coordinate expression for the canonical one-form. Denoting a tangent vector $\xi \in T_pFM$ as $\xi = \xi^\mu\partial_\mu + \bar{\xi}^\mu{}_i\bar{\partial}_\mu{}^i$, we have

$$\varpi_*(\xi) = \xi^\mu\partial_\mu \in T_{\varpi(p)}M , \tag{22.6.12}$$

and hence

$$\theta(\xi) = p^{-1}(\varpi_*(\xi)) = p^{-1\,i}{}_\mu\xi^\mu \mathbb{e}_i \in \mathbb{R}^n . \tag{22.6.13}$$

Hence, it follows that the canonical one-form is given by

$$\theta = p^{-1\,i}{}_\mu \mathrm{d}x^\mu \otimes \mathbb{e}_i \in \Omega^1(FM, \mathbb{R}^n) , \tag{22.6.14}$$

and we will also write

$$\theta^i = p^{-1\,i}{}_\mu \mathrm{d}x^\mu \tag{22.6.15}$$

for convenience.

Recall that the structure group $\mathrm{GL}(n, \mathbb{R})$ of FM acts on $\mathbb{R}^n$ via the natural representation $\rho = \mathrm{id}_{\mathrm{GL}(n,\mathbb{R})}$. Following definition 27.2.1, we now show the following.

Theorem 22.6.5. *The canonical form $\theta \in \Omega^1(FM, \mathbb{R}^n)$ on the general linear frame bundle FM of a manifold M of dimension n is a basic form of type $\rho = \mathrm{id}_{\mathrm{GL}(n,\mathbb{R})}$.*

Proof. We need to show that θ is both horizontal and equivariant. For the former, let $p \in FM$ and $\xi \in V_pFM$ a vertical tangent vector. By the definition of the canonical form we then have

$$\theta_p(\xi) = p^{-1}(\varpi_*(\xi)) = 0\,, \tag{22.6.16}$$

since $\varpi_*(\xi) = 0$ for a vertical tangent vector. Further, let $g \in \mathrm{GL}(n,\mathbb{R})$ and $\zeta \in T_pFM$. Then we have

$$\begin{aligned}(R_g^*\theta)_p(\zeta) &= \theta_{p\cdot g}(R_{g*}(\zeta))\\ &= (p\cdot g)^{-1}(\varpi_*(R_{g*}(\zeta)))\\ &= (g^{-1}\cdot p^{-1})(\varpi_*(\zeta))\\ &= g^{-1}(\theta_p(\zeta))\,,\end{aligned} \tag{22.6.17}$$

which shows that θ is equivariant with respect to $\rho = \mathrm{id}_{\mathrm{GL}(n,\mathbb{R})}$, and thus a basic form. ∎

This can also be illustrated using coordinates. Hence, we have

$$\theta_{p\cdot g} = {g^{-1}}^i{}_j {p^{-1}}^j{}_\mu \mathrm{d}x^\mu \otimes \mathbb{e}_i\,, \tag{22.6.18}$$

and also

$$(R_g^*\theta)_p = {g^{-1}}^i{}_j {p^{-1}}^j{}_\mu \mathrm{d}x^\mu \otimes \mathbb{e}_i\,, \tag{22.6.19}$$

since the action on the x^μ coordinates is trivial. Since g acts by

$$g \;:\; \begin{array}{ccc}\mathbb{R}^n & \to & \mathbb{R}^n\\ v^i\mathbb{e}_i & \mapsto & g^i{}_j v^j \mathbb{e}_i\end{array}\,, \tag{22.6.20}$$

we have

$$g^{-1}(\theta_p) = {g^{-1}}^i{}_j {p^{-1}}^j{}_\mu \mathrm{d}x^\mu \otimes \mathbb{e}_i = (R_g^*\theta)_p\,, \tag{22.6.21}$$

as expected.

Chapter 23

Densities

23.1 Density bundles and scalar densities

Let us now consider the following linear (left) actions of the general linear group $\mathrm{GL}(k, \mathbb{R})$ with $k \in \mathbb{N}$ on the vector space $F = \mathbb{R}$ of real numbers:

1. $\rho^+_w(A, c) = (\det A)^{-w} c$ for $w \in \mathbb{Z}$,
2. $\rho^-_w(A, c) = \operatorname{sgn}(\det A)(\det A)^{-w} c$ for $w \in \mathbb{Z}$,
3. $|\rho|^+_w(A, c) = |\det A|^{-w} c$ for $w \in \mathbb{R}$,
4. $|\rho|^-_w(A, c) = \operatorname{sgn}(\det A)|\det A|^{-w} c$ for $w \in \mathbb{R}$.

We add a few remarks:

- One might be worried that, except for the first one, the right hand sides of the expressions above are functions of $\det A$ which are non-differentiable and even discontinuous at $\det A = 0$. However, recall that $\det A \neq 0$ for $A \in \mathrm{GL}(k, \mathbb{R})$, and so all defined functions are smooth on $\mathrm{GL}(k, \mathbb{R}) \times \mathbb{R}$.
- Note that the former two actions are special cases of the latter two. Namely, if w is even, we have $\rho^+_w = |\rho|^+_w$ and $\rho^-_w = |\rho|^-_w$. If w is odd, the opposite holds, i.e., $\rho^+_w = |\rho|^-_w$ and $\rho^-_w = |\rho|^+_w$.

Consider now a vector bundle $\pi : E \to M$ of rank k over a manifold M, and recall that its frame bundle $F(E)$ is a principal $\mathrm{GL}(k, \mathbb{R})$-bundle. Hence, for any of the actions ρ listed above, we may define a corresponding associated bundle. This leads to the following definition:

Definition 23.1.1 (Densities)**.** Let $\pi : E \to M$ be a vector bundle and $F(E)$ its general linear frame bundle. For the representations listed above one defines the *density bundles* as follows:

1. the bundle of *authentic scalar densities* of weight $w \in \mathbb{Z}$:

$$D^+_w(E) = F(E) \times_{\rho^+_w} \mathbb{R}\,, \tag{23.1.1}$$

2. the bundle of *scalar pseudo-densities* of weight $w \in \mathbb{Z}$:

$$D^-_w(E) = F(E) \times_{\rho^-_w} \mathbb{R}\,, \tag{23.1.2}$$

3. the bundle of *even scalar densities* of weight $w \in \mathbb{R}$:

$$|D|^+_w(E) = F(E) \times_{|\rho|^+_w} \mathbb{R}\,, \tag{23.1.3}$$

4. the bundle of *odd scalar densities* of weight $w \in \mathbb{R}$:

$$|D|^-_w(E) = F(E) \times_{|\rho|^-_w} \mathbb{R}\,. \tag{23.1.4}$$

We call a section of a density bundle as defined above a *scalar density* of weight s.

In the literature, one finds also different names for these objects, some of them contradicting the ones we use here:

- In [Spi99, vol. I, ch. 4], the following nomenclature is used:
  1. Sections of D^+_w are called *even relative scalars*.
  2. Sections of $|D|^+_w$ are called *odd relative scalars*.
  3. Sections of D^+_1 are called *even scalar densities*.
  4. Sections of $|D|^+_1$ are called *odd scalar densities*.

There are a few special cases that deserve particular attention. We start with the following case.

Theorem 23.1.1. *The bundle $D^+_0(E) = |D|^+_0(E)$ of even (authentic) scalar densities of weight 0 is canonically isomorphic to the trivial line bundle $M \times \mathbb{R}$.*

Proof. Recall that an element of the associated bundle is given by an equivalence class

$$[p, c]_{\rho^+_0} = \{[p \cdot g, \rho^+_0(g^{-1}, c)], g \in G\} = \{[p \cdot g, c], g \in G\}\,, \tag{23.1.5}$$

where $p \in F(E)$, $c \in \mathbb{R}$ and $G = \mathrm{GL}(k, \mathbb{R})$. We see that there is a one-to-one correspondence between such equivalence classes and pairs $(x, c) \in M \times \mathbb{R}$, where $x = \pi(p \cdot g) = \pi(p) \in M$. ∎

Hence, any (even, authentic) scalar density of weight 0 may simply be identified with a real function. Another important case is the following.

Theorem 23.1.2. *The bundle $D^+_{-1}(E) = |D|^-_{-1}(E)$ is canonically isomorphic to the exterior power bundle $\Lambda^k E$.*

Proof. Recall that an element of the associated bundle is given by an equivalence class

$$[p, c]_{\rho^+_{-1}} = \{[p \cdot g, \rho^+_{-1}(g^{-1}, c)], g \in G\} = \{[p \cdot g, (\det g)^{-1}c], g \in G\}\,, \tag{23.1.6}$$

where $p \in F(E)$, $c \in \mathbb{R}$ and $G = \mathrm{GL}(k, \mathbb{R})$. We then construct an element

$$cp_1 \wedge p_2 \wedge \ldots \wedge p_k \in \Lambda^k E\,. \tag{23.1.7}$$

To see that this is independent of the representative, we choose a different representative given by $p' = p \cdot g$ and $c' = (\det g)^{-1}c$ and calculate

$$\begin{aligned} c'p'_1 \wedge p'_2 \wedge \ldots \wedge p'_k &= (\det g)^{-1}c(p_{i_1}g^{i_1}{}_1) \wedge \ldots \wedge (p_{i_k}g^{i_k}{}_k) \\ &= (\det g)^{-1}cp_1 \wedge \ldots \wedge p_k \sum_{\sigma \in S_k}\left(\operatorname{sgn}\sigma \prod_{i=1}^k g^{\sigma(i)}{}_i\right) \\ &= cp_1 \wedge p_2 \wedge \ldots \wedge p_k\,. \end{aligned} \tag{23.1.8}$$

To get the second line from the first, we used the fact that the exterior product is totally antisymmetric, and so we could reorder the terms such that the basis elements p_i are in canonical order. The third line is obtained by realizing that the sum appearing in the second line is just the determinant formula in terms of matrix components. ■

Also densities of weight 1 we already encountered, despite in a different form, which we will see as follows.

Theorem 23.1.3. *The bundle $D_1^+(E) = |D|_1^-(E)$ is canonically isomorphic to the dual exterior power bundle $\Lambda^k E^*$.*

Proof. In this case an element of the associated bundle is given by an equivalence class

$$[p, c]_{\rho_1^+} = \{[p \cdot g, \rho_1^+(g^{-1}, c)], g \in G\} = \{[p \cdot g, (\det g)c], g \in G\}\,, \tag{23.1.9}$$

where $p \in F(E)$, $c \in \mathbb{R}$ and $G = \mathrm{GL}(k, \mathbb{R})$. Now we construct an element

$$cp^{*1} \wedge p^{*2} \wedge \ldots \wedge p^{*k} \in \Lambda^k E^*\,, \tag{23.1.10}$$

where p^* denotes the basis of $E^*_{\pi(p)}$ dual to p. Similarly to the previous case, one must show that this is independent of the representative. Here we choose a different representative given by $p' = p \cdot g$ and $c' = (\det g)c$ and calculate

$$\begin{aligned} c'p'^{*1} \wedge p'^{*2} \wedge \ldots \wedge p'^{*k} &= (\det g)c[(g^{-1})^1{}_{i_1} p^{*i_1}] \wedge \ldots \wedge [(g^{-1})^k{}_{i_k} p^{*i_k}] \\ &= (\det g)cp^{*1} \wedge \ldots \wedge p^{*k} \sum_{\sigma \in S_k} \left(\operatorname{sgn} \sigma \prod_{i=1}^k (g^{-1})^i{}_{\sigma(i)}\right) \\ &= cp^{*1} \wedge p^{*2} \wedge \ldots \wedge p^{*k}\,, \end{aligned} \tag{23.1.11}$$

where we essentially used the same steps as in the preceding proof. ■

We see that the bundles $\Lambda^k E$ and $\Lambda^k E^*$ can hence be regarded as particular cases of density bundles. Note that also these two bundles are dual to each other. One may therefore ask whether also for the duals of other density bundles one can find an explicit rule. This is the case, and we state it as follows.

Theorem 23.1.4. *The dual vector bundles of the scalar density bundles are given by:*

$$[D_w^+(E)]^* = D_w^+(E^*) = D_{-w}^+(E)\,, \qquad [|D|_w^+(E)]^* = |D|_w^+(E^*) = |D|_{-w}^+(E)\,, \tag{23.1.12a}$$

$$[D_w^-(E)]^* = D_w^-(E^*) = D_{-w}^-(E)\,, \qquad [|D|_w^-(E)]^* = |D|_w^-(E^*) = |D|_{-w}^-(E)\,. \tag{23.1.12b}$$

Proof. From the relation $\det A^{-1} = (\det A)^{-1}$ of the determinant and the fact that the transpose is trivial for a one-dimensional representation follows that the dual representations are given by

$$(\rho_w^+)^*(A, c) = \rho_w^+(A^{-1}, c) = (\det A)^w c = \rho_{-w}^+(A, c)\,, \tag{23.1.13a}$$

$$(\rho_w^-)^*(A, c) = \rho_w^-(A^{-1}, c) = \operatorname{sgn}(\det A)(\det A)^w c = \rho_{-w}^-(A, c)\,, \tag{23.1.13b}$$

$$(|\rho|_w^+)^*(A, c) = |\rho|_w^+(A^{-1}, c) = |\det A|^w c = |\rho|_{-w}^+(A, c)\,, \tag{23.1.13c}$$

$$(|\rho|_w^-)^*(A, c) = |\rho|_w^-(A^{-1}, c) = \operatorname{sgn}(\det A)|\det A|^w c = |\rho|_{-w}^-(A, c)\,. \tag{23.1.13d}$$

Using theorem 20.4.2, we therefore find the relations for the dual bundles. ►Show relations for E^*.◄ ■

Theorem 23.1.5. *The tensor product between two scalar density bundles satisfies the following rules:*

$$D_w^+(E) \otimes D_{w'}^+(E) \cong D_w^-(E) \otimes D_{w'}^-(E) \cong D_{w+w'}^+(E)\,, \tag{23.1.14a}$$

$$D^-_w(E) \otimes D^+_{w'}(E) \cong D^+_w(E) \otimes D^-_{w'}(E) \cong D^-_{w+w'}(E)\,, \tag{23.1.14b}$$
$$|D|^+_w(E) \otimes |D|^+_{w'}(E) \cong |D|^-_w(E) \otimes |D|^-_{w'}(E) \cong |D|^+_{w+w'}(E)\,, \tag{23.1.14c}$$
$$|D|^-_w(E) \otimes |D|^+_{w'}(E) \cong |D|^+_w(E) \otimes |D|^-_{w'}(E) \cong |D|^-_{w+w'}(E)\,. \tag{23.1.14d}$$

Proof. Since the representations involved in the construction of these bundles are one-dimensional, their tensor product reduces to the ordinary product. For these we find the relations

$$(\rho^+_w \otimes \rho^+_{w'})(A,c) = (\rho^-_w \otimes \rho^-_{w'})(A,c) = (\det A)^{-w-w'}c = \rho^+_{w+w'}(A,c)\,, \tag{23.1.15a}$$
$$(\rho^-_w \otimes \rho^+_{w'})(A,c) = (\rho^+_w \otimes \rho^-_{w'})(A,c) = \operatorname{sgn}(\det A)(\det A)^{-w-w'}c = \rho^-_{w+w'}(A,c)\,, \tag{23.1.15b}$$
$$(|\rho|^+_w \otimes |\rho|^+_{w'})(A,c) = (|\rho|^-_w \otimes |\rho|^-_{w'})(A,c) = |\det A|^{-w-w'}c = |\rho|^+_{w+w'}(A,c)\,, \tag{23.1.15c}$$
$$(|\rho|^-_w \otimes |\rho|^+_{w'})(A,c) = (|\rho|^+_w \otimes |\rho|^-_{w'})(A,c) = \operatorname{sgn}(\det A)|\det A|^{-w-w'}c = |\rho|^-_{w+w'}(A,c)\,. \tag{23.1.15d}$$

The corresponding relations for the density bundles then follow from theorem 20.4.2. ■

Since the density bundles are vector bundles, we can describe their elements and sections by introducing a basis on their fibers. This basis consists of only one (non-zero) element of each fiber, since these bundles are of rank 1. Note that such a basis element is not always canonically defined, except in special cases, such as $D^+_0(E)$. However, given a basis of E, it can be constructed by using the defining properties of the associated bundle construction. Note that a local basis defined on an open set $U \subset M$ is simply a local section $e : U \to F(E)$ of the frame bundle. This allows us to define a local, nowhere vanishing section

$$\begin{array}{rcccc} [e,1]_{|\rho|^\pm_w} & : & U & \to & |D|^\pm_w \\ & & x & \mapsto & [e(x),1]_{|\rho|^\pm_w} \end{array} \tag{23.1.16}$$

of any of the density bundles we introduced. Since these bundles are one-dimensional, this section constitutes a local basis. We will denote these bases as

$$e^\pm_w = [e,1]_{\rho^\pm_w}\,, \quad |e|^\pm_w = [e,1]_{|\rho|^\pm_w}\,. \tag{23.1.17}$$

To demonstrate the use of these bases, we apply them to write out the bundle isomorphisms whose existence we proved above. Using the identifications from theorems 23.1.1, 23.1.2 and 23.1.3 and setting $p = e(x)$ for $x \in U$, we see that the bundle isomorphisms simply map the bases as

$$e^+_0 \cong (\bullet \mapsto 1)\,, \quad e^+_1 \cong e^{*1} \wedge \ldots \wedge e^{*k}\,, \quad e^+_{-1} \cong e_1 \wedge \ldots \wedge e_k\,, \tag{23.1.18}$$

where by $\bullet \mapsto 1$ we denoted the constant function on U which has the value 1 everywhere. Also it is helpful to note that the dual bases of the dual bundles, following theorem 23.1.4, are given by

$$(e^+_w)^* = e^+_{-w}\,, \quad (e^-_w)^* = e^-_{-w}\,, \quad (|e|^+_w)^* = |e|^+_{-w}\,, \quad (|e|^-_w)^* = |e|^-_{-w}\,. \tag{23.1.19}$$

Finally, following theorem 23.1.5, the tensor products of these bases satisfy

$$e^+_w \otimes e^+_{w'} = e^-_w \otimes e^-_{w'} = e^+_{w+w'}\,, \tag{23.1.20a}$$
$$e^-_w \otimes e^+_{w'} = e^+_w \otimes e^-_{w'} = e^-_{w+w'}\,, \tag{23.1.20b}$$
$$|e|^+_w \otimes |e|^+_{w'} = |e|^-_w \otimes |e|^-_{w'} = |e|^+_{w+w'}\,, \tag{23.1.20c}$$
$$|e|^-_w \otimes |e|^+_{w'} = |e|^+_w \otimes |e|^-_{w'} = |e|^-_{w+w'}\,. \tag{23.1.20d}$$

We will not prove these formulas here, but make use of them later.

We finally discuss the question how the bases of density bundles are related which are derived from different bases of the underlying vector bundle E. We thus consider a new basis $\tilde{e}$ defined on the same open subset $U \subset M$ as above, which is related to e by a basis transformation of the form

$$\tilde{e} = e \cdot g\,, \quad \tilde{e}_a(x) = e_b(x) g^b{}_a(x) \tag{23.1.21}$$

for all $x \in U$, where $g : U \to \mathrm{GL}(k, \mathbb{R})$ defines the basis transformation. One then easily derives the transformations

$$\begin{aligned} \tilde{e}^+_w &= [\tilde{e}, 1]_{\rho^+_w} \\ &= [e \cdot g, 1]_{\rho^+_w} \\ &= [e, \rho^+_w(g, 1)]_{\rho^+_w} \\ &= [e, (\det g)^{-w}]_{\rho^+_w} \\ &= (\det g)^{-w}[e, 1]_{\rho^+_w} \\ &= (\det g)^{-w} e^+_w\,, \end{aligned} \tag{23.1.22}$$

and in the same way also

$$\tilde{e}^-_w = \operatorname{sgn}(\det g)(\det g)^{-w} e^-_w\,, \quad |\tilde{e}|^+_w = |\det g|^{-w}|e|^+_w\,, \quad |\tilde{e}|^-_w = \operatorname{sgn}(\det g)|\det g|^{-w}|e|^-_w\,. \tag{23.1.23}$$

Using these transformation formulas, we can now answer the question how the component expression of a scalar density depends on the choice of the basis. We find the following formulas:

$$\mathfrak{q} = q e^+_w = \tilde{q}\tilde{e}^+_w \in \Gamma(D^+_w(E))\,, \qquad \tilde{q} = (\det g)^w q \tag{23.1.24a}$$

$$\mathfrak{t} = t e^-_w = \tilde{t}\tilde{e}^-_w \in \Gamma(D^-_w(E))\,, \qquad \tilde{t} = \operatorname{sgn}(\det g)(\det g)^w t \tag{23.1.24b}$$

$$\mathfrak{u} = u|e|^+_w = \tilde{u}|\tilde{e}|^+_w \in \Gamma(|D|^+_w(E))\,, \qquad \tilde{u} = |\det g|^w u \tag{23.1.24c}$$

$$\mathfrak{v} = v|e|^-_w = \tilde{v}|\tilde{e}|^-_w \in \Gamma(|D|^-_w(E))\,, \qquad \tilde{v} = \operatorname{sgn}(\det g)|\det g|^w v\,. \tag{23.1.24d}$$

We see that the component of a scalar density of weight w with respect to a basis changes under a basis transformation with the determinant of the transformation matrix to the power w. Note that this is the reason for introducing the negative sign in the definition of the representations $|\rho|^\pm_w$ at the beginning of this chapter.

23.2 Pseudotensors and tensor densities

Given a vector bundle $\pi : E \to M$, we have now found essentially two types of vector bundles which we can obtain as associated bundles of the frame bundle $F(E)$: the tensor product bundles E^r_s discussed in chapter 4, which are associated to the frame bundle by the tensor product of the canonical representation of the structure group $\mathrm{GL}(k, \mathbb{R})$ and its dual, as well as the scalar density bundles shown in section 23.1. We may combine these two notions, and thus obtain another important type of bundle by taking their tensor product. We then arrive at the following definition.

Definition 23.2.1 (Tensor densities). Let $\pi : E \to M$ be a vector bundle and the density bundles as given in definition 23.1.1. For $r, s \in \mathbb{N}$ we define the following bundles:

1. the bundle of *authentic tensor densities* of weight $w \in \mathbb{Z}$:

$$D^+_w(E) \otimes E^r_s\,, \tag{23.2.1}$$

2. the bundle of *tensor pseudo-densities* of weight $w \in \mathbb{Z}$:

$$D^-_w(E) \otimes E^r_s\,, \tag{23.2.2}$$

3. the bundle of *even tensor densities* of weight $w \in \mathbb{R}$:

$$|D|^+_w(E) \otimes E^r_s\,, \tag{23.2.3}$$

4. the bundle of *odd tensor densities* of weight $w \in \mathbb{R}$:

$$|D|^-_w(E) \otimes E^r_s\,. \tag{23.2.4}$$

We call a section of a density bundle as defined above a *tensor density* of weight w.

In the literature one also finds other conventions for naming the objects defined above, as it is also the case for the scalar densities we encountered before:

- In [HO01, sec. A.1.8], sections of $D^+_w(E) \otimes E^r_s$ are called *tensor densities*, while sections of $D^-_w(E) \otimes E^r_s$ are called *twisted tensor densities*.
- In [Spi99, vol. I, ch. 4], sections of $D^+_w(E) \otimes E^r_s$ are called *even relative tensors*, while sections of $|D|^+_w(E) \otimes E^r_s$ are called *odd relative tensors*.

Two special cases which we previously discussed are easily recovered:

1. For $w = 0$ and the positive sign, i.e., $D^+_0 = |D|^+_0$, one obtains an ordinary tensor bundle.
2. For $r = s = 0$ one obtains a scalar density.

Since also the tensor density bundles are vector bundles, one may of course discuss their duals and tensor products. These are straightforward to proof, and so we will summarize their most important properties in the following few statements.

Theorem 23.2.1. *The dual vector bundles of the tensor density bundles are given by:*

$$[D^+_w(E) \otimes E^r_s]^* \cong D^+_{-w}(E) \otimes E^s_r\,, \tag{23.2.5a}$$
$$[|D|^+_w(E) \otimes E^r_s]^* \cong |D|^+_{-w}(E) \otimes E^s_r\,, \tag{23.2.5b}$$
$$[D^-_w(E) \otimes E^r_s]^* \cong D^-_{-w}(E) \otimes E^s_r\,, \tag{23.2.5c}$$
$$[|D|^-_w(E) \otimes E^r_s]^* \cong |D|^-_{-w}(E) \otimes E^s_r\,. \tag{23.2.5d}$$

Proof. This follows immediately from theorems 4.3.2, 4.3.4 and 23.1.4. ■

Theorem 23.2.2. *The tensor product between two tensor density bundles satisfies the following rules:*

$$D^+_w(E) \otimes D^+_{w'}(E) \otimes E^r_s \cong D^-_w(E) \otimes D^-_{w'}(E) \otimes E^r_s \cong D^+_{w+w'}(E) \otimes E^s_r\,, \tag{23.2.6a}$$
$$D^-_w(E) \otimes D^+_{w'}(E) \otimes E^r_s \cong D^+_w(E) \otimes D^-_{w'}(E) \otimes E^r_s \cong D^-_{w+w'}(E) \otimes E^s_r\,, \tag{23.2.6b}$$
$$|D|^+_w(E) \otimes |D|^+_{w'}(E) \otimes E^r_s \cong |D|^-_w(E) \otimes |D|^-_{w'}(E) \otimes E^r_s \cong |D|^+_{w+w'}(E) \otimes E^s_r\,, \tag{23.2.6c}$$
$$|D|^-_w(E) \otimes |D|^+_{w'}(E) \otimes E^r_s \cong |D|^+_w(E) \otimes |D|^-_{w'}(E) \otimes E^r_s \cong |D|^-_{w+w'}(E) \otimes E^s_r\,. \tag{23.2.6d}$$

Proof. This follows from theorem 23.1.5. ■

Theorem 23.2.3. *The contraction tr^k_l of a tensor density is a tensor density of the same weight and parity.*

Proof. ►...◄ ■

From the definition 23.2.1 one can easily derive how to construct bases for the tensor density bundles, given a basis e on the underlying vector bundle. These are of the form

$$|e|^{\pm}_{w} \otimes e_{a_1} \otimes \ldots \otimes e_{a_r} \otimes e^{*b_1} \otimes \ldots \otimes e^{*b_s} \,, \tag{23.2.7}$$

so that we can write a tensor density in the form

$$\mathfrak{T} = \mathfrak{T}^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} |e|^{\pm}_{w} \otimes e_{a_1} \otimes \ldots \otimes e_{a_r} \otimes e^{*b_1} \otimes \ldots \otimes e^{*b_s} \,, \tag{23.2.8}$$

where the first factor $|e|^{\pm}_{w}$ must be chosen such that it corresponds to the weight and parity of the tensor density.

We finally pose the question how the components of a tensor density change under a basis transformation of the form (23.1.21), which we already considered for scalar densities. Using the basis transformation

$$\tilde{e}_a = e_b g^b{}_a \,, \quad \tilde{e}^{*a} = (g^{-1})^a{}_b e^{*b} \,, \tag{23.2.9}$$

as well as the transformation rules (23.1.24) for the basis transformation of scalar densities derived in the section 23.1, we find

$$\tilde{\mathfrak{T}}^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} = \begin{Bmatrix} (\det g)^w \\ \operatorname{sgn}(\det g)(\det g)^w \\ |\det g|^w \\ \operatorname{sgn}(\det g)|\det g|^w \end{Bmatrix} (g^{-1})^{a_1}{}_{c_1} \cdots (g^{-1})^{a_r}{}_{c_r} g^{d_1}{}_{b_1} \cdots g^{d_s}{}_{b_s} \mathfrak{T}^{c_1 \cdots c_r}{}_{d_1 \cdots d_s} \,, \tag{23.2.10}$$

where the factor in braces depends on the parity of the tensor density.

23.3 Canonical tensor densities

The vector bundle isomorphisms shown in theorems 23.1.1, 23.1.2 and 23.1.3, which allow the identifications of certain density bundles with other canonically constructed one-dimensional vector bundles, as well as the tensor product rules in theorem 23.1.5, allow the construction of a number of tensor densities, which are canonically defined for any vector bundle $\pi : E \to M$. From these follow the canonical isomorphism

$$D^{+}_{1}(E) \otimes \Lambda^k E \cong D^{+}_{-1}(E) \otimes \Lambda^k E^* \cong D^{+}_{0}(E) \cong \Lambda^0 E \cong M \times \mathbb{R} \tag{23.3.1}$$

with the trivial line bundle. Further, the latter has a canonical, nowhere vanishing section, namely the constant function $\bullet \mapsto 1$, which is an element of $C^{\infty}(M, \mathbb{R}) \cong \Gamma(M \times \mathbb{R})$. Via the canonical isomorphism (23.3.1) one can therefore find canonical, nowhere vanishing sections also for the two listed bundles of tensor densities. These are defined as follows.

Definition 23.3.1 (Levi-Civita densities)**.** Let $\pi : E \to M$ be a vector bundle of rank k. The *Levi-Civita densities* on E are the tensor densities $\mathfrak{E} \in \Gamma(D^{+}_{1}(E) \otimes \Lambda^k E)$ and $\mathfrak{e} \in \Gamma(D^{+}_{-1}(E) \otimes \Lambda^k E^*)$ which are obtained by identifying the function $\bullet \mapsto 1$ on M via the bundle isomorphisms (23.3.1).

It is helpful to write these sections using the bases (23.1.17) we derived in section 23.1, given a local basis e of E. We start with $\mathfrak{E}$, which we identify as

$$\begin{aligned}\mathfrak{E} &= (\bullet \mapsto 1)\\ &= e_0^+\\ &= e_1^+ \otimes e_{-1}^+\\ &= e_1^+ \otimes (e_1 \wedge \ldots \wedge e_k)\\ &= \frac{1}{k!}\epsilon^{a_1\cdots a_k} e_1^+ \otimes (e_{a_1} \wedge \ldots \wedge e_{a_k})\\ &= \epsilon^{a_1\cdots a_k} e_1^+ \otimes e_{a_1} \otimes \ldots \otimes e_{a_k}\,,\end{aligned} \tag{23.3.2}$$

where $\epsilon^{a_1\cdots a_k}$ is the Levi-Civita symbol, which is totally antisymmetric in its indices. It takes the values 1 if $(a_1, \ldots, a_k)$ are an even permutation of $(1, \ldots, k)$, -1 for an odd permutation and 0 otherwise. It is a remarkable fact that the components of $\mathfrak{E}$, which is a tensor density of rank $(k, 0)$ and weight 1, are the same in *any* basis e, which can be seen from the fact that we did not use any properties of e in order to derive its component expression. We can proceed analogously with $\mathfrak{e}$ and find

$$\begin{aligned}\mathfrak{e} &= (\bullet \mapsto 1)\\ &= e_0^+\\ &= e_{-1}^+ \otimes e_1^+\\ &= e_{-1}^+ \otimes (e^{*1} \wedge \ldots \wedge e^{*k})\\ &= \frac{1}{k!}\epsilon_{a_1\cdots a_k} e_{-1}^+ \otimes (e^{*a_1} \wedge \ldots \wedge e^{*a_k})\\ &= \epsilon_{a_1\cdots a_k} e_{-1}^+ \otimes e^{*a_1} \otimes \ldots \otimes e^{*a_k}\,,\end{aligned} \tag{23.3.3}$$

where $\epsilon_{a_1\cdots a_k}$ takes the same values as $\epsilon^{a_1\cdots a_k}$. Also the components of $\mathfrak{e}$, which is a tensor density of rank $(0, k)$ and weight -1, are the same in *any* basis e.

23.4 Determinant of tensor densities

Using the canonical Levi-Civita densities defined in the previous section, as well as the relations between different tensor bundles, we can define another useful operation acting on second rank (covariant, contravariant or mixed) tensors and tensor densities, which is analogue to a similarly defined operation in linear algebra, namely that of a determinant. While in linear algebra the determinant of a matrix is simple a number, the intuitive expectation that the determinant of a second rank tensor is a scalar function, does *not* hold true. Instead, it turns out to be a tensor *density*. The reason for this false intuition comes from the fact that the determinant in linear algebra is commonly defined in terms of the totally antisymmetric Levi-Civita symbol, whose entries are numbers, but the analogue totally antisymmetric object in differential geometry is given by the Levi-Civita densities. This becomes clear in the following definition.

Definition 23.4.1 (Determinant of a tensor density)**.** Let $\pi : E \to M$ be a vector bundle of rank k. For a second rank tensor density on E, the *determinant* is defined as follows:

1. For a covariant tensor density $\mathfrak{T} \in \Gamma(|D|_w^{\pm}(E) \otimes E_2^0)$:

$$\det \mathfrak{T} = \frac{1}{k!} \operatorname{tr}_1^1 \operatorname{tr}_3^2 \cdots \operatorname{tr}_{2k-1}^k \operatorname{tr}_2^{k+1} \operatorname{tr}_4^{k+2} \cdots \operatorname{tr}_{2k}^{2k} \mathfrak{E} \otimes \mathfrak{E} \otimes \underbrace{\mathfrak{T} \otimes \ldots \otimes \mathfrak{T}}_{k \text{ times}}\,. \tag{23.4.1}$$

2. For a contravariant tensor density $\mathfrak{U} \in \Gamma(|D|_w^{\pm}(E) \otimes E_0^2)$:

$$\det \mathfrak{U} = \frac{1}{k!} \operatorname{tr}_1^1 \operatorname{tr}_2^3 \cdots \operatorname{tr}_k^{2k-1} \operatorname{tr}_{k+1}^2 \operatorname{tr}_{k+2}^4 \cdots \operatorname{tr}_{2k}^{2k} \mathfrak{e} \otimes \mathfrak{e} \otimes \underbrace{\mathfrak{U} \otimes \ldots \otimes \mathfrak{U}}_{k \text{ times}}\,. \tag{23.4.2}$$

3. For a mixed tensor density $\mathfrak{V} \in \Gamma(|D|^{\pm}_{w}(E) \otimes E^1_1)$:

$$\det \mathfrak{V} = \frac{1}{k!} \operatorname{tr}_1^{k+1} \operatorname{tr}_2^{k+2} \cdots \operatorname{tr}_k^{2k} \operatorname{tr}_{k+1}^1 \operatorname{tr}_{k+2}^2 \cdots \operatorname{tr}_{2k}^k \mathfrak{e} \otimes \mathfrak{E} \otimes \underbrace{\mathfrak{V} \otimes \ldots \otimes \mathfrak{V}}_{k \text{ times}} . \tag{23.4.3}$$

To see that the obtained determinants are scalar densities and calculate their weight and parity, we use theorem 23.2.2 on the tensor product of tensor densities, as well as theorem 23.2.3 on their contraction. With the notations of definition 23.4.1 we thus find the following.

1. For a covariant tensor density $\mathfrak{T} \in \Gamma(|D|^{\pm}_{w}(E) \otimes E^0_2)$ we have

$$\underbrace{\mathfrak{T} \otimes \ldots \otimes \mathfrak{T}}_{k \text{ times}} \in \Gamma\left(|D|^{\pm^k}_{kw}(E) \otimes E^0_{2k}\right) . \tag{23.4.4}$$

   Together with the Levi-Civita densities this yields

$$\mathfrak{E} \otimes \mathfrak{E} \otimes \underbrace{\mathfrak{T} \otimes \ldots \otimes \mathfrak{T}}_{k \text{ times}} \in \Gamma\left(|D|^{\pm^k}_{kw+2}(E) \otimes E^{2k}_{2k}\right) . \tag{23.4.5}$$

   Finally, after contraction, we find $\det \mathfrak{T} \in \Gamma\left(|D|^{\pm^k}_{kw+2}(E)\right)$.

2. For a contravariant tensor density $\mathfrak{U} \in \Gamma(|D|^{\pm}_{w}(E) \otimes E^2_0)$ we have

$$\underbrace{\mathfrak{U} \otimes \ldots \otimes \mathfrak{U}}_{k \text{ times}} \in \Gamma\left(|D|^{\pm^k}_{kw}(E) \otimes E^{2k}_{0}\right) . \tag{23.4.6}$$

   Together with the Levi-Civita densities this yields

$$\mathfrak{e} \otimes \mathfrak{e} \otimes \underbrace{\mathfrak{U} \otimes \ldots \otimes \mathfrak{U}}_{k \text{ times}} \in \Gamma\left(|D|^{\pm^k}_{kw-2}(E) \otimes E^{2k}_{2k}\right) . \tag{23.4.7}$$

   Finally, after contraction, we find $\det \mathfrak{U} \in \Gamma\left(|D|^{\pm^k}_{kw-2}(E)\right)$.

3. For a mixed tensor density $\mathfrak{V} \in \Gamma(|D|^{\pm}_{w}(E) \otimes E^1_1)$ we have

$$\underbrace{\mathfrak{V} \otimes \ldots \otimes \mathfrak{V}}_{k \text{ times}} \in \Gamma\left(|D|^{\pm^k}_{kw}(E) \otimes E^{k}_{k}\right) . \tag{23.4.8}$$

   Together with the Levi-Civita densities this yields

$$\mathfrak{e} \otimes \mathfrak{E} \otimes \underbrace{\mathfrak{V} \otimes \ldots \otimes \mathfrak{V}}_{k \text{ times}} \in \Gamma\left(|D|^{\pm^k}_{kw}(E) \otimes E^{2k}_{2k}\right) . \tag{23.4.9}$$

   Finally, after contraction, we find $\det \mathfrak{V} \in \Gamma\left(|D|^{\pm^k}_{kw}(E)\right)$.

We see that the weight depends on the type of the original tensor density, due to the contribution from the Levi-Civita densities. To give a more intuitive formula for the tensor contractions, we also give the coordinate formulas with respect to a local basis e of E below, making use of the product (23.1.20) of the basis elements.

1. A covariant tensor density $\mathfrak{T} \in \Gamma(|D|^{\pm}_{w}(E) \otimes E^0_2)$ can be written as

$$\mathfrak{T} = \mathfrak{T}_{ab}|e|^{\pm}_{w} \otimes e^{*a} \otimes e^{*b} . \tag{23.4.10}$$

   Its determinant is given by

$$\det \mathfrak{T} = \frac{1}{k!} \epsilon^{a_1 \cdots a_k} \epsilon^{b_1 \cdots b_k} \mathfrak{T}_{a_1 b_1} \cdots \mathfrak{T}_{a_k b_k} |e|^{\pm^k}_{kw+2} = \det \mathfrak{T}_{\bullet\bullet} |e|^{\pm^k}_{kw+2} . \tag{23.4.11}$$

2. A contravariant tensor density $\mathfrak{U} \in \Gamma(|D|^{\pm}_{w}(E) \otimes E^2_0)$ can be written as

$$\mathfrak{U} = \mathfrak{U}^{ab}|e|^{\pm}_{w} \otimes e_a \otimes e_b\,. \tag{23.4.12}$$

Its determinant is given by

$$\det \mathfrak{U} = \frac{1}{k!}\epsilon_{a_1\cdots a_k}\epsilon_{b_1\cdots b_k}\mathfrak{U}^{a_1b_1}\cdots\mathfrak{U}^{a_kb_k}|e|^{\pm^k}_{kw-2} = \det \mathfrak{U}^{\bullet\bullet}|e|^{\pm^k}_{kw-2}\,. \tag{23.4.13}$$

3. A mixed tensor density $\mathfrak{V} \in \Gamma(|D|^{\pm}_{w}(E) \otimes E^1_1)$ can be written as

$$\mathfrak{V} = \mathfrak{V}^{a}{}_{b}|e|^{\pm}_{w} \otimes e_a \otimes e^{*b}\,. \tag{23.4.14}$$

Its determinant is given by

$$\det \mathfrak{V} = \frac{1}{k!}\epsilon_{a_1\cdots a_k}\epsilon^{b_1\cdots b_k}\mathfrak{V}^{a_1}{}_{b_1}\cdots\mathfrak{V}^{a_k}{}_{b_k}|e|^{\pm^k}_{kw} = \det \mathfrak{V}^{\bullet}{}_{\bullet}|e|^{\pm^k}_{kw}\,. \tag{23.4.15}$$

Note that we denoted by $\det \mathfrak{T}, \det \mathfrak{U}, \det \mathfrak{V}$ the determinant of the tensor density, which is a scalar density, following definition 23.4.1, i.e., a section of a density bundle. This is different from most other literature, where this notation is instead used to denote the determinant of the *component expression* $\mathfrak{T}_{ab}, \mathfrak{U}^{ab}, \mathfrak{V}^{a}{}_{b}$ with respect to some given basis e, understood as a matrix. To distinguish these different objects, we use the notation $\det \mathfrak{T}_{\bullet\bullet}, \det \mathfrak{U}^{\bullet\bullet}, \det \mathfrak{V}^{\bullet}{}_{\bullet}$ for the latter, to indicate that these are components with respect to a basis.

For matrices, we are used to the fact that we can obtain the determinant of a product of matrices as the product of their individual determinants. A similar statement also holds for the determinant of tensor densities. We will thus show the following statement.

Theorem 23.4.1. *The determinant of the product of two densities is given by the product of their determinants.*

Proof. ►. . . ◄ ■

23.5 Densities in the tangent bundle

If the vector bundle we consider is the tangent bundle $\tau : TM \to M$, so that the frame bundle becomes the tangent frame bundle FM, another set of operations on densities becomes available. This comes from the fact that in this case the (scalar and tensor) density bundles are associated to the frame bundle, and are thus natural bundles, which allows the construction of pullbacks along diffeomorphisms, as shown for tensor bundles in section 12.1, and more general in section 22.6. A possible way to introduce a pullback is by making use of the lift 22.6.6 to the associated bundle, and its relation 22.6.4, which we now also demand for densities. In the most simple case of a scalar density, this leads to the following definition.

Definition 23.5.1 (Pullback of a scalar tangent density). Let M, N be manifolds of dimension $\dim M = \dim N = n$ and $\varphi : M \to N$ a diffeomorphism. For a scalar density $\mathfrak{q} \in \Gamma(|D|^{\pm}_{w}(TN))$ of weight w, the *pullback* along φ is the scalar density

$$\varphi^*(\mathfrak{q}) = (|\varphi|^{\pm}_{w})^{-1} \circ \mathfrak{q} \circ \varphi \in \Gamma(|D|^{\pm}_{w}(TM))\,, \tag{23.5.1}$$

where $|\varphi|^{\pm}_{w} : |D|^{\pm}_{w}(TM) \to |D|^{\pm}_{w}(TN)$ denotes the associated bundle lift.

Having defined a notion of a pullback, one can proceed as in section 16.2 and introduce a Lie derivative, by taking the pullback along the flow of a vector field. This definition proceeds in full analogy to the Lie derivative 16.2.1 of tensor fields. We thus define:

Definition 23.5.2 (Lie derivative of a scalar tangent density)**.** Let $\mathfrak{q} \in \Gamma(|D|^{\pm}_{w}(TM))$ be a scalar density and $X \in \mathrm{Vect}(M)$ a vector field on a manifold M. Let $\phi : \mathbb{R} \times M \supseteq U \to M$ be the flow of X. The *Lie derivative* of $\mathfrak{q}$ with respect to X is the scalar density defined by

$$\mathcal{L}_X \mathfrak{q} = \lim_{t \to 0} \frac{\phi_t^* \mathfrak{q} - \mathfrak{q}}{t} . \tag{23.5.2}$$

It is instructive to derive coordinate expressions for the pullback of a scalar density and its Lie derivative. We start with the former, and introduce local coordinates (x^a) on M and (y^a) on N. For the corresponding induced bases of $|D|^{\pm}_{w}(TM)$ and $|D|^{\pm}_{w}(TN)$, we will use the notation $|\eth|^{\pm}_{w}$ and $|\eth'|^{\pm}_{w}$, respectively. Using these bases, we will write the original tensor density as $\mathfrak{q} = q|\eth'|^{\pm}_{w}$ and its pullback as $\mathfrak{q}' = \varphi^*(\mathfrak{q}) = q'|\eth|^{\pm}_{w}$. Our aim is to derive the relation between the component expressions q' and q. For this purpose, recall that for $x \in M$, we can write a frame $p \in F_x M$ in coordinates as the linear function

$$\begin{array}{rccc} p & : & \mathbb{R}^n & \to & T_x M \\ & & v^i \mathbb{e}_i & \mapsto & p(v^i \mathbb{e}_i) = p^a{}_i v^i \partial_a \end{array} , \tag{23.5.3}$$

where we used the notation ∂_a for the induced coordinate basis of TM. Applying the lift map $\varphi_\circ : FM \to FN$, which is defined via the differential $\varphi_* : TM \to TN$, we find

$$\varphi_\circ(p)(v^i \mathbb{e}_i) = \frac{\partial y^a}{\partial x^b} p^b{}_i v^i \partial'_a . \tag{23.5.4}$$

Now we recall that the coordinate bases on the density bundles are obtained as

$$|\eth|^{\pm}_{w} = [\delta, 1]_{|\rho|^{\pm}_{w}} , \quad |\eth'|^{\pm}_{w} = [\delta', 1]_{|\rho|^{\pm}_{w}} , \tag{23.5.5}$$

where by δ and δ' we denoted the local sections of the frame bundles FM and FN induced by the coordinate charts, as $\delta(\mathbb{e}_a) = \partial_a$ and $\delta'(\mathbb{e}_a) = \partial'_a$. By comparison with the formula for the frame bundle lift, we find that for $x \in M$ the coordinate frames at x and $\varphi(x) = y$ are related by

$$\varphi_\circ(\delta(x))(v^a \mathbb{e}_a) = \frac{\partial y^a}{\partial x^b} v^b \partial'_a = \delta'(y) \left(\frac{\partial y^a}{\partial x^b} v^b \mathbb{e}_a \right) = (\delta'(y) \circ g)(v^a \mathbb{e}_a) , \tag{23.5.6}$$

where $g \in \mathrm{GL}(n, \mathbb{R})$ is defined by

$$g(v^a \mathbb{e}_a) = g^a{}_b v^b \mathbb{e}_a = \frac{\partial y^a}{\partial x^b} v^b \mathbb{e}_a . \tag{23.5.7}$$

Hence, we have found the element $g \in \mathrm{GL}(n, \mathbb{R})$ which relates $\varphi_\circ(\delta(x))$ and $\delta'(y)$. With this knowledge, we can use the lift 22.6.6 to obtain

$$\begin{aligned} q(y)|\eth'|^{\pm}_{w} &= \mathfrak{q}(y) \\ &= |\varphi|^{\pm}_{w}(\mathfrak{q}'(x)) \\ &= |\varphi|^{\pm}_{w}(q'(x)|\eth'|^{\pm}_{w}) \\ &= |\varphi|^{\pm}_{w}([\delta(x), q'(x)]_{|\rho|^{\pm}_{w}}) \\ &= [\varphi_\circ(\delta(x)), q'(x)]_{|\rho|^{\pm}_{w}} \\ &= [\delta'(y) \cdot g, q'(x)]_{|\rho|^{\pm}_{w}} \\ &= [\delta'(y), |\rho|^{\pm}_{w}(g, q'(x))]_{|\rho|^{\pm}_{w}} \\ &= |\rho|^{\pm}_{w}(g, q'(x))|\eth'|^{\pm}_{w} . \end{aligned} \tag{23.5.8}$$

In other words, the component expressions for densities transform with the representation of the Jacobian $g^a{}_b = \partial y^a / \partial x^b$ of the map φ expressed in the respective coordinates on M and N. In particular, we thus find

$$\mathfrak{q} = q\eth'^{+}_{w} \in \Gamma(D^{+}_{w}(TN)) , \qquad q' = (\det g)^w q \tag{23.5.9a}$$

$$\mathfrak{t} = t\eth'^{-}_{w} \in \Gamma(D^{-}_{w}(TN))\,, \qquad t' = \operatorname{sgn}(\det g)(\det g)^{w} t \tag{23.5.9b}$$

$$\mathfrak{u} = u|\eth'|^{+}_{w} \in \Gamma(|D|^{+}_{w}(TN))\,, \qquad u' = |\det g|^{w} u \tag{23.5.9c}$$

$$\mathfrak{v} = v|\eth'|^{-}_{w} \in \Gamma(|D|^{-}_{w}(TN))\,, \qquad v' = \operatorname{sgn}(\det g)|\det g|^{w} v\,. \tag{23.5.9d}$$

Note that we obtain a positive sign for w, since we defined the density representations with a negative sign in section 23.1. We can then make use of these formulas in order to derive an expression for the Lie derivative, following the same prescription as in section 16.2, where we discussed the Lie derivative of tensor fields. Using coordinates (x^a) on M, let $X = X^a\partial_a$ be a vector field, $\phi : \mathbb{R} \times M \supseteq U \to M$ its flow and

$$\mathfrak{q} = q|\eth|^{\pm}_{w} \in \Gamma(|D|^{\pm}_{w}(TM))\,. \tag{23.5.10}$$

We then write the pullback $\mathfrak{q}' = \phi_t^*(\mathfrak{q})$ in the same coordinates as

$$\phi_t^*(\mathfrak{q}) = \mathfrak{q}'_t = q'_t|\eth|^{\pm}_{w} \in \Gamma(|D|^{\pm}_{w}(TM))\,. \tag{23.5.11}$$

Now the coordinate expression for the Lie derivative takes the form

$$\mathcal{L}_X\mathfrak{q}(x) = \lim_{t\to 0}\frac{q'_t(x) - q(x)}{t}|\eth|^{\pm}_{w} = \frac{\mathrm{d}}{\mathrm{d}t}q'_t(x)\bigg|_{t=0}|\eth|^{\pm}_{w}\,. \tag{23.5.12}$$

We then make use of the pullback formulas (23.5.9) we derived above, where now

$$\det g = \det\frac{\partial x'^{\bullet}_t}{\partial x^{\bullet}}\,, \tag{23.5.13}$$

and the vector field is related to the flow by

$$X^a(x) = \frac{\mathrm{d}}{\mathrm{d}t}x'^a_t(x)\bigg|_{t=0}\,. \tag{23.5.14}$$

Hence, we find the derivative

$$\frac{\mathrm{d}}{\mathrm{d}t}\det g\bigg|_{t=0} = \operatorname{tr}\frac{\mathrm{d}}{\mathrm{d}t}g\bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial x'^a_t}{\partial x^a}\bigg|_{t=0} = \partial_a\frac{\mathrm{d}}{\mathrm{d}t}x'^a_t(x)\bigg|_{t=0} = \partial_a X^a(x)\,. \tag{23.5.15}$$

Note that we evaluate this derivative at $t = 0$, where $\det g \to 1$, and hence we can set $\operatorname{sgn}(\det g) = 1$ and $|\det g| = \det g$ a priori, so that we only need the derivative

$$\frac{\mathrm{d}}{\mathrm{d}t}(\det g)^w\bigg|_{t=0} = w\partial_a X^a(x)\,. \tag{23.5.16}$$

Together with the term

$$\frac{\mathrm{d}}{\mathrm{d}t}q(x'_t(x))\bigg|_{t=0} = X^a(x)\partial_a q(x) \tag{23.5.17}$$

we then find the derivative

$$\frac{\mathrm{d}}{\mathrm{d}t}q'_t(x)\bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}(\det g)^w q(x'_t(x))\bigg|_{t=0} = X^a(x)\partial_a q(x) + w\partial_a X^a(x)q(x)\,. \tag{23.5.18}$$

Since this holds for all $x \in M$, we find that the Lie derivative is given as the density

$$\mathcal{L}_X\mathfrak{q} = (X^a\partial_a q + w\partial_a X^a q)|\eth|^{\pm}_{w}\,. \tag{23.5.19}$$

This is the coordinate expression for the Lie derivative of a scalar density. Note that in addition to the term for a scalar function which we derived in section 16.3 we find a contribution which scales with the weight w. Also note that the same formula holds for all cases (23.5.9). We finally remark that in the literature one conventionally omits the basis element, and thus finds only the term in brackets as the coordinate expression. Here we included the basis element for clarity, to distinguish the density component q from a scalar function.

It is now straightforward to apply the same construction also to tensor densities. One can define their pullback in full analogy to the case of scalar densities using associated bundles, or use the fact that tensor density bundles arise as a tensor product between scalar density bundles and ordinary tensor bundles. Here we do the latter, and remark that this definition is equivalent to the associated bundle construction, as we have shown in section 22.6.

Definition 23.5.3 (Pullback of a tensor tangent density)**.** Let M and N be manifolds and $\varphi : M \to N$ a diffeomorphism. The *pullback* of tensor densities on N to tensor densities on M is defined as the linear function $\varphi^* : \Gamma(|D|^{\pm}_{w}(TN) \otimes T^r_s N) \to \Gamma(|D|^{\pm}_{w}(TM) \otimes T^r_s M)$ that for any scalar density $\mathfrak{t} \in \Gamma(|D|^{\pm}_{w}(TN))$ and tensor field $T \in \Gamma(T^r_s N)$ holds

$$\varphi^*(\mathfrak{t} \otimes T) = \varphi^*(\mathfrak{t}) \otimes \varphi^*(T) \,. \tag{23.5.20}$$

Also a coordinate expression for the pullback is now easily derived from the formulas for the pullback of a tensor field (12.1.7) and a scalar density (23.5.9). Writing the original tensor density as

$$\mathfrak{T} = \mathfrak{T}^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} |\partial'|^{\pm}_{w} \otimes \partial'_{a_1} \otimes \ldots \otimes \partial'_{a_r} \otimes \mathrm{d}y^{b_1} \otimes \ldots \otimes \mathrm{d}y^{b_s} \tag{23.5.21}$$

and its pullback as

$$\varphi^*(\mathfrak{T}) = \mathfrak{T}' = \mathfrak{T}'^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} |\partial|^{\pm}_{w} \otimes \partial_{a_1} \otimes \ldots \otimes \partial_{a_r} \otimes \mathrm{d}x^{b_1} \otimes \ldots \otimes \mathrm{d}x^{b_s} \,, \tag{23.5.22}$$

one finds that the components are related by

$$\mathfrak{T}'^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} = \left\{ \begin{matrix} (\det g)^w \\ \operatorname{sgn}(\det g)(\det g)^w \\ |\det g|^w \\ \operatorname{sgn}(\det g)|\det g|^w \end{matrix} \right\} \frac{\partial x^{a_1}}{\partial y^{c_1}} \cdots \frac{\partial x^{a_r}}{\partial y^{c_r}} \frac{\partial y^{d_1}}{\partial x^{b_1}} \cdots \frac{\partial y^{d_s}}{\partial x^{b_s}} \mathfrak{T}^{c_1 \cdots c_r}{}_{d_1 \cdots d_s} \,, \tag{23.5.23}$$

where the term in braces must be chosen according to the parity of the density, and we wrote g for the Jacobian $g^a{}_b = \partial y^a / \partial x^b$ of the map φ expressed in the respective coordinates on M and N.

Given the pullback, it is now obvious how to construct the Lie derivative of tensor densities. As for scalars and ordinary densities, it is defined as follows.

Definition 23.5.4 (Lie derivative of a tensor tangent density)**.** Let $\mathfrak{T} \in \Gamma(|D|^{\pm}_{w}(TM) \otimes T^r_s M)$ be a tensor density and $X \in \mathrm{Vect}(M)$ a vector field on a manifold M. Let $\phi : \mathbb{R} \times M \supseteq U \to M$ be the flow of X. The *Lie derivative* of $\mathfrak{T}$ with respect to X is the tensor density defined by

$$\mathcal{L}_X \mathfrak{T} = \lim_{t \to 0} \frac{\phi_t^* \mathfrak{T} - \mathfrak{T}}{t} \,. \tag{23.5.24}$$

Using the previous findings for the coordinate expressions of the Lie derivative of tensor fields and scalar densities, one can easily derive a coordinate expression for the Lie derivative of a tensor density. Observe that the factors attained in the pullback (23.5.23) of a tensor density are simply a product of the factors we found for a tensor field (12.1.7) and a scalar density (23.5.9). Hence, in the coordinate expression for the Lie derivative, the respective terms obtained from these factors in the formulas (16.2.12) and (23.5.19) must be added, as a consequence of the Leibniz rule. We thus obtain the coordinate expression

$$\begin{aligned}(\mathcal{L}_X \mathfrak{T})^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} &= X^c \partial_c \mathfrak{T}^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} + w \partial_c X^c \mathfrak{T}^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} \\ &\quad - \partial_c X^{a_1} \mathfrak{T}^{c a_2 \cdots a_r}{}_{b_1 \cdots b_s} - \ldots - \partial_c X^{a_r} \mathfrak{T}^{a_1 \cdots a_{r-1} c}{}_{b_1 \cdots b_s} \\ &\quad + \partial_{b_1} X^c \mathfrak{T}^{a_1 \cdots a_r}{}_{c b_2 \cdots b_s} + \ldots + \partial_{b_s} X^c \mathfrak{T}^{a_1 \cdots a_r}{}_{b_1 \cdots b_{s-1} c} \,,\end{aligned} \tag{23.5.25}$$

where we have now omitted the basis element for brevity, as its form should be clear by now.

We finally introduce another property of manifolds, which can conveniently be obtained from the notion of tangent densities. Recall from definition 24.2.2 that we defined a vector bundle to be orientable if its orientation bundle possesses a global section. Since every fiber bundle is canonically equipped with a vector bundle, namely its tangent bundle, we can make the following definition.

Definition 23.5.5 (Orientable manifold)**.** A manifold is called *orientable* if and only if its tangent bundle is orientable.

The orientability of a manifold is a fundamental property, which is deeply connected to its geometry. It turns out that it can be related to the properties of atlases as follows.

Theorem 23.5.1. *A manifold is orientable if and only if it possesses an atlas such that the determinants of the Jacobian matrices of all transition functions are positive.*

Proof. ►...◄ ■

23.6 Twisted differential forms

In chapter 9 we have introduced and intensively discussed differential forms, which are one of the most important and foundational classes of objects in differential geometry. We have defined and studied a number of operations on them, in particular the exterior product, interior product and exterior derivative. In chapter 17 we have seen that the latter two are just particular examples of graded derivations on differential forms. We will now see that differential forms have a close relative, which turns out to be useful in physics, but is much less covered in the literature on differential geometry, despite having numerous applications in mathematical as well as in physics. We start with the following definition.

Definition 23.6.1 (Twisted differential form)**.** A *twisted differential form of rank* k (or *twisted* k*-form*) on a manifold M is a section of the vector bundle $\Lambda^k T^*M \otimes D_0^-(T^*M)$ for $k \in \mathbb{N}$. The space of all twisted k-forms on M is denoted $\bar{\Omega}^k(M)$, while the space of all twisted differential forms is denoted

$$\bar{\Omega}^\bullet(M) = \bigoplus_{k=0}^{\dim M} \bar{\Omega}^k(M)\,. \tag{23.6.1}$$

Recalling the canonical bundle isomorphisms

$$D_0^-(TM) \cong D_0^-(T^*M) \cong [D_0^-(TM)]^* \cong [D_0^-(T^*M)]^*\,, \tag{23.6.2}$$

and comparing with definition 23.2.1, we see that we can also understand twisted k-forms as totally antisymmetric pseudotensors of rank $(0,k)$. For this reason, also the term k*-pseudoform* can be sometimes be found in the literature, while the term twisted k-form is used for a more general notion. Also note that we have already encountered two special cases. A twisted zero-form is simply a section of the bundle $D_0^-(T^*M)$ and hence a pseudoscalar, while for $k = m = \dim M$ we recall the canonical isomorphism $\Lambda^m T^M \cong D_{-1}^+(T^*M)$, so that

$$\Lambda^m T^M \otimes D_0^-(T^*M) \cong D_{-1}^-(T^*M) \cong D_1^-(TM) \tag{23.6.3}$$

is the bundle of scalar pseudo-densities.

Given coordinates (x^a) on M, it follows that any twisted k-form $\mathfrak{A} \in \bar{\Omega}^k(M)$ can be written as

$$\mathfrak{A} = \frac{1}{k!}\mathfrak{A}_{a_1\cdots a_k}\mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k} \otimes \mathfrak{d}_0^- \,, \tag{23.6.4}$$

similarly to (untwisted) differential forms. The important difference, which is often also used as the defining property in the physics literature, is the fact that the components $\mathfrak{A}_{a_1\cdots a_k}$ of a twisted k-form incur an additional factor $\operatorname{sgn}(\det \partial x'^\bullet/\partial x^\bullet)$ under a coordinate change from (x^a) to (x'^a), which comes from the change of the basis element from $\mathfrak{d}_0^-$ to $\mathfrak{d}_0'^-$, as given by the relations (23.5.9) and (23.5.23).

The question arises whether one can apply any of the operations we have introduced for differential forms also to twisted forms. This can be seen most easily for the exterior product. The crucial observation which makes this possible is the fact that the bundle $D_0^-(T^*M)$ is its own dual, and that

$$D_0^-(T^*M) \otimes D_0^-(T^*M) \cong D_0^+(T^*M) \cong M \times \mathbb{R} \tag{23.6.5}$$

is the trivial line bundle. Hence, the tensor product of two pseudoscalars is canonically identified with a real function. This allows us to give the following definition.

Definition 23.6.2 (Exterior product of twisted forms)**.** For any differential forms $\alpha \in \Omega^k(M)$, $\beta \in \Omega^l(M)$ and pseudoscalars $\mathfrak{a}, \mathfrak{b} \in \bar{\Omega}^0(M)$, the *exterior product* is defined as

$$(\alpha \otimes \mathfrak{a}) \wedge \beta = (\alpha \wedge \beta) \otimes \mathfrak{a} \in \bar{\Omega}^{k+l}(M)\,, \tag{23.6.6a}$$

$$\alpha \wedge (\beta \otimes \mathfrak{b}) = (\alpha \wedge \beta) \otimes \mathfrak{b} \in \bar{\Omega}^{k+l}(M)\,, \tag{23.6.6b}$$

$$(\alpha \otimes \mathfrak{a}) \wedge (\beta \otimes \mathfrak{b}) = (\alpha \wedge \beta) \otimes (\mathfrak{a} \otimes \mathfrak{b}) \in \Omega^{k+l}(M)\,, \tag{23.6.6c}$$

and extended to arbitrary twisted forms $\mathfrak{A} \in \bar{\Omega}^k(M)$, $\mathfrak{B} \in \bar{\Omega}^l(M)$ by linearity in each factor.

In other words, with regard to the exterior product, we may treat pseudoscalars exactly as we would treat scalar functions, by simply pulling them out of the product, and let the exterior product act only on the untwisted differential form factors. It is thus not surprising that the exterior product of twisted forms inherits the same properties as for untwisted forms.

Theorem 23.6.1. *For any twisted or untwisted differential forms* $\mathfrak{A} \in \overset{(-)}{\Omega}{}^k(M)$*,* $\mathfrak{B} \in \overset{(-)}{\Omega}{}^l(M)$ *and* $\mathfrak{C} \in \overset{(-)}{\Omega}{}^r(M)$ *holds:*

1. *Graded anticommutativity:*
$$\mathfrak{A} \wedge \mathfrak{B} = (-1)^{kl}\mathfrak{B} \wedge \mathfrak{A}\,. \tag{23.6.7}$$
2. *Associativity:*
$$\mathfrak{A} \wedge (\mathfrak{B} \wedge \mathfrak{C}) = (\mathfrak{A} \wedge \mathfrak{B}) \wedge \mathfrak{C} = \mathfrak{A} \wedge \mathfrak{B} \wedge \mathfrak{C}\,. \tag{23.6.8}$$
3. *$\mathbb{R}$-linearity in each factor.*

Proof. ▶...◀ ∎

We then continue with the interior product. From theorem ▶...◀ we know that also in this case we can pull out a scalar factor which multiplies either the differential form or the vector field. Extending this definition to allow also for pseudoscalar factors thus suggests itself. In fact, we may even consider the case that the vector field carries a pseudoscalar factor, which leads to the following definition, which is again useful in physics.

Definition 23.6.3 (Pseudovector field)**.** Let M be a manifold. A *pseudovector field* on M is a section of the bundle $TM \otimes D_0^-(TM)$. The space of all pseudovector fields on M is denoted $\bar{\mathrm{Vect}}(M)$.

With this definition at hand, it is now straightforward to generalize the interior product as follows.

Definition 23.6.4 (Interior product of twisted forms and pseudovector fields)**.** For any vector fields $X \in \mathrm{Vect}(M)$, differential forms $\alpha \in \Omega^k(M)$ and pseudoscalars $\mathfrak{x}, \mathfrak{a}$, the *interior product* is defined such that

$$\iota_{X \otimes \mathfrak{x}} \alpha = \iota_X \alpha \otimes \mathfrak{x} \in \bar{\Omega}^{k-1}(M)\,, \tag{23.6.9a}$$

$$\iota_X(\alpha \otimes \mathfrak{a}) = \iota_X \alpha \otimes \mathfrak{a} \in \bar{\Omega}^{k-1}(M)\,, \tag{23.6.9b}$$

$$\iota_{X \otimes \mathfrak{x}}(\alpha \otimes \mathfrak{a}) = \iota_X \alpha \otimes (\mathfrak{x} \otimes \mathfrak{a}) \in \Omega^{k-1}(M)\,, \tag{23.6.9c}$$

and extended to arbitrary twisted forms $\mathfrak{A} \in \bar{\Omega}^k(M)$ by linearity.

Just like the exterior product, also this generalized notion of an interior product inherits a number of properties from the untwisted case. ►...◄

Finally, we come to the exterior derivative. ►...◄

Theorem 23.6.2. *Let $\mathfrak{q} \in \bar{\Omega}^0(M)$ a pseudoscalar. Then there exists a unique twisted one-form $\mathrm{d}\mathfrak{q} \in \bar{\Omega}^1(M)$ such that for all vector fields $X \in \mathrm{Vect}(M)$ holds*

$$\mathcal{L}_X \mathfrak{q} = \iota_X \mathrm{d}\mathfrak{q}\,. \tag{23.6.10}$$

Proof. ►...◄ ■

We can thus define:

Definition 23.6.5 (Total differential of pseudoscalars)**.** For a pseudoscalar $\mathfrak{q} \in \bar{\Omega}^0(M)$, the *total differential* is the twisted one-form $\mathrm{d}\mathfrak{q} \in \Omega^1(M)$ constructed in theorem 23.6.2.

Chapter 24

G-structures

24.1 Volume forms

It follows from the properties of the exterior power bundle that for a vector bundle $\pi : E \to M$ of rank k the highest non-trivial exterior power is given by $\Lambda^k E$, and that this is a vector bundle of rank 1. We have seen in theorem 23.1.2 that we can canonically identify this bundle with the density bundle $D^+_{-1}(E) = |D|^-_{-1}(E)$, which we will do in this section. Sections of this bundle play an important role, as they are related to defining integrals, and therefore deserve special attention. Of particular interest are nowhere vanishing sections of this type, which we will now study, and which we define as follows.

Definition 24.1.1 (Volume form)**.** A *volume form* on a vector bundle $\pi : E \to M$ of rank k is a nowhere vanishing section of the bundle $D^+_{-1}(E)$, i.e., a section $\omega \in \Gamma(D^+_{-1}(E))$ such that $\omega(x) \neq 0$ for all $x \in M$.

Using a local basis $(e_a, a = 1, \ldots, k)$ on an open set $U \subset M$, which is simply a local section $e : U \to F(E)$ of the frame bundle, a volume form can always be written in the form $\omega = we^+_{-1} \cong we_1 \wedge \ldots \wedge e_k$ with $w(x) \neq 0$ for all $x \in U$, with respect to the induced basis e^+_{-1}. Note that although at first sight it looks like w is simply a real function on M, so that one could identify volume forms and nowhere vanishing functions, this is not the case - the value of w in this definition depends on the choice of the basis e, while the value of a real function $f \in \Gamma(D^+_0(E)) \cong C^\infty(M, \mathbb{R})$ depends only on a point on M, but not on the choice of coordinates used for its description. However, functions can be used to compare volume forms. If ω is a volume form and $f \in C^\infty(M, \mathbb{R})$ is nowhere vanishing, then obviously also $f\omega$ is a volume form. In fact, every volume form can be expressed by any other volume form and a function:

Theorem 24.1.1. *Let ω and ω' be volume forms on a vector bundle $\pi : E \to M$. Then there exists a unique nowhere vanishing function $f \in C^\infty(M, \mathbb{R})$ such that $\omega' = f\omega$.*

There is another crucial difference between the bundles $D^+_0(E) \cong \Lambda^0 E \cong M \times \mathbb{R}$ and $D^+_{-1}(E) \cong \Lambda^k E$: although both are vector bundles of rank 1 over M, they are in general not isomorphic. While the former is simply the trivial line bundle $M \times \mathbb{R}$ and therefore always has nowhere vanishing sections such as the constant function $f : \bullet \mapsto 1$, this does not necessarily hold for the latter, and there are examples which do not possess nowhere vanishing sections. Hence, not every vector bundle allows for a volume form. In fact, volume forms can be identified with another structure, as we will see next.

Theorem 24.1.2. *There is a one-to-one correspondence between volume forms of a vector bundle $\pi : E \to M$ of rank k and $\mathrm{SL}(k, \mathbb{R})$-reductions of its general linear frame bundle $\varpi : F(E) \to M$.*

Proof. Recall from definition 20.6.1 that a reduction of a principal bundle is a particular homomorphism of principal bundles covering the identity on M. We will sketch how to construct such a homomorphism from a volume form and vice versa, and show how these constructions are related to each other.

First, consider a volume form $\omega \in \Gamma(D^+_{-1}(E))$ and $x \in M$. We call a frame $p \in F_x(E)$ over x a *unit frame* if and only if

$$[p, 1]_{\rho^+_{-1}} = \omega(\varpi(p)) = \omega(x) . \tag{24.1.1}$$

Note that such unit frames exist for all $x \in M$. To see this, recall that by definition of the bundle $D^+_{-1}(E)$, for *any* frame $p \in F_x(E)$ there exists a unique $c \in \mathbb{R}$ such that

$$[p, c]_{\rho^+_{-1}} = \omega(x) . \tag{24.1.2}$$

From the condition that ω nowhere vanishes follows that $c \neq 0$. We may thus consider the frame p' obtained from p by replacing its first component p_1 by $p'_1 = cp_1$, while all other components remain unchanged, and again obtain a frame. One easily checks that this new frame is a unit frame, since

$$\omega(x) \cong cp_1 \wedge \ldots \wedge p_k = p'_1 \wedge \ldots \wedge p'_k . \tag{24.1.3}$$

Further, one finds that if $p \in F_x(E)$ is a unit frame, then $p \cdot g$ is a unit frame if and only if $\det g = 1$, and hence $g \in \mathrm{SL}(k, \mathbb{R})$. For every $x \in M$, we thus find that the set of unit frames over x carries a right action of $\mathrm{SL}(k, \mathbb{R})$, which is obtained by the restriction of the action of $\mathrm{GL}(k, \mathbb{R})$ on $F_x(E)$, and that this action is free and transitive. From the smoothness of ω one finds that these sets constitute the fibers of a principal $\mathrm{SL}(k, \mathbb{R})$-bundle over M, which we will denote $\mathrm{SL}(E, \omega)$. Finally, one finds that the canonical inclusion $\mathrm{SL}(E, \omega) \hookrightarrow F(E)$, together with the inclusion $\mathrm{SL}(k, \mathbb{R}) \hookrightarrow \mathrm{GL}(k, \mathbb{R})$ define a principal bundle reduction of $F(E)$.

To show the converse direction, let $\chi : Q \to M$ be a principal $\mathrm{SL}(k, \mathbb{R})$ bundle and $\phi : Q \to F(E)$ a $\mathrm{SL}(k, \mathbb{R})$ reduction over the canonical inclusion $\mathrm{SL}(k, \mathbb{R}) \hookrightarrow \mathrm{GL}(k, \mathbb{R})$. For $q \in Q$ with $\chi(q) = x$, define

$$\omega(x) = [\phi(q), 1]_{\rho^+_{-1}} . \tag{24.1.4}$$

This is independent of the choice of the representative, since for any other $q' = q \cdot g$ with $g \in \mathrm{SL}(k, \mathbb{R})$ we have $\det g = 1$ and hence

$$[\phi(q'), 1]_{\rho^+_{-1}} = [\phi(q \cdot g), 1]_{\rho^+_{-1}} = [\phi(q) \cdot g, 1]_{\rho^+_{-1}} = [\phi(q), 1]_{\rho^+_{-1}} = \omega(x) . \tag{24.1.5}$$

One finds that this defines a nowhere vanishing section ω of $D^+_{-1}(E)$, and hence a volume form on $\pi : E \to M$.

Finally, one can see from the two constructions that the bundle $\mathrm{SL}(E, \omega)$ constructed above is simply the image of Q under ϕ in $F(E)$, which can canonically be identified with Q, since ϕ restricts to an isomorphism of principal $\mathrm{SL}(k, \mathbb{R})$ bundles. ∎

In the proof we have constructed a particular subbundle of the frame bundle $F(E)$, which we will study further below, and denote as follows.

Definition 24.1.2 (Unit frame bundle). Let $\pi : E \to M$ be a vector bundle of rank k equipped with a volume form ω and $\varpi : F(E) \to M$ its general linear frame bundle. Its *unit frame bundle* is the principal $\mathrm{SL}(k, \mathbb{R})$ bundle

$$\mathrm{SL}(E, \omega) = \left\{ p \in F(E), [p, 1]_{\rho^+_{-1}} = \omega(\varpi(p)) \right\} \tag{24.1.6}$$

of unit frames, together with the right action obtained by restricting the right action of $\mathrm{GL}(k,\mathbb{R})$ on $F(E)$ to $\mathrm{SL}(k,\mathbb{R})$.

Using the unit frame bundle, we can obtain yet another interpretation of a volume form. Recall from theorem 20.6.1 that we can identify a principal bundle reduction to a closed subgroup with a section of a suitable associated bundle, whose fiber is a coset space. Now we aim to identify this section with the volume form. It turns out that we can achieve this as follows.

Theorem 24.1.3. *Let $\pi : E \to M$ be a vector bundle and $\varpi : F(E) \to M$ its general linear frame bundle. Then there exists a canonical isomorphism between the slit density bundle*

$$D^{*+}_{-1}(E) = F(E) \times_{\rho^+_{-1}} \mathbb{R}^* \tag{24.1.7}$$

and the coset bundle

$$F(E) \times_\rho G/H\,, \tag{24.1.8}$$

where $G = \mathrm{GL}(k,\mathbb{R})$, $H = \mathrm{SL}(k,\mathbb{R})$ and ρ denotes the canonical left action of G on G/H.

Proof. It is obviously enough to show that $\mathbb{R}^* \cong G/H$ and that this identification relates ρ and ρ^+_{-1}, so that one obtains the same associated bundle. First note that ρ^+_{-1} acts on $\mathbb{R}^*$ by restriction, since $\rho^+_{-1}(g,0) = 0$ for all $g \in \mathrm{GL}(k,\mathbb{R})$, so that $D^{*+}_{-1}(E)$ as defined above is well-defined. Further, we can write every $g \in \mathrm{GL}(k,\mathbb{R})$ uniquely as

$$g = \mathrm{diag}(\det g, 1, \ldots, 1) \cdot h \tag{24.1.9}$$

with $h \in \mathrm{SL}(k,\mathbb{R})$. Hence, every coset gH is uniquely characterized by $\det g$. Since $\det g \in \mathbb{R}^*$, this yields an identification $\mathbb{R}^* \cong G/H$. Further, from the relation $\det(gg') = \det g \cdot \det g'$ follows that G acts on G/H by multiplication with $\det g$, which is exactly the action defined by ρ^+_{-1}. ■

With this identification, we can see why there exist vector bundles which do not admit a volume form. Note that $\mathbb{R}^*$ is a one-dimensional manifold with two connected components, namely the positive and negative real numbers. Hence, also the fibers of $D^{*+}_{-1}(E)$ have two connected components. If we consider, for example, a vector bundle over the circle $M = S^1$, which is constructed such that one passes from one connected component to the other when going around the circle, it is not possible to find a smooth section of this bundle. An example is the infinite Möbius strip $\pi : E \to M$ we constructed in example 3.1.1. Here we have $D^+_{-1}(E) \cong \Lambda^1 E \cong E$, and if we remove the zero section, we obtain a bundle which has no global sections, since every section of E must pass through 0.

24.2 Orientations

In the previous section we have seen that not every vector bundle admits a volume form, due to the fact that the fibers of the bundle $D^{*+}_{-1}(E)$ consist of two connected components, and that these fibers might be "glued the wrong way", so that the resulting fiber bundle does not have any global sections. Intuitively, one could try to simplify this geometric picture by shrinking every connected component of the fibers to a single point, and considering a bundle whose fiber is the zero-dimensional manifold $\mathbb{Z}_2 = \{1, -1\}$ consisting of only two points, instead of $\mathbb{R}^*$. It turns out that this type of construction is indeed possible, and yields a structure which is just as rich as the volume forms we discussed in the previous section. To arrive at this structure, we proceed by analogy. Recall that $\mathrm{GL}(k,\mathbb{R})$ acts on $\mathbb{R}^*$ by multiplication with the determinant, and that we could identify this action as the canonical action on the coset space $\mathrm{GL}(k,\mathbb{R})/\mathrm{SL}(k,\mathbb{R})$. Looking for a similar action on $\mathbb{Z}_2$, we find that $\mathrm{GL}(k,\mathbb{R})$ naturally acts by multiplication with *the sign of* the determinant. Proceeding as in the previous section, one finds that this allows an identification of $\mathbb{Z}_2$ with the coset space $\mathrm{GL}(k,\mathbb{R})/\mathrm{GL}^+(k,\mathbb{R})$, where

$$\mathrm{GL}^+(k,\mathbb{R}) = \{g \in \mathrm{GL}(k,\mathbb{R}), \det g > 0\} \tag{24.2.1}$$

is the positive general linear group. The associated bundle construction then yields a bundle, whose sections we denote as follows.

Definition 24.2.1 (Orientation)**.** Let $\pi : E \to M$ be a vector bundle and $F(E)$ its general linear frame bundle. An *orientation* on E is a section of the bundle

$$F(E) \times_\rho G/H \,, \tag{24.2.2}$$

where $G = \mathrm{GL}(k,\mathbb{R})$, $H = \mathrm{GL}^+(k,\mathbb{R})$ and ρ denotes the canonical left action of G on G/H.

As it is also the case with volume forms, it is intuitively clear that not every vector bundle admits an orientation. We thus introduce the following notion.

Definition 24.2.2 (Orientable vector bundle)**.** A vector bundle is called *orientable* if and only if it admits an orientation.

In analogy to the case of volume forms, we can find equivalent descriptions of orientations. We first take a closer look at the associated bundle we constructed in definition 24.2.1. As mentioned earlier, the fibers of this bundle are given by the space $\mathbb{Z}_2$, on which $\mathrm{GL}(k,\mathbb{R})$ acts by multiplication with the sign of the determinant. By comparison with the density bundles listed in definition 23.1.1, we find the same action under the name $\rho_0^- = |\rho|_0^-$, albeit acting on the larger space $\mathbb{R}$. As it turns out, the bundle $D_0^-(E) = |D|_0^-(E)$ is a curious special case. This bundle is its own dual, and the tensor product with itself yields the trivial line bundle $D_0^+(E) = |D|_0^+(E) \cong M \times \mathbb{R}$, and it is sometimes denoted as follows.

Definition 24.2.3 (Orientation line bundle)**.** Let $\pi : E \to M$ be a vector bundle. Its *orientation line bundle* is the bundle $D_0^-(E) = |D|_0^-(E)$.

Also the sections of this bundle carry a particular name in the literature.

Definition 24.2.4 (Pseudoscalar)**.** Let $\pi : E \to M$ be a vector bundle. A *pseudoscalar* on E is a section of the orientation line bundle $D_0^-(E) = |D|_0^-(E)$.

It is easy to check that the action ρ_0^- preserves the absolute value $|c|$ of $c \in \mathbb{R}$, as it multiplies only with ± 1, and so restricts to an action on $\mathbb{Z}_2 = \{1, -1\}$. Hence, we may define another associated bundle s follows.

Definition 24.2.5 (Orientation bundle)**.** Let $\pi : E \to M$ be a vector bundle. Its *orientation bundle* is the discrete density bundle

$$F(E) \times_{\rho_0^-} \mathbb{Z}_2 \,. \tag{24.2.3}$$

The name already suggests that we may regard an orientation also as a section of this bundle. This follows from the following statement.

Theorem 24.2.1. *Let $\pi : E \to M$ be a vector bundle and $\varpi : F(E) \to M$ its general linear frame bundle. Then there exists a canonical isomorphism between the orientation bundle $F(E) \times_{\rho_0^-} \mathbb{Z}_2$ and the coset bundle $F(E) \times_\rho G/H$, where $G = \mathrm{GL}(k, \mathbb{R})$, $H = \mathrm{GL}^+(k, \mathbb{R})$ and ρ denotes the canonical left action of G on G/H.*

Proof. One can proceed in analogy to the proof of theorem 24.1.3, and show that the two fiber spaces and the left actions they carry are identical. For this purpose, note that one can write every $g \in \mathrm{GL}(k, \mathbb{R})$ uniquely as

$$g = \mathrm{diag}(\mathrm{sgn}(\det g), 1, \dots, 1) \cdot h \tag{24.2.4}$$

with $h \in \mathrm{GL}^+(k, \mathbb{R})$. Hence, every coset gH is uniquely characterized by $\mathrm{sgn}(\det g) \in \mathbb{Z}_2$. Using the relation $\det(gg') = \det g \cdot \det g'$, it follows that G acts on G/H by multiplication with $\mathrm{sgn}(\det g)$, which corresponds to the action on $\mathbb{Z}_2$ defined by ρ_0^-. ■

Since $F(E) \times_{\rho_0^-} \mathbb{Z}_2 \subset D_0^-(E)$, we can also consider an orientation as a section of the bundle $D_0^-(E)$, which in a given basis e_0^-, induced by a basis e on E, takes only the values ± 1. This is similar to the definition of a volume form, which in an induced local basis takes only non-zero values. Note that this property is independent of the choice of the basis, since the induced basis of $D_0^-(E)$ changes only by a sign, as can be seen from the corresponding transformation law (23.1.24). Since we can identify a basis of E with a frame, we arrive at the following notion.

Definition 24.2.6 (Oriented frame bundle)**.** Let $\pi : E \to M$ be a vector bundle of rank k equipped with an orientation $\sigma \in \Gamma(F(E) \times_{\rho_0^-} \mathbb{Z}_2)$ and $\varpi : F(E) \to M$ its general linear frame bundle. Its *oriented frame bundle* is the principal $\mathrm{GL}^+(k, \mathbb{R})$ bundle

$$\mathrm{GL}^+(E, \sigma) = \left\{ p \in F(E), [p, 1]_{\rho_0^-} = \sigma(\varpi(p)) \right\} \tag{24.2.5}$$

of oriented frames, together with the right action obtained by restricting the right action of $\mathrm{GL}(k, \mathbb{R})$ on $F(E)$ to $\mathrm{GL}^+(k, \mathbb{R})$.

Obviously, choosing an orientation is now equivalent to choosing an oriented frame bundle, i.e., declaring one of the two connected components of each fiber of $F(E)$ as containing the oriented frames, and thus turning this connected component into the fiber of $\mathrm{GL}^+(E, \sigma)$. We formalize this equivalence in the following statement.

Theorem 24.2.2. *There is a one-to-one correspondence between orientations of a vector bundle $\pi : E \to M$ of rank k and $\mathrm{GL}^+(k, \mathbb{R})$-reductions of its general linear frame bundle $\varpi : F(E) \to M$.*

Proof. We can proceed in analogy to the proof of theorem 24.1.2, by first considering an orientation $\sigma \in \Gamma(F(E) \times_{\rho_0^-} \mathbb{Z}_2)$ and $x \in M$. Then we call a frame $p \in F_x(E)$ over x an *oriented* *frame* if and only if

$$[p,1]_{\rho_0^-} = \sigma(\varpi(p)) = \sigma(x)\,. \tag{24.2.6}$$

We need to show that oriented frames exist for all $x \in M$. It follows from the definition of the associated bundle $F(E) \times_{\rho_0^-} \mathbb{Z}_2$ that for *any* frame $p \in F_x(E)$ there exists a unique $c \in \mathbb{Z}_2$ such that

$$[p,c]_{\rho_0^-} = \sigma(x)\,. \tag{24.2.7}$$

We may then consider the frame p' obtained from p by replacing its first component p_1 by $p'_1 = cp_1$, while all other components remain unchanged, and again obtain a frame. It follows that this new frame p' is oriented, since

$$[p,c]_{\rho_0^-} = [p' \cdot g, c]_{\rho_0^-} = [p', \operatorname{sgn}(\det g)c]_{\rho_0^-} = [p', c^2]_{\rho_0^-} = [p',1]_{\rho_0^-} = \sigma(x)\,, \tag{24.2.8}$$

where $g = \operatorname{diag}(c,1,\ldots,1)$, and we repeatedly used the fact that $\det g = c = \pm 1$, so that $c = \operatorname{sgn}(c)$ and $c^2 = 1$. Further, it holds that if $p \in F_x(E)$ is oriented, then $p \cdot g$ is oriented if and only if $\det g > 0$, and hence $g \in \mathrm{GL}^+(k,\mathbb{R})$. For every $x \in M$, we thus find that the set of oriented frames over x carries a right action of $\mathrm{GL}^+(k,\mathbb{R})$, which is obtained by the restriction of the action of $\mathrm{GL}(k,\mathbb{R})$ on $F_x(E)$, and that this action is free and transitive. From the smoothness of σ one finds that these sets constitute the fibers of a principal $\mathrm{GL}^+(k,\mathbb{R})$-bundle over M, which we will denote $\mathrm{GL}^+(E,\sigma)$. Finally, one finds that the canonical inclusion $\mathrm{GL}^+(E,\sigma) \hookrightarrow F(E)$, together with the inclusion $\mathrm{GL}^+(k,\mathbb{R}) \hookrightarrow \mathrm{GL}(k,\mathbb{R})$ define a principal bundle reduction of $F(E)$.

To show the converse direction, let $\chi : Q \to M$ be a principal $\mathrm{GL}^+(k,\mathbb{R})$ bundle and $\phi : Q \to F(E)$ a $\mathrm{GL}^+(k,\mathbb{R})$ reduction over the canonical inclusion $\mathrm{GL}^+(k,\mathbb{R}) \hookrightarrow \mathrm{GL}(k,\mathbb{R})$. For $q \in Q$ with $\chi(q) = x$, define

$$\sigma(x) = [\phi(q),1]_{\rho_0^-}\,. \tag{24.2.9}$$

This is independent of the choice of the representative, since for any other $q' = q \cdot g$ with $g \in \mathrm{GL}^+(k,\mathbb{R})$ we have $\det g > 0$ and hence

$$[\phi(q'),1]_{\rho_0^-} = [\phi(q \cdot g),1]_{\rho_0^-} = [\phi(q) \cdot g,1]_{\rho_0^-} = [\phi(q), \operatorname{sgn}(\det g)]_{\rho_0^-} = [\phi(q),1]_{\rho_0^-} = \sigma(x)\,. \tag{24.2.10}$$

One finds that this defines a section σ of $F(E) \times_{\rho_0^-} \mathbb{Z}_2$, and hence an orientation on $\pi : E \to M$.

Finally, one can see from the two constructions that the bundle $\mathrm{GL}^+(E,\omega)$ constructed above is simply the image of Q under ϕ in $F(E)$, which can canonically be identified with Q, since ϕ restricts to an isomorphism of principal $\mathrm{GL}^+(k,\mathbb{R})$ bundles. ■

Intuitively, one may now expect that if one can choose an orientation, and hence an oriented frame bundle, one can also choose a volume form, by a further bundle reduction. Indeed, it turns out that this is the case.

Theorem 24.2.3. *A vector bundle admits a volume form if and only if it is orientable.*

Proof. ►. . . ◄ ■

24.3 Twisted volume forms

In the previous section we have constructed a bundle by taking the bundle $D^{*+}_{-1}(E)$ of volume forms over a vector bundle, and "forgetting" the magnitude of its elements and keeping only the orientation, i.e., by identifying all elements which belong to the same connected component of each fiber. We thus reduced the fiber space from $\mathbb{R}^*$ to $\mathbb{Z}_2$. We now pose the question whether one can also "forget" the complementary piece of information, namely the orientation, and keep only the magnitude. In other words, we will now identify antipodal pairs of elements of $D^{*+}_{-1}(E)$, in order to obtain a bundle whose fiber is diffeomorphic to the connected manifold $\mathbb{R}^+$. Looking for a suitable action of $\mathrm{GL}(k,\mathbb{R})$ on this manifold, one finds that in this case one needs to multiply with *the absolute value of* the determinant. In this case it turns out that we can identify $\mathbb{R}^+$ with the coset space $\mathrm{GL}(k,\mathbb{R})/\mathrm{SL}^{\pm}(k,\mathbb{R})$, where

$$\mathrm{SL}^{\pm}(k,\mathbb{R}) = \{g \in \mathrm{GL}(k,\mathbb{R}), |\det g| = 1\} \tag{24.3.1}$$

is the indefinite special linear group. Then we may again employ the associated bundle construction to obtain a bundle, whose sections we denote as follows.

Definition 24.3.1 (Twisted volume form)**.** Let $\pi : E \to M$ be a vector bundle and $F(E)$ its general linear frame bundle. A *twisted volume form* on E is a section of the bundle

$$F(E) \times_\rho G/H\,, \tag{24.3.2}$$

where $G = \mathrm{GL}(k,\mathbb{R})$, $H = \mathrm{SL}^{\pm}(k,\mathbb{R})$ and ρ denotes the canonical left action of G on G/H.

In the literature, a twisted volume form is also often called simply a *density*. We will not use this terminology here, since we use the term "density" for general sections of density bundles. One may wonder whether passing to this new bundle, whose fibers have only one connected component, always allows us to find sections, or whether there happen to be other obstructions in certain cases. It turns out that the former is true:

Theorem 24.3.1. *Every vector bundle admits twisted volume forms.*

Proof. ►. . . ◄ ■

We can now apply the same kind of constructions as in the previous sections for the (untwisted) volume forms and orientations. First, we relate the associated bundle built from coset spaces to a suitable density bundle, whose fiber is given by the positive real numbers $\mathbb{R}^+$ and which can be obtained from the general linear frame bundle by acting on these fibers by multiplication with the absolute value of the determinant. Taking a look at the representations used in the definition 23.1.1 of the density bundles, we see that this action is denoted $\rho^-_{-1} = |\rho|^+_{-1}$, and that we may restrict it from $\mathbb{R}$ to $\mathbb{R}^+$. Hence, we find the following bundle isomorphism.

Theorem 24.3.2. *Let $\pi : E \to M$ be a vector bundle and $\varpi : F(E) \to M$ its general linear frame bundle. Then there exists a canonical isomorphism between the positive density bundle*

$$F(E) \times_{\rho^-_{-1}} \mathbb{R}^+ \tag{24.3.3}$$

and the coset bundle

$$F(E) \times_\rho G/H\,, \tag{24.3.4}$$

where $G = \mathrm{GL}(k,\mathbb{R})$, $H = \mathrm{SL}^{\pm}(k,\mathbb{R})$ and ρ denotes the canonical left action of G on G/H.

Proof. One can proceed in analogy to the proof of theorem 24.1.3, and show that the two fiber spaces and the left actions they carry are identical. For this purpose, note that one can write every $g \in \mathrm{GL}(k,\mathbb{R})$ uniquely as

$$g = \mathrm{diag}(|\det g|, 1, \ldots, 1) \cdot h \tag{24.3.5}$$

with $h \in \mathrm{SL}^{\pm}(k,\mathbb{R})$. Hence, every coset gH is uniquely characterized by $|\det g| \in \mathbb{R}^+$. Using the relation $\det(gg') = \det g \cdot \det g'$, it follows that G acts on G/H by multiplication with $|\det g|$, which corresponds to the action on $\mathbb{R}^+$ defined by ρ^-_{-1}. ■

Using this isomorphism, we can identify twisted volume forms with sections of the density bundle D^-_{-1}, which are represented by positive numbers in any basis e^-_{-1} induced by a basis e of E. As in the case of orientations, this condition is independent of the choice of the basis e, since the sign does not change under a basis transformation, as can also be seen from the transformation rules (23.1.24). With this knowledge, we can identify those bases, or frames, with respect to which a given twisted volume form μ is represented by the number 1. We define them as follows.

Definition 24.3.2 (Normalized frame bundle)**.** Let $\pi : E \to M$ be a vector bundle of rank k equipped with a twisted volume form $\mu \in \Gamma(F(E) \times_{\rho^-_{-1}} \mathbb{R}^+)$ and $\varpi : F(E) \to M$ its general linear frame bundle. Its *normalized frame bundle* is the principal $\mathrm{SL}^{\pm}(k, \mathbb{R})$ bundle

$$\mathrm{SL}^{\pm}(E, \sigma) = \left\{p \in F(E), [p, 1]_{\rho^-_{-1}} = \mu(\varpi(p))\right\} \tag{24.3.6}$$

of normalized frames, together with the right action obtained by restricting the right action of $\mathrm{GL}(k, \mathbb{R})$ on $F(E)$ to $\mathrm{SL}^{\pm}(k, \mathbb{R})$.

Note that in contrast to the oriented and unit frame bundles, the fibers of the normalized frame bundle have two connected components, as also the structure group $\mathrm{SL}^{\pm}(k, \mathbb{R})$ acting on these fibers has two connected components. This relates to the fact that the twisted volume form prescribes a magnitude for normalized frames, but no preferred orientation, and hence the normalized frame bundle contains elements from each connected component of the fibers of the general linear frame bundle. As for the other bundles we considered, we can then state the following relation between twisted volume forms and frame bundle reductions.

Theorem 24.3.3. *There is a one-to-one correspondence between twisted volume forms of a vector bundle $\pi : E \to M$ of rank k and $\mathrm{SL}^{\pm}(k, \mathbb{R})$-reductions of its general linear frame bundle $\varpi : F(E) \to M$.*

Proof. We can once again proceed in analogy to the proof of theorem 24.1.2, by now considering a twisted volume form $\mu \in \Gamma(F(E) \times_{\rho^-_{-1}} \mathbb{R}^+)$ and $x \in M$. Here we call a frame $p \in F_x(E)$ over x a *normalized frame* if and only if

$$[p, 1]_{\rho^-_{-1}} = \mu(\varpi(p)) = \mu(x)\,. \tag{24.3.7}$$

It is easy to show that oriented frames exist for all $x \in M$. It follows from the definition of the associated bundle $F(E) \times_{\rho^-_0} \mathbb{R}^+$ that for *any* frame $p \in F_x(E)$ there exists a unique $c \in \mathbb{R}^+$ such that

$$[p, c]_{\rho^-_{-1}} = \mu(x)\,. \tag{24.3.8}$$

We may then consider the frame p' obtained from p by replacing its first component p_1 by $p'_1 = cp_1$, while all other components remain unchanged, and again obtain a frame. It follows that this new frame p' is normalized, since

$$[p, c]_{\rho^-_{-1}} = [p' \cdot g, c]_{\rho^-_{-1}} = [p', c|\det g|]_{\rho^-_{-1}} = [p', cc^{-1}]_{\rho^-_{-1}} = [p', 1]_{\rho^-_{-1}} = \mu(x)\,, \tag{24.3.9}$$

where $g = \mathrm{diag}(c^{-1}, 1, \ldots, 1)$, and used the fact that $\det g = c > 0$, so that $|c| = c$. Further, it holds that if $p \in F_x(E)$ is normalized, then $p \cdot g$ is normalized if and only if $|\det g| = 1$, and hence $g \in \mathrm{SL}^{\pm}(k, \mathbb{R})$. For every $x \in M$, we thus find that the set of normalized frames over x carries a right action of $\mathrm{SL}^{\pm}(k, \mathbb{R})$, which is obtained by the restriction of the action of $\mathrm{GL}(k, \mathbb{R})$ on $F_x(E)$, and that this action is free and transitive. From the smoothness of μ one finds that these sets constitute the fibers of a principal $\mathrm{SL}^{\pm}(k, \mathbb{R})$-bundle over M, which we will denote

$\mathrm{SL}^{\pm}(E,\mu)$. Finally, one finds that the canonical inclusion $\mathrm{SL}^{\pm}(E,\mu) \hookrightarrow F(E)$, together with the inclusion $\mathrm{SL}^{\pm}(k,\mathbb{R}) \hookrightarrow \mathrm{GL}(k,\mathbb{R})$ define a principal bundle reduction of $F(E)$.

To show the converse direction, let $\chi : Q \to M$ be a principal $\mathrm{SL}^{\pm}(k,\mathbb{R})$ bundle and $\phi : Q \to F(E)$ a $\mathrm{SL}^{\pm}(k,\mathbb{R})$ reduction over the canonical inclusion $\mathrm{SL}^{\pm}(k,\mathbb{R}) \hookrightarrow \mathrm{GL}(k,\mathbb{R})$. For $q \in Q$ with $\chi(q) = x$, define

$$\mu(x) = [\phi(q), 1]_{\rho^-_{-1}} . \tag{24.3.10}$$

This is independent of the choice of the representative, since for any other $q' = q \cdot g$ with $g \in \mathrm{SL}^{\pm}(k,\mathbb{R})$ we have $|\det g| =$ and hence

$$[\phi(q'), 1]_{\rho^-_{-1}} = [\phi(q \cdot g), 1]_{\rho^-_{-1}} = [\phi(q) \cdot g, 1]_{\rho^-_{-1}} = [\phi(q), |\det g|]_{\rho^-_{-1}} = [\phi(q), 1]_{\rho^-_{-1}} = \mu(x) . \tag{24.3.11}$$

One finds that this defines a section μ of $F(E) \times_{\rho^-_{-1}} \mathbb{R}^+$, and hence a twisted volume form on $\pi : E \to M$.

Finally, one can see from the two constructions that the bundle $\mathrm{SL}^{\pm}(E,\mu)$ constructed above is simply the image of Q under ϕ in $F(E)$, which can canonically be identified with Q, since ϕ restricts to an isomorphism of principal $\mathrm{SL}^{\pm}(k,\mathbb{R})$ bundles. ∎

Finally, we discuss an interesting relation between (untwisted) volume forms, orientations and twisted volume forms, which we state as follows.

Theorem 24.3.4. *On every vector bundle $\pi : E \to M$, there exists a one-to-one correspondence between volume forms ω and pairs (σ,μ) of orientations and twisted volume forms.*

Proof. Given a twisted volume form $\mu \in \Gamma(D^-_{-1}(E))$ and an orientation $\sigma \in \Gamma(D^-_0(E))$, we may define $\omega = \sigma \otimes \mu \in \Gamma(D^+_{-1}(E))$, using the bundle isomorphism

$$D^-_0(E) \otimes D^-_{-1}(E) \cong D^+_{-1}(E) , \tag{24.3.12}$$

which follows from theorem 23.1.5. Since μ and σ, by definition, are nowhere vanishing sections of their respective density bundles, the same holds also for ω, and so ω is a volume form. Using the equivalence between reductions of the frame bundle and the aforementioned sections, we can obtain another, more geometric picture for this construction. One finds that the bundles, and corresponding reductions, defined by these sections are related by

$$\mathrm{SL}(E,\omega) = \mathrm{GL}^+(E,\sigma) \cap \mathrm{SL}^{\pm}(E,\mu) . \tag{24.3.13}$$

Hence, a frame is a unit frame, if it is both oriented and normalized.

Conversely, we show that every volume form $\omega \in \Gamma(D^+_{-1}(E))$ uniquely decomposes into a twisted volume form $\mu \in \Gamma(D^-_{-1}(E))$ and an orientation $\sigma \in \Gamma(D^-_0(E))$ in the form $\omega = \sigma \otimes \mu$. This decomposition can also most easily found from the geometric picture of bundle reductions. Given the unit frame bundle $\mathrm{SL}(E,\omega)$, one may consider the bundles

$$\mathrm{GL}^+(E,\sigma) = \{p \cdot g, p \in \mathrm{SL}(E,\omega), g \in \mathrm{GL}^+(k,\mathbb{R})\} , \tag{24.3.14a}$$

$$\mathrm{SL}^{\pm}(E,\mu) = \{p \cdot g, p \in \mathrm{SL}(E,\omega), g \in \mathrm{SL}^{\pm}(k,\mathbb{R})\} . \tag{24.3.14b}$$

It follows that these bundles define reductions of the frame bundle, so that σ is an orientation, while μ is a twisted volume form, and $\omega = \sigma \otimes \mu$. More explicitly, for any $x \in M$, we can pick a unit frame $p \in F_x(E) \cap \mathrm{SL}(E,\omega)$, and define

$$\mu(x) = [p, 1]_{\rho^-_{-1}} , \quad \sigma(x) = [p, 1]_{\rho^-_0} . \tag{24.3.15}$$

These definitions are independent of the choice of p, since any other unit frame p' over x is related by $p' = p \cdot g$ with $\det g = 1$. ∎

The construction above conveys the intuition that a non-zero real number uniquely decomposes into its sign and absolute value to volume forms, whose basis representation is likewise given by non-zero real numbers. The remarkable fact is that this decomposition is independent of the choice of a basis.

Finally, we remark that the bundles (24.3.14) can also be obtained as extensions of $\mathrm{SL}(E, \omega)$ via the canonical inclusions $\mathrm{SL}(k, \mathbb{R}) \hookrightarrow \mathrm{GL}^+(k, \mathbb{R})$ and $\mathrm{SL}(k, \mathbb{R}) \hookrightarrow \mathrm{SL}^\pm(k, \mathbb{R})$, following the treatment detailed in section 20.7. To see this, note that we can canonically identify $p \cdot g$ with the equivalence class

$$[p, g] = \{(p \cdot h, h^{-1} g), h \in \mathrm{SL}(k, \mathbb{R})\} \tag{24.3.16}$$

for $p \in \mathrm{SL}(E, \omega)$ and $g \in \mathrm{GL}^+(k, \mathbb{R})$ or $g \in \mathrm{SL}^\pm(k, \mathbb{R})$, respectively.

24.4 Metrics

24.5 Almost symplectic structures

24.6 Almost complex structures

24.7 Almost Hermitian structures

24.8 Almost product structures

Chapter 25

Integration

25.1 Integrals over curve segments

We start our discussion of integrals with the simplest possible case, which should remind us to an integral over a single variable. From elementary calculus we know the meaning of integrals of the form

$$F(b) - F(a) = \int_a^b f(x)\mathrm{d}x\,. \tag{25.1.1}$$

This looks like an integral of a function $f(x)$ over the one-dimensional manifold $\mathbb{R}$, and so one may be tempted to define a way to integrate functions over (one-dimensional) manifolds. However, doing so would very soon cause a lot of trouble. To see this, consider the famous change-of-variable formula for integrals. Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a smooth function with $\varphi'(x) > 0$ for all $x \in \mathbb{R}$. Then we have

$$\int_{\varphi(a)}^{\varphi(b)} f(y)\mathrm{d}y = \int_a^b (f \circ \varphi)(x)\varphi'(x)\mathrm{d}x\,. \tag{25.1.2}$$

Note the appearance of the factor $\varphi'(x) = \mathrm{d}y/\mathrm{d}x$, which tells us that $f(y)\mathrm{d}y$ is a 1-form. In other words, the objects we can and should integrate over a curve segments will not be functions, but 1-forms. This leads us to the following definition.

Definition 25.1.1 (Integration on $\mathbb{R}$)**.** Let $\omega = f\,\mathrm{d}x$ be a 1-form on the one-dimensional manifold $\mathbb{R}$. Its *integral* over the interval $[a, b] \subset \mathbb{R}$ is defined as

$$\int_{[a,b]} \omega = \int_a^b f(x)\mathrm{d}x\,, \tag{25.1.3}$$

where the right hand side is to be interpreted in the obvious way.

Note that the integral above of course depends on the order of the bounds a and b. Changing them would reverse the sign of the integral. For an (oriented) interval one has $a < b$, so that there is a unique prescription how the bounds on the integral must be ordered.

To see how this definition works together with the change-of-variable formula, note that the function $\varphi : \mathbb{R} \to \mathbb{R}$ defined above is simply a diffeomorphism from the manifold $\mathbb{R}$ to itself. It then follows that

$$\int_{\varphi([a,b])} \omega = \int_{\varphi(a)}^{\varphi(b)} f(y)\mathrm{d}y = \int_a^b f(\varphi(x))\varphi'(x)\mathrm{d}x = \int_{[a,b]} \varphi^*(\omega)\,, \tag{25.1.4}$$

which shows that the integral as we defined it is invariant under a pullback by φ. This is an important property which enters the proofs of many of the statements we encounter during this chapter. It is also important that we demanded $\varphi' > 0$, since otherwise the order of the bounds would get reversed.

Of course we want to perform integration not only on $\mathbb{R}$ (or intervals), but also on general manifolds. To see how we can get there, we first take something we already know (an interval) and stick it into a manifold.

Definition 25.1.2 (Singular curve segment)**.** A *singular curve segment* on a smooth manifold M is a smooth function $c : [0, 1] \to M$. We denote the space of singular curve segments on M by $\mathcal{K}_1(M)$.

By "singular" we mean that we make no assumption on c being injective, but leave it arbitrary. We already know that we can integrate 1-forms on an interval, so we need a prescription which yields us a 1-form on $[0, 1]$ if we have a curve segment. If we have a 1-form on the target manifold M, then we can simply take its pullback. This leads us to the following definition.

Definition 25.1.3 (Integral over a curve segment)**.** Let M be a manifold and $\omega \in \Omega^1(M)$. For a curve segment $c \in \mathcal{K}_1(M)$, the *integral* of ω over c is defined as

$$\int_c \omega = \int_{[0,1]} c^*(\omega) \,. \tag{25.1.5}$$

This prescription now allows us to integrate 1-forms on a manifold along a curve segment. Again we raise the question how the change-of-variable formula and diffeomorphisms work together with this definition. Since we have fixed the interval of integration to be $[0, 1]$, we will consider only diffeomorphisms of $\mathbb{R}$ which leave this interval unchanged. We define them as follows.

Definition 25.1.4 (Reparametrization of the unit interval)**.** A reparametrization of the unit interval $[0, 1]$ is a smooth function $\varphi : [0, 1] \to [0, 1]$ such that $\varphi(0) = 0$, $\varphi(1) = 1$ and $\varphi'(x) > 0$ for all $x \in [0, 1]$.

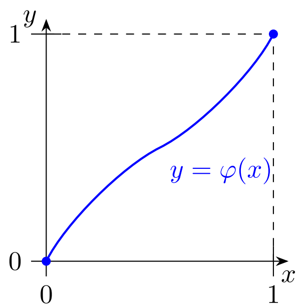


Figure 25.1: A reparametrization $\varphi : [0, 1] \to [0, 1]$ of the unit interval.

We have restricted ourselves to *orientation preserving* reparametrizations here by demanding $\varphi' > 0$ everywhere. If we would also allow the case $\varphi' < 0$ everywhere, this would be an *orientation reversing* reparametrization, which would also swap the endpoints of the interval, as we discussed above. We keep things more simple by not considering this case. Now it is easy to see what happens with our integral if we apply a reparametrization.

Theorem 25.1.1. *The integral over a curve segment is invariant under reparametrization,*

$$\int_c \omega = \int_{c\circ\varphi} \omega \tag{25.1.6}$$

for all curve segments c, 1-forms ω and reparametrizations φ.

Proof. By definition we have

$$\int_{c\circ\varphi} \omega = \int_{[0,1]} (c\circ\varphi)^*(\omega) = \int_{[0,1]} \varphi^*(c^*(\omega)) = \int_{[0,1]} c^*(\omega) = \int_c \omega\,, \tag{25.1.7}$$

where we used the fact that $\varphi([0,1]) = [0,1]$ by definition. ■

Intuitively this means that the integral depends only on the path traced out by the curve segment, but not on the velocity with which this path is transversed.

25.2 Integrals over k-cubes

We now generalize our knowledge from the previous section from curve segments to k-cubes, and from intervals on the real line to boxes in Euclidean space $\mathbb{R}^k$. It should be clear from what we have learned that the objects which we can integrate on $\mathbb{R}^k$ must be k-forms, since they have the correct transformation behavior under a change of integration variables. We define integrals in analogy to the one-dimensional case.

Definition 25.2.1 (Integration over boxes on $\mathbb{R}^k$)**.** Let $\omega = f\,\mathrm{d}x^1\wedge\ldots\wedge\mathrm{d}x^k$ be a k-form on the k-dimensional manifold $\mathbb{R}^k$. Its *integral* over the rectangular box $[a^1,b^1]\times\ldots\times[a^k,b^k]\subset\mathbb{R}^k$ is defined as

$$\int_{[a^1,b^1]\times\ldots\times[a^k,b^k]} \omega = \int_{a^k}^{b^k}\cdots\int_{a^1}^{b^1} f(x)\mathrm{d}x^1\cdots\mathrm{d}x^k\,, \tag{25.2.1}$$

where the integrals on the right hand side are evaluated from the inside outwards.

In order to apply this knowledge to the integration of k-forms on manifolds, we first need to transfer the box over which we integrate into the manifold. In other words, we need to define the analogy of a singular curve segment. We will use cubes here, because the formulas will become easier - an alternative formulation uses simplices, shown in section 25.3.

Definition 25.2.2 (Singular k-cube)**.** A *singular k-cube* on a smooth manifold M is a smooth function $c:[0,1]^k\to M$. We denote the space of singular k-cubes on M by $\mathcal{K}_k(M)$.

Again by singular mean that c will not necessarily be injective. It should now be clear how to integrate a k-form over a k-cube, so we just provide the definition.

Definition 25.2.3 (Integral over a k-cube)**.** Let M be a manifold and $\omega \in \Omega^k(M)$. For a k-cube $c \in \mathcal{K}_k(M)$, the *integral* of ω over c is defined as

$$\int_c \omega = \int_{[0,1]^k} c^*(\omega)\,. \tag{25.2.2}$$

We finally come to the question how this definition of integrals behaves under reparametrizations. For this purpose we first need to generalize our definition of reparametrizations.

Definition 25.2.4 (Reparametrization of the unit cube)**.** A reparametrization of the unit cube $[0,1]^k$ is a smooth, bijective function $\varphi : [0,1]^k \to [0,1]^k$ such that $\det \mathrm{D}\varphi(x) > 0$ for all $x \in [0,1]^k$.

Again we restrict ourselves to orientation preserving reparametrizations, which in this case means that the Jacobi determinant $\det \mathrm{D}\varphi$ must be everywhere positive. Without this restriction the sign of the integral would change. We give an illustrative example.

Example 25.2.1 (Rotation of a 2-cube). The function

$$\begin{array}{cccc} \varphi & : & [0,1]^2 & \to & [0,1]^2 \\ & & (x^1,x^2) & \mapsto & (1-x^2,x^1) \end{array}\,, \tag{25.2.3}$$

which turns the unit cube by 90°, is a reparametrization. It changes a cube $c : [0,1]^2 \to M$ to $c \circ \varphi$, i.e., such that

$$c(y^1,y^2) = (c\circ\varphi)(x^1,x^2) = c(1-x^2,x^1)\,. \tag{25.2.4}$$

This is illustrated in figure 25.2.

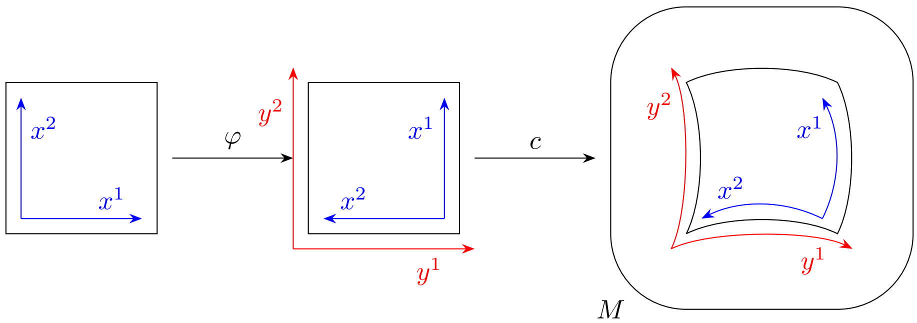


Figure 25.2: A reparametrization $\varphi : [0,1]^2 \to [0,1]^2$ of the unit cube and its application to a singular 2-cube $c : [0,1]^2 \to M$.

Note that we did not demand that φ restricts to the identity on the boundary of the cube, but we allowed for boundary points to be displaced; in contrast, for the reparametrization of the unit interval in definition 25.1.4 we demanded that the endpoints (which are the boundary of the interval) remain fixed. However, it turns out that the latter is actually a consequence of the remaining requirements in the case $k = 1$. In this case the Jacobian $D\varphi$ and its determinant simply reduce to φ'. From $\varphi' > 0$ then follows that φ is a monotonically increasing function, and so it must map the minimum 0 of the domain $[0, 1]$ to the minimum of the image; by demanding that φ is bijective, the latter is also $[0, 1]$, and hence $\varphi(0) = 0$. Analogously follows $\varphi(1) = 1$.

We finally come to the important result of this section.

Theorem 25.2.1. *The integral over a k-cube is invariant under reparametrization,*

$$\int_c \omega = \int_{c\circ\varphi} \omega \tag{25.2.5}$$

for all k-cubes c, k-forms ω and reparametrizations φ.

Proof. ►...◄ ■

25.3 Integrals over k-simplices

While integration over cubes is most easily defined, it is sometimes more practical to work with simplices instead. Recall from Euclidean geometry that a k-simplex is a k-dimensional polytope spanned by $k + 1$ points $p_0, \ldots, p_k \in \mathbb{R}^k$, where each pair of points is connected by an edge. For simplicity, we will choose the coordinate system such that the x^i coordinate points along the edge from p_{i-1} to p_i, and that the points $p_0 = (a, \ldots, a)$ and $p_k = (b, \ldots, b)$ lie at opposite corners of a cube with $a < b$. Then the simplex contains the points

$$\{x \in \mathbb{R}^k, a \leq x^1 \leq \ldots \leq x^k \leq b\}. \tag{25.3.1}$$

Now we can integrate a k-form as follows.

Definition 25.3.1 (Integration over simplices on $\mathbb{R}^k$)**.** Let $\omega = f\,dx^1 \wedge \ldots \wedge dx^k$ be a k-form on the k-dimensional manifold $\mathbb{R}^k$. Its *integral* over the simplex (25.3.1) is defined as

$$\int_{\{x\in\mathbb{R}^k, a\leq x^1\leq\ldots\leq x^k\leq b\}} \omega = \int_a^b \int_a^{x^k} \cdots \int_a^{x^3} \int_a^{x^2} f(x)dx^1 \cdots dx^k\,, \tag{25.3.2}$$

where the integrals on the right hand side are evaluated from the inside outwards.

In order to make use of this relation to integrate differential forms on manifolds, we reproduce the same steps which we have performed for cubes in section 25.2. Instead of a unit cube, we will now need a "unit simplex", which we define as follows.

Definition 25.3.2 (Unit k-simplex)**.** For $k \in \mathbb{N}$, the *unit k-simplex* is the set

$$\Delta^k = \{x \in \mathbb{R}^k, 0 \leq x^1 \leq \ldots \leq x^k \leq 1\}. \tag{25.3.3}$$

The unit simplices Δ^2 and Δ^3 are displayed in figure 25.3. With this definition in place, we can proceed with the following definition.

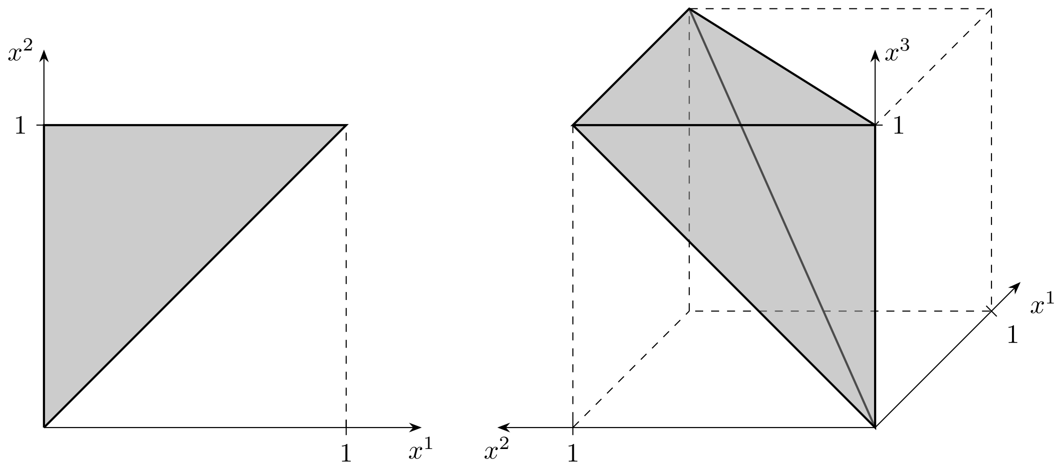


Figure 25.3: Unit simplices Δ^2 and Δ^3.

Definition 25.3.3 (Singular k-simplex)**.** A *singular k-simplex* on a smooth manifold M is a smooth function $c : \Delta^k \to M$. We denote the space of singular k-simplices on M by $\mathcal{S}_k(M)$.

As in the previous cases, by singular we mean that the function c does not have to be injective. Now the definition of the integral is straightforward.

Definition 25.3.4 (Integral over a k-simplex)**.** Let M be a manifold and $\omega \in \Omega^k(M)$. For a k-simplex $c \in \mathcal{S}_k(M)$, the *integral* of ω over c is defined as

$$\int_c \omega = \int_{\Delta^k} c^*(\omega)\,. \tag{25.3.4}$$

Finally, also for simplices we discuss how this definition of integrals behaves under reparametrizations. We therefore need to provide a definition for the reparametrization of a simplex.

Definition 25.3.5 (Reparametrization of the unit simplex)**.** A reparametrization of the unit simplex Δ^k is a smooth, bijective function $\varphi : \Delta^k \to \Delta^k$ such that $\det \mathrm{D}\varphi(x) > 0$ for all $x \in \Delta^k$.

When we write a reparametrization in coordinates, we have to keep in mind that their values form an ascending sequence for points in Δ^k. Here we give an illustrating example.

Example 25.3.1 (Rotation of a 2-simplex)*.* The functions

$$\varphi \quad : \quad \begin{array}{ccc} \Delta^2 & \to & \Delta^2 \\ (x^1, x^2) & \mapsto & (1 - x^2, 1 + x^1 - x^2) \end{array}\,, \tag{25.3.5}$$

and

$$\vartheta : \begin{array}{ccc} \Delta^2 & \to & \Delta^2 \\ (x^1, x^2) & \mapsto & (x^2 - x^1, 1 - x^1) \end{array} , \tag{25.3.6}$$

which turn the unit simplex by one third in either direction, are reparametrizations. They changes a simplex $c : \Delta^2 \to M$ to $c \circ \varphi$ and $c \circ \vartheta$, respectively, i.e., such that

$$c(y^1, y^2) = (c \circ \varphi)(x^1, x^2) = c(1 - x^2, 1 + x^1 - x^2) , \tag{25.3.7a}$$

$$c(y^1, y^2) = (c \circ \vartheta)(x^1, x^2) = c(x^2 - x^1, 1 - x^1) . \tag{25.3.7b}$$

This is illustrated in figure 25.4.

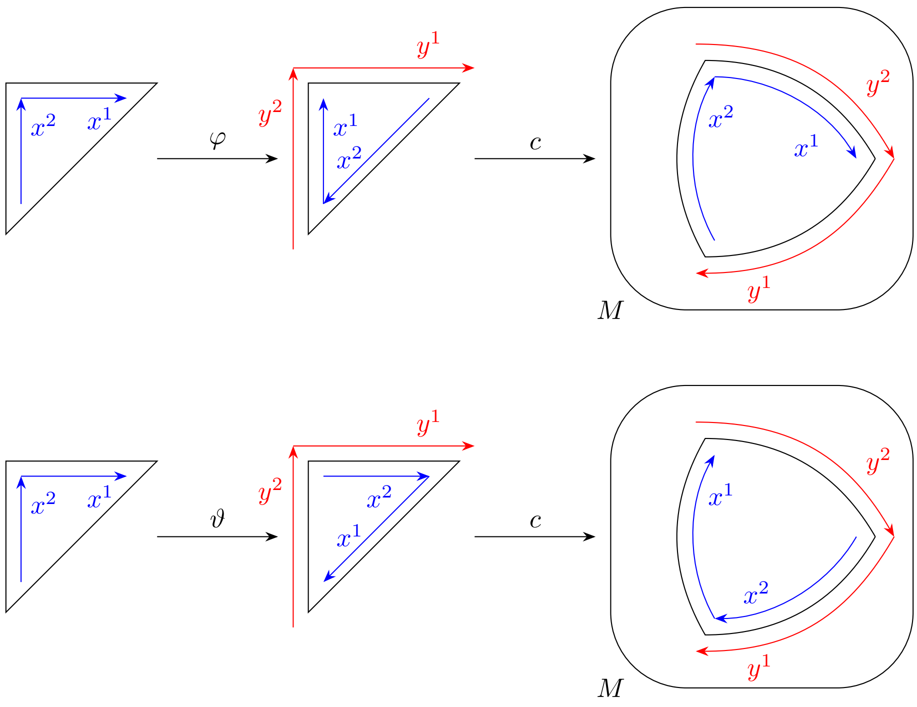


Figure 25.4: Two reparametrizations $\varphi : \Delta^2 \to \Delta^2$ of the unit simplex and their application to a singular 2-simplex $c : \Delta^2 \to M$.

Finally, we can show the following.

Theorem 25.3.1. *The integral over a k-simplex is invariant under reparametrization,*

$$\int_c \omega = \int_{c \circ \varphi} \omega \tag{25.3.8}$$

for all k-simplices c, k-forms ω and reparametrizations φ.

Proof. ▶. . . ◀ ■

25.4 Integrals over k-chains

So far we have learned how to integrate k-forms over regions which can be parametrized by singular k-cubes or singular k-simplices. We now wish to further generalize this concept to more general regions. One straightforward possibility would be to consider other polytopes. While it is obviously possible to integrate over arbitrary polytopes in $\mathbb{R}^k$ by suitably restricting the domain of the integral, and decompose it into several pieces if necessary, it is more fruitful to follow a different approach, by considering domains which may be composed of several k-cubes or k-simplices, some of which may be transversed multiple times or in different orientation, as it turns out that such domains will occur naturally in certain constructions, and we can always decompose more involved polytopes into simplices. The mathematical object corresponding to such a composite integration domain is a chain, which we define as follows.

Definition 25.4.1 (Chain)**.** A *k-chain* over a set $\mathcal{P}_k(M)$ of singular k-polytopes on a smooth manifold M is an element of the free abelian group $\mathcal{C}_k(M)$ generated by the set $\mathcal{P}_k(M)$. If $\mathcal{P}_k(M) = \mathcal{K}_k(M)$ is the set of k-cubes, the elements of $\mathcal{C}_k$ are called *cubical chains*. If $\mathcal{P}_k(M) = \mathcal{S}_k(M)$ is the set of k-simplices, the elements of $\mathcal{C}_k$ are called *simplical chains*.

We first need to clarify the notion of a free abelian group. The elements of $\mathcal{C}_k(M)$ are finite formal sums of elements of $\mathcal{P}_k(M)$ with integer coefficients, i.e., a chain $C \in \mathcal{C}_k(M)$ can be written as

$$C = \sum_{c \in \mathcal{P}_k(M)} C_c c \tag{25.4.1}$$

with integer coefficients C_c such that only finitely many C_c are non-zero. The group operation is the addition of formal sums and the group inversion is the negative,

$$C + C' = \sum_{c \in \mathcal{P}_k(M)} (C_c + C'_c)c\,, \quad -C = \sum_{c \in \mathcal{P}_k(M)} (-C_c)c\,. \tag{25.4.2}$$

Hence, a k-chain C is uniquely determined by a function

$$C_\bullet \quad : \quad \begin{array}{ccc} \mathcal{P}_k(M) & \to & \mathbb{Z} \\ c & \mapsto & C_c \end{array}\,, \tag{25.4.3}$$

such that only finitely many C_c are non-zero, and addition is defined pointwise.

A chain can thus be interpreted as a prescription which k-cubes should be transversed, how often and with which orientation. Note that every k-cube can also be interpreted as a k-chain which simply prescribes to transverse only this k-cube and exactly once with positive orientation. It is clear that this chain with $c \in \mathcal{P}_k(M) \subset \mathcal{C}_k(M)$ takes the form

$$c = \sum_{c' \in \mathcal{P}_k(M)} c' \cdot \begin{cases} 1 & \text{if } c = c' \\ 0 & \text{otherwise} \end{cases}\,. \tag{25.4.4}$$

Integration over a chain is then simply the integration over all of its constituents c, where each integral is weighted with the integer coefficient C_c.

Definition 25.4.2 (Integral over a chain)**.** Let M be a manifold and $\omega \in \Omega^k(M)$. For a k-chain $C \in \mathcal{C}_k(M)$, the *integral* of ω over C is defined as

$$\int_C \omega = \sum_{c \in \mathcal{P}_k(M)} C_c \int_c \omega\,. \tag{25.4.5}$$

25.5 Boundary of a k-chain

An interesting property of polytopes, which also transcends to k-chains, is the existence of a boundary, which is a $(k-1)$-chain. To construct this boundary, we will start by constructing the boundary of its constituents. For a cubical chain, these are cubes, and so we start with the following definition.

Definition 25.5.1 (Boundary of a k-cube)**.** The boundary of a singular k-cube $c \in \mathcal{K}_k(M)$ is the cubical $(k-1)$-chain $\partial c \in \mathcal{C}_{k-1}(M)$ defined as

$$\partial c = \sum_{i=1}^{k} (-1)^i (c_{(i,0)} - c_{(i,1)}) \,, \tag{25.5.1}$$

where $c_{(i,y)} \in \mathcal{K}_{k-1}(M)$ is the *facet* of c defined by

$$c_{(i,y)} = c \circ F_{(i,y)} \tag{25.5.2}$$

with

$$F_{(i,y)} \; : \; \begin{array}{ccc} [0,1]^{k-1} & \to & [0,1]^k \\ (x^1, \ldots, x^{k-1}) & \mapsto & (x^1, \ldots, x^{i-1}, y, x^i, \ldots, x^{k-1}) \end{array} . \tag{25.5.3}$$

The maps $F_{(i,y)}$ defining the facets of the singular k-cube can also be understood as fixing the i'th coordinate of the unit cube $[0,1]^k$ to be equal to y, and relabelling the remaining $k-1$ coordinates by the coordinates of $[0,1]^{k-1}$. We can illustrate this with a simple example.

Example 25.5.1 (Boundary of a 2-cube). Let $c : [0,1]^2 \to M$ be a singular 2-cube. Following definition 25.5.1, its boundary is the chain

$$\partial c = c_{(1,1)} - c_{(1,0)} + c_{(2,0)} - c_{(2,1)} \,, \tag{25.5.4}$$

where the facets $c_{(i,y)} : [0,1] \to M$ are defined by

$$c_{(1,0)}(x) = c(0,x) \,, \quad c_{(1,1)}(x) = c(1,x) \,, \quad c_{(2,0)}(x) = c(x,0) \,, \quad c_{(2,1)}(x) = c(x,1) \,. \tag{25.5.5}$$

This is illustrated in figure 25.5.

In case we have a simplical chain, the constituents will be simplices instead of cubes. In this case, we will use the following definition of a boundary.

Definition 25.5.2 (Boundary of a k-simplex)**.** The boundary of a singular k-simplex $c \in \mathcal{S}_k(M)$ is the simplical $(k-1)$-chain $\partial c \in \mathcal{C}_{k-1}(M)$ defined as

$$\partial c = \sum_{i=0}^{k} (-1)^{i+1} c_i \,, \tag{25.5.6}$$

where $c_i \in \mathcal{S}_{k-1}(M)$ is the *facet* of c defined by

$$c_i = c \circ F_i \,, \tag{25.5.7}$$

where the maps $F_i : \Delta^{k-1} \to \Delta^k$ are defined by

$$F_0(x^1, \ldots, x^{k-1}) = (0, x^1, x^2, \ldots, x^{k-1}) \,, \tag{25.5.8a}$$

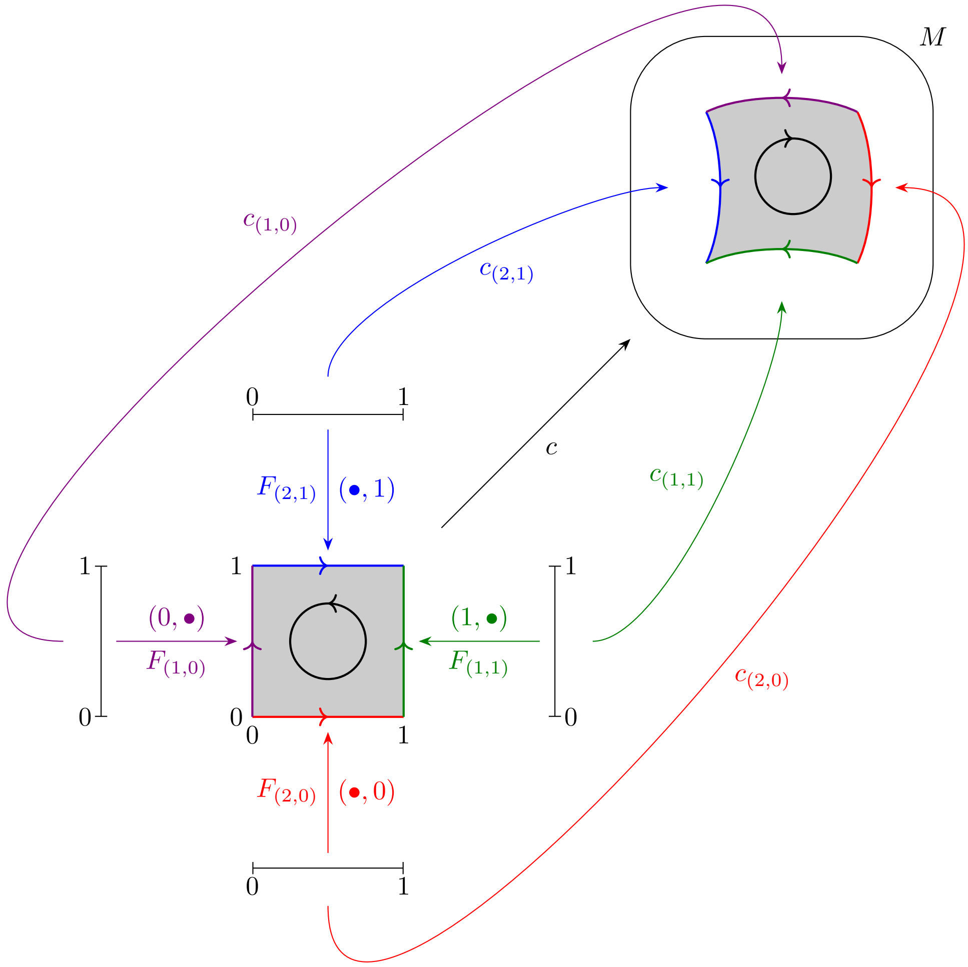


Figure 25.5: Illustration of the boundary $\partial c = c_{(1,1)} - c_{(1,0)} + c_{(2,0)} - c_{(2,1)}$ of a singular 2-cube c. The latter defines an oriented "patch" of the manifold M, whose orientation is indicated by a circular arrow. The arrows on the facets $c_{(i,y)}$ indicate their orientation. Note that the orientation of the facets $c_{(1,0)}$ and $c_{(2,1)}$ is against the orientation of the boundary, and so they must contribute with a negative sign.

$$F_1(x^1, \ldots, x^{k-1}) = (x^1, x^1, x^2, \ldots, x^{k-1}), \tag{25.5.8b}$$
$$F_2(x^1, \ldots, x^{k-1}) = (x^1, x^2, x^2, \ldots, x^{k-1}), \tag{25.5.8c}$$
$$\vdots \tag{25.5.8d}$$
$$F_{k-2}(x^1, \ldots, x^{k-1}) = (x^1, x^2, x^3, \ldots, x^{k-2}, x^{k-2}, x^{k-1}), \tag{25.5.8e}$$
$$F_{k-1}(x^1, \ldots, x^{k-1}) = (x^1, x^2, x^3, \ldots, x^{k-2}, x^{k-1}, x^{k-1}), \tag{25.5.8f}$$
$$F_k(x^1, \ldots, x^{k-1}) = (x^1, x^2, x^3, \ldots, x^{k-2}, x^{k-1}, 1). \tag{25.5.8g}$$

To understand the maps F_i, recall that the unit k-simplex Δ^k by definition contains those points $x \in \mathbb{R}^k$ which satisfy

$$0 \leq x^1 \leq \ldots \leq x^k \leq 1, \tag{25.5.9}$$

and so it is the intersection of the $k+1$ half-spaces

$$\Delta^k = \{x \in \mathbb{R}^k, 0 \leq x^1\} \cap \{x \in \mathbb{R}^k, x^1 \leq x^2\} \cap \ldots \cap \{x \in \mathbb{R}^k, x^{k-1} \leq x^k\} \cap \{x \in \mathbb{R}^k, x^k \leq 1\} . \tag{25.5.10}$$

Note that there are $k+1$ inequalities in the condition (25.5.9), which we shall label from 0 to k. The map F_i is obtained if we replace the i'th inequality sign by an equality, thus fixing either the first or last coordinate equal to 0 or 1, respectively, or two successive coordinates to be equal to each other, and relabel the remaining $k-1$ independent coordinates in ascending order to match those of Δ^{k-1}. Also this definition can be illustrated with a simple example.

Example 25.5.2 (Boundary of a 2-simplex). Let $c : \Delta^2 \to M$ be a singular 2-simplex. Following definition 25.5.2, its boundary is the chain

$$\partial c = -c_0 + c_1 - c_2 , \tag{25.5.11}$$

where the facets $c_i : \Delta^1 \to M$ are defined by

$$c_0(x) = c(0, x) , \quad c_1(x) = c(x, x) , \quad c_2(x) = c(x, 1) . \tag{25.5.12}$$

This is illustrated in figure 25.6.

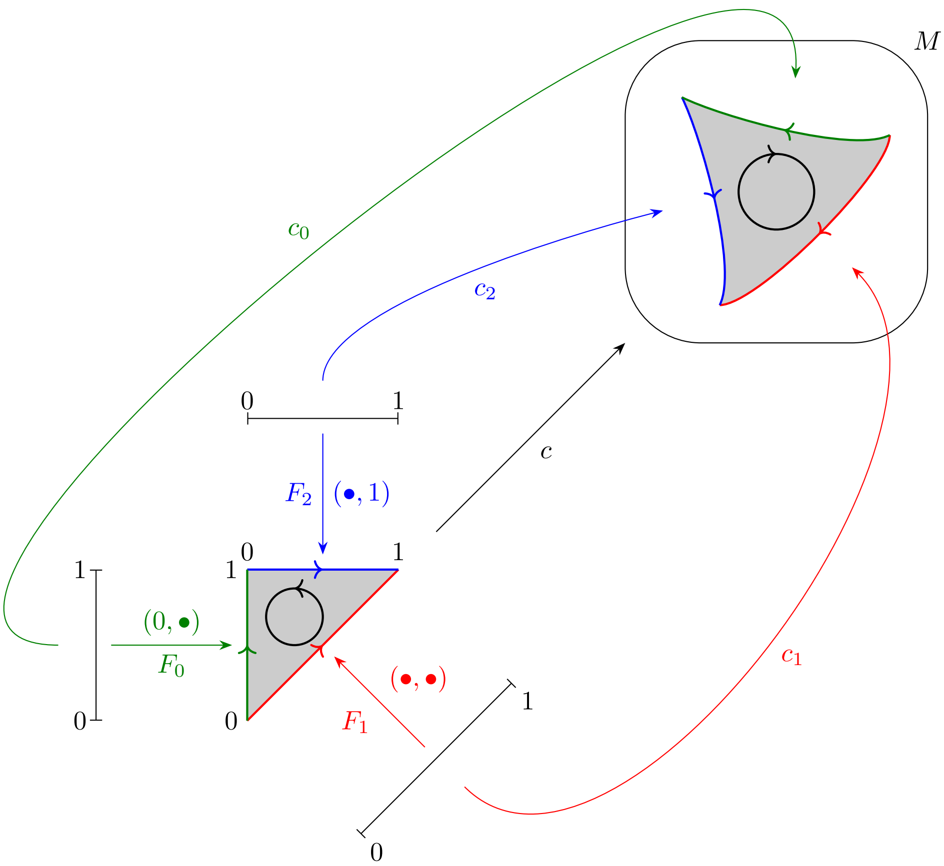


Figure 25.6: Illustration of the boundary $\partial c = -c_0 + c_1 - c_2$ of a singular 2-simplex c, whose orientation is indicated by a circular arrow. The arrows on the facets c_i indicate their orientation. Note that the orientation of the facets c_0 and c_2 is against the orientation of the boundary, and so they must contribute with a negative sign.

Having defined the boundary of the constituents, we can generalize these definitions to the boundary of a chain. This is straightforward.

Definition 25.5.3 (Boundary of a k-chain)**.** Let

$$C = \sum_{c \in \mathcal{P}_k(M)} C_c c \in \mathcal{C}_k(M)\,. \tag{25.5.13}$$

be a k-chain on a manifold M. Its *boundary* is the $(k-1)$-chain $\partial C \in \mathcal{C}_{k-1}(M)$ given by

$$\partial C = \sum_{c \in \mathcal{P}_k(M)} C_c \partial c\,. \tag{25.5.14}$$

Recall from the definitions 25.5.1 and 25.5.2 that we introduced particular signs for each of the facets which constitute the boundary. As we have seen in the examples 25.5.1 and 25.5.2, these signs ensure that when calculating the integral over the boundary, the contribution of each facet whose orientation is opposite to the orientation of the boundary is accounted with a negative sign. An interesting consequence of this accounting is a rather important property of the boundary operator, which we will explore next. We start again with the case of cubes.

Theorem 25.5.1. *The double boundary of a singular cube vanishes, $\partial^2 c = 0$ for all $c \in \mathcal{K}_k(M)$.*

Proof. By direct calculation, following the definitions given above, we have

$$\begin{aligned}
\partial\partial c &= \sum_{i=1}^{k} (-1)^i (\partial c_{(i,0)} - \partial c_{(i,1)}) \\
&= \sum_{i=1}^{k} \sum_{j=1}^{k-1} (-1)^{i+j} (c_{(i,0)(j,0)} - c_{(i,0)(j,1)} - c_{(i,1)(j,0)} + c_{(i,1)(j,1)})\,,
\end{aligned} \tag{25.5.15}$$

where the appearing terms are the $(k-2)$-cubes given by

$$\begin{aligned}
c_{(i,y)(j,z)}(x^1, \ldots, x^{k-2}) &= c_{(i,y)}(x^1, \ldots, x^{j-1}, z, x^j, \ldots, x^{k-2}) \\
&= \begin{cases} c(x^1, \ldots, x^{i-1}, y, x^i, \ldots, x^{j-1}, z, x^j, \ldots, x^{k-2}) & i \leq j\,, \\ c(x^1, \ldots, x^{j-1}, z, x^j, \ldots, x^{i-2}, y, x^{i-1}, \ldots, x^{k-2}) & i > j\,. \end{cases}
\end{aligned} \tag{25.5.16}$$

Now we see that there are always two $(k-2)$-cubes which only differ by the order in which y and z appear as arguments of c, while their dependence on the remaining arguments is the same. They can thus be transformed into each other by swapping y and z, together with an appropriate change of i and j. In particular, we have

$$c_{(i,y)(j,z)} = c_{(i',y')(j',z')} = \begin{cases} c_{(j+1,z)(i,y)} & i \leq j\,, \\ c_{(j,z)(i-1,y)} & i > j\,, \end{cases} \tag{25.5.17}$$

where $y' = z$, $z' = y$ and i', j' follow from the two cases given above. Note that this operation is symmetric: by exchanging primed and non-primed indices, we get the same relation in reverse. Also, we have $i' \leq j'$ if and only if $i > j$ and vice versa. Applying this to the double boundary formula, we have

$$
\begin{aligned}
\partial\partial c &= \sum_{i=1}^{k}\sum_{j=1}^{k-1}(-1)^{i+j}(c_{(i,0)(j,0)} - c_{(i,0)(j,1)} - c_{(i,1)(j,0)} + c_{(i,1)(j,1)}) \\
&= \sum_{i=2}^{k}\sum_{j=1}^{i-1}(-1)^{i+j}(c_{(i,0)(j,0)} - c_{(i,0)(j,1)} - c_{(i,1)(j,0)} + c_{(i,1)(j,1)}) \\
&\quad + \sum_{i=1}^{k-1}\sum_{j=i}^{k-1}(-1)^{i+j}(c_{(i,0)(j,0)} - c_{(i,0)(j,1)} - c_{(i,1)(j,0)} + c_{(i,1)(j,1)}) \\
&= \sum_{i=2}^{k}\sum_{j=1}^{i-1}(-1)^{i+j}(c_{(j,0)(i-1,0)} - c_{(j,1)(i-1,0)} - c_{(j,0)(i-1,1)} + c_{(j,1)(i-1,1)}) \\
&\quad + \sum_{i=1}^{k-1}\sum_{j=i}^{k-1}(-1)^{i+j}(c_{(i,0)(j,0)} - c_{(i,0)(j,1)} - c_{(i,1)(j,0)} + c_{(i,1)(j,1)}) \\
&= \sum_{j'=1}^{k-1}\sum_{i'=1}^{j'}(-1)^{j'+1+i'}(c_{(i',0)(j',0)} - c_{(i',1)(j',0)} - c_{(i',0)(j',1)} + c_{(i',1)(j',1)}) \\
&\quad + \sum_{i=1}^{k-1}\sum_{j=i}^{k-1}(-1)^{i+j}(c_{(i,0)(j,0)} - c_{(i,0)(j,1)} - c_{(i,1)(j,0)} + c_{(i,1)(j,1)}) \\
&= -\sum_{i'=1}^{k-1}\sum_{j'=i'}^{k-1}(-1)^{i'+j'}(c_{(i',0)(j',0)} - c_{(i',1)(j',0)} - c_{(i',0)(j',1)} + c_{(i',1)(j',1)}) \\
&\quad + \sum_{i=1}^{k-1}\sum_{j=i}^{k-1}(-1)^{i+j}(c_{(i,0)(j,0)} - c_{(i,0)(j,1)} - c_{(i,1)(j,0)} + c_{(i,1)(j,1)}) \\
&= 0\,. \quad \blacksquare
\end{aligned}
\tag{25.5.18}
$$

We illustrate this with the following example.

Example 25.5.3 (Double boundary of a 3-cube). Let $c : [0,1]^3 \to M$ be a singular 3-cube. Its boundary is the chain

$$\partial c = c_{(1,1)} - c_{(1,0)} + c_{(2,0)} - c_{(2,1)} + c_{(3,1)} - c_{(3,0)}\,, \tag{25.5.19}$$

where the facets $c_{(i,z)} : [0,1]^2 \to M$ are given by

$$c_{(1,0)}(x,y) = c(0,x,y)\,, \qquad c_{(1,1)}(x,y) = c(1,x,y)\,, \tag{25.5.20a}$$
$$c_{(2,0)}(x,y) = c(x,0,y)\,, \qquad c_{(2,1)}(x,y) = c(x,1,y)\,, \tag{25.5.20b}$$
$$c_{(3,0)}(x,y) = c(x,y,0)\,, \qquad c_{(3,1)}(x,y) = c(x,y,1)\,. \tag{25.5.20c}$$

Now their boundaries are given by

$$\partial c_{(i,z)} = c_{(i,z)(1,1)} - c_{(i,z)(1,0)} + c_{(i,z)(2,0)} - c_{(i,z)(2,1)}\,. \tag{25.5.21}$$

Hence, we have

$$
\begin{aligned}
\partial\partial c = {} & c_{(1,1)(1,1)} - c_{(1,0)(1,1)} + c_{(2,0)(1,1)} - c_{(2,1)(1,1)} + c_{(3,1)(1,1)} - c_{(3,0)(1,1)} \\
& - c_{(1,1)(1,0)} + c_{(1,0)(1,0)} - c_{(2,0)(1,0)} + c_{(2,1)(1,0)} - c_{(3,1)(1,0)} + c_{(3,0)(1,0)} \\
& + c_{(1,1)(2,0)} - c_{(1,0)(2,0)} + c_{(2,0)(2,0)} - c_{(2,1)(2,0)} + c_{(3,1)(2,0)} - c_{(3,0)(2,0)} \\
& - c_{(1,1)(2,1)} + c_{(1,0)(2,1)} - c_{(2,0)(2,1)} + c_{(2,1)(2,1)} - c_{(3,1)(2,1)} + c_{(3,0)(2,1)}\,.
\end{aligned}
\tag{25.5.22}
$$

Here the facets $c_{(i,z)(j,y)} : [0,1] \to M$ are given by the 1-cubes

$$c_{(1,0)(1,0)}(x) = c(0,0,x) = c_{(2,0)(1,0)}(x)\,, \qquad c_{(1,0)(1,1)}(x) = c(0,1,x) = c_{(2,1)(1,0)}(x)\,, \tag{25.5.23a}$$

$$c_{(1,0)(2,0)}(x) = c(0,x,0) = c_{(3,0)(1,0)}(x)\,, \qquad c_{(1,0)(2,1)}(x) = c(0,x,1) = c_{(3,1)(1,0)}(x)\,, \tag{25.5.23b}$$

$$c_{(1,1)(1,0)}(x) = c(1,0,x) = c_{(2,0)(1,1)}(x)\,, \qquad c_{(1,1)(1,1)}(x) = c(1,1,x) = c_{(2,1)(1,1)}(x)\,, \tag{25.5.23c}$$

$$c_{(1,1)(2,0)}(x) = c(1,x,0) = c_{(3,0)(1,1)}(x)\,, \qquad c_{(1,1)(2,1)}(x) = c(1,x,1) = c_{(3,1)(1,1)}(x)\,, \tag{25.5.23d}$$

$$c_{(2,0)(2,0)}(x) = c(x,0,0) = c_{(3,0)(2,0)}(x)\,, \qquad c_{(2,0)(2,1)}(x) = c(x,0,1) = c_{(3,1)(2,0)}(x)\,, \tag{25.5.23e}$$

$$c_{(2,1)(2,0)}(x) = c(x,1,0) = c_{(3,0)(2,1)}(x)\,, \qquad c_{(2,1)(2,1)}(x) = c(x,1,1) = c_{(3,1)(2,1)}(x)\,. \tag{25.5.23f}$$

Hence, we see that they appear in pairs which cancel each other, and so $\partial\partial c = 0$. This is illustrated in figure 25.7.

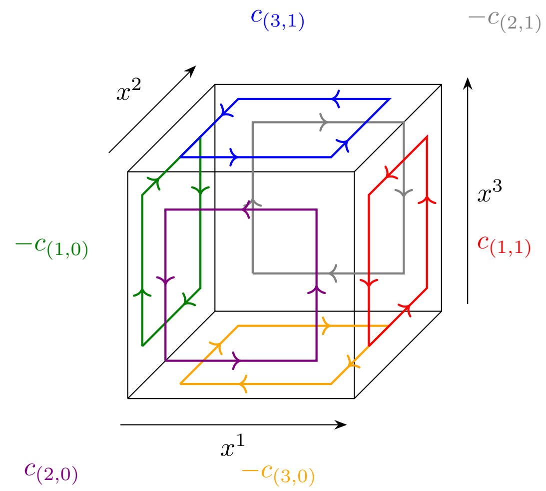


Figure 25.7: The double boundary of a 3-cube. Arrows indicate the orientation of the (possibly flipped) facets $\pm c_{(i,y)}$, i.e., they take into account the sign with which they contribute to the boundary.

Naturally one may ask whether the same holds also for simplices, which is what we consider next.

Theorem 25.5.2. *The double boundary of a singular simplex vanishes, $\partial^2 c = 0$ for all $c \in \mathcal{S}_k(M)$.*

Proof. By direct calculation, following the definitions given above, we have

$$\partial\partial c = \sum_{i=0}^{k}(-1)^{i+1}\partial c_i = \sum_{i=1}^{k}\sum_{j=1}^{k-1}(-1)^{i+j}c_{ij}\,. \tag{25.5.24}$$

To understand the appearing $(k-2)$-simplices c_{ij}, we can come back to their definition in terms of functions $F_{ij} : \Delta k - 2 \to \Delta^k$, which are defined by first replacing the i'th inequality in the condition (25.5.9) by an equality, and then replacing also the j'th of the remaining k inequalities by an equality. In total, this means that if $i \leq j$, we have replaced the inequalities originally labeled i and $j+1$ (since at the time we perform the second replacement, we have already replaced the i'th inequality, and so what has now become the j'th inequality was originally labeled as $(j+1)$'th inequality) by equalities, while if $i > j$, we have replaced the inequalities

i and j. In any case we replace two inequalities and relabel all remaining free coordinates in ascending order. Since it does not matter which of the two inequalities is replaced in the first step and which one in the second step, we can conclude that

$$F_{ij} = \begin{cases} F_{(j+1)i} & i \leq j\,, \\ F_{j(i-1)} & i > j\,, \end{cases} \tag{25.5.25}$$

and hence also

$$c_{ij} = \begin{cases} c_{(j+1)i} & i \leq j\,, \\ c_{j(i-1)} & i > j\,. \end{cases} \tag{25.5.26}$$

Applying this to the double boundary formula, we have

$$\begin{aligned}\partial\partial c &= \sum_{i=1}^{k}\sum_{j=1}^{k-1}(-1)^{i+j}c_{ij} \\ &= \sum_{i=2}^{k}\sum_{j=1}^{i-1}(-1)^{i+j}c_{ij} + \sum_{i=1}^{k-1}\sum_{j=i}^{k-1}(-1)^{i+j}c_{ij} \\ &= \sum_{i=2}^{k}\sum_{j=1}^{i-1}(-1)^{i+j}c_{j(i-1)} + \sum_{i=1}^{k-1}\sum_{j=i}^{k-1}(-1)^{i+j}c_{ij} \\ &= \sum_{j'=1}^{k-1}\sum_{i'=1}^{j'}(-1)^{j'+1+i'}c_{i'j'} + \sum_{i=1}^{k-1}\sum_{j=i}^{k-1}(-1)^{i+j}c_{ij} \\ &= -\sum_{i'=1}^{k-1}\sum_{j'=i'}^{k-1}(-1)^{i'+j'}c_{i'j'} + \sum_{i=1}^{k-1}\sum_{j=i}^{k-1}(-1)^{i+j}c_{ij} \\ &= 0\,. \quad \blacksquare\end{aligned} \tag{25.5.27}$$

Also this is easily illustrated by an example.

Example 25.5.4 (Double boundary of a 3-simplex). Let $c : \Delta^3 \to M$ be a singular 3-cube. Its boundary is the chain

$$\partial c = -c_0 + c_1 - c_2 + c_3\,, \tag{25.5.28}$$

where the facets $c_i : \Delta^2 \to M$ are given by

$$c_0(x,y) = c(0,x,y)\,, \qquad c_1(x,y) = c(x,x,y)\,, \tag{25.5.29a}$$

$$c_2(x,y) = c(x,y,y)\,, \qquad c_3(x,y) = c(x,y,1)\,. \tag{25.5.29b}$$

Now their boundaries are given by

$$\partial c_i = -c_{i0} + c_{i1} - c_{i2}\,. \tag{25.5.30}$$

Hence, we have

$$\begin{aligned}\partial\partial c = c_{00} - c_{01} + c_{02} - c_{10} + c_{11} - c_{12} \\ + c_{20} - c_{21} + c_{22} - c_{30} + c_{31} - c_{32}\,.\end{aligned} \tag{25.5.31}$$

Here the facets $c_{ij} : \Delta^1 \to M$ are given by the 1-simplices

$$c_{00}(x) = c(0,0,x) = c_{10}(x)\,, \qquad c_{01}(x) = c(0,x,x) = c_{20}(x)\,, \tag{25.5.32a}$$

$$c_{02}(x) = c(0,x,1) = c_{30}(x)\,, \qquad c_{11}(x) = c(x,x,x) = c_{21}(x)\,, \tag{25.5.32b}$$

$$c_{12}(x) = c(x,x,1) = c_{31}(x)\,, \qquad c_{22}(x) = c(x,1,1) = c_{32}(x)\,. \tag{25.5.32c}$$

Hence, we see that they appear in pairs which cancel each other, and so $\partial\partial c = 0$. This is illustrated in figure 25.8.

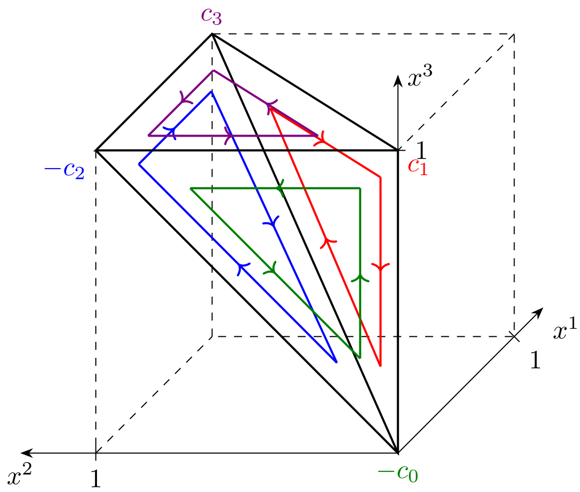


Figure 25.8: The double boundary of a 3-simplex. Arrows indicate the orientation of the (possibly flipped) facets $\pm c_i$, i.e., they take into account the sign with which they contribute to the boundary.

Finally, we can come back to our original motivation and apply our findings to the case of chains. The following is now straightforward.

Theorem 25.5.3. *The double boundary of a chain vanishes, $\partial^2 C = 0$ for all $C \in \mathcal{C}_k(M)$.*

Proof. Since the boundary of a chain is given by the boundary of its constituents, we have

$$\partial\partial C = \partial\partial \sum_{c \in \mathcal{P}_k(M)} C_c c = \sum_{c \in \mathcal{P}_k(M)} C_c \partial\partial c = 0\,, \tag{25.5.33}$$

where $\partial\partial c = 0$ as shown in the proofs of theorems 25.5.1 and 25.5.2. ∎

Intuitively, it means that a boundary must always be "closed" in the sense that it does by itself not have a boundary. Otherwise it could not enclose its interior.

25.6 Integrals over manifolds

In the previous sections we have now learned how to integrate k-forms over regions of manifolds parametrized be k-chains. We finally come to the point of integrating over all of a manifold. In order to do this, we must decompose the integral into pieces which we can map into $\mathbb{R}^k$. This decomposition can be done using a partition of unity, as discussed in section 1.5. Given a partition and unity and an atlas which are chosen to fit together in a suitable way, we can define integration on manifolds as follows.

Definition 25.6.1 (Integration over manifolds)**.** Let M be an orientable manifold of dimension n together with an oriented atlas $\mathcal{A}$ and a partition of unity R, such that for every $\rho \in R$ the support $\operatorname{supp}\rho$ is compact and there exists a chart (U_ρ, ϕ_ρ) such that $\operatorname{supp}\rho \subset U_\rho$. Let $B_\rho \subset \mathbb{R}^n$ be a box such that $\phi_\rho(\operatorname{supp}\rho) \subset B_\rho$. For a compactly supported n-form $\omega \in \Omega^n(M)$ the *integral* over M is defined as

$$\int_M \omega = \sum_{\rho \in R} \int_{B_\rho} (\phi_\rho^{-1})^*(\rho\omega)\,. \tag{25.6.1}$$

There are a few remarks regarding this definition. First of all, note that a partition of unity of this type does not exist on every manifold (using our definition of manifolds). One needs additional conditions (such as metrizability or paracompactness) to guarantee the existence of such a partition of unity. We will not go into these details, because all examples of manifolds we consider in this lecture course have these properties and suitable partitions of unity.

We further remark that although R will in general be an infinite set, the sum in the definition above is actually finite and thus well-defined. The reason is that the compact support of ω intersects only the supports of a finite number of $\rho \in R$, so that $\rho \cdot \omega$ will be non-zero only for these finitely many ρ.

Of course the important question arises how the value of the integral depends on the choice of the partition of unity and the atlas. This is answered by the following theorem.

Theorem 25.6.1. *Let M be a manifold and $(R, \mathcal{A})$ and $(R', \mathcal{A}')$ two choices of a partition of unity and an oriented atlas as given in the definition above, such that $\mathcal{A} \cup \mathcal{A}'$ is oriented. Then the integrals defined by $(R, \mathcal{A})$ and $(R', \mathcal{A}')$ are the same.*

Proof. ▶...◀ ■

The condition that the union of both atlases is oriented is important. It means that both atlases define the same orientation on M. If they define opposite orientation, their integrals will differ by a minus sign.

25.7 Stokes' theorem

The boundary operator $\partial : \mathcal{C}_k(M) \to \mathcal{C}_{k-1}(M)$ is in some sense similar to the exterior derivative $\mathrm{d} : \Omega^k(M) \to \Omega^{k+1}(M)$. Both satisfy $\partial^2 = 0$ resp. $\mathrm{d}^2 = 0$. This already suggests that both operators are related to each other. Indeed there exists a close relationship, which is given by the following famous theorem.

Theorem 25.7.1 (Stokes' theorem)**.** *The integral of a $(k-1)$-form ω over the boundary ∂C of a k-chain C equals the integral of its exterior derivative $\mathrm{d}\omega$ over the k-chain C,*

$$\int_{\partial C} \omega = \int_C \mathrm{d}\omega \,. \tag{25.7.1}$$

Proof. ▶...◀ ■

25.8 Integration by parts

From Stokes' theorem 25.7.1 and the properties of differential forms a number of helpful formulas can be derived, which may be summarized under the name integration by parts. The most basic of these formulas is the following.

Theorem 25.8.1. *Let M be a manifold of dimension n, $\omega \in \Omega^p(M)$ and $\sigma \in \Omega^q(M)$ differential forms such that $k = p + q + 1 \leq n$ and $C \in \mathcal{C}_k(M)$ a k-chain. Then the following holds:*

$$\int_C \mathrm{d}\omega \wedge \sigma = (-1)^{p+1} \int_C \omega \wedge \mathrm{d}\sigma + \int_{\partial C} \omega \wedge \sigma \,. \tag{25.8.1}$$

Proof. Recall the formula

$$\mathrm{d}(\omega \wedge \sigma) = \mathrm{d}\omega \wedge \sigma + (-1)^p \omega \wedge \mathrm{d}\sigma\,. \tag{25.8.2}$$

Together with Stokes' theorem follows

$$\begin{aligned}\int_{\partial C} \omega \wedge \sigma &= \int_C \mathrm{d}(\omega \wedge \sigma)\\ &= \int_C \mathrm{d}\omega \wedge \sigma + (-1)^p \int_C \omega \wedge \mathrm{d}\sigma\,,\end{aligned} \tag{25.8.3}$$

from which the statement of the theorem can be obtained. ■

Another helpful formula can be derived for Lie derivatives, in case the differential form to be integrated has maximal rank.

Theorem 25.8.2. *Let M be a manifold of dimension n, $\omega \in \Omega^p(M)$ and $\sigma \in \Omega^q(M)$ differential forms such that $p + q = n$, $X \in \mathrm{Vect}(M)$ a vector field and $C \in \mathcal{C}_n(M)$ a n-chain. Then the following holds:*

$$\int_C \mathcal{L}_X\omega \wedge \sigma = \int_{\partial C} \iota_X(\omega \wedge \sigma) - \int_C \omega \wedge \mathcal{L}_X\sigma\,. \tag{25.8.4}$$

Proof. We first use the Leibniz rule (16.5.7) for Lie derivatives and then Cartan's formula (16.5.1), from which we find

$$\mathcal{L}_X\omega \wedge \sigma + \omega \wedge \mathcal{L}_X\sigma = \mathcal{L}_X(\omega \wedge \sigma) = \mathrm{d}\iota_X(\omega \wedge \sigma) + \iota_X\mathrm{d}(\omega \wedge \sigma)\,. \tag{25.8.5}$$

The second term vanishes, since $\omega \wedge \sigma$ is already a form of maximal rank $p + q = n$, and thus has vanishing exterior derivative. The first term is an exact form, and so we can use Stokes' theorem to find

$$\begin{aligned}\int_{\partial C} \iota_X(\omega \wedge \sigma) &= \int_C \mathrm{d}\iota_X(\omega \wedge \sigma)\\ &= \int_C \mathcal{L}_X(\omega \wedge \sigma)\\ &= \int_C \mathcal{L}_X\omega \wedge \sigma + \int_C \omega \wedge \mathcal{L}_X\sigma\,,\end{aligned} \tag{25.8.6}$$

from which the statement of the theorem follows. ■

25.9 Dirac distributions

In the physics literature one often encounters a situation in which a quantity which is confined to a submanifold of lower dimension is supposed to be described by a density on the surrounding manifold. The most common example is that of a point charge of total charge Q located at a point $p \in M$ of some space M, which is described by a charge density ρ on M such that ρ vanishes everywhere except in p, and its integral over M equals Q. This density is then formally written in a form similar to $\rho = Q\delta_p$, where δ_p is often simply called a Dirac delta function centered at p, and defined to vanish everywhere except in p and have integral equal to one when integrated over a subvolume of M which contains p, and zero otherwise. Equivalently, one can demand that

$$\int_M f\delta_p = f(p) \tag{25.9.1}$$

for all functions $f \in C^\infty(M, \mathbb{R})$. More generally, one considers charges confined to a wire or a surface, with a given charge density per line element or area element, and again expresses these in terms of a volume density, with the help of an appropriate Dirac delta. Usually these are expressed in a given set of coordinates, and integration is performed using the known rules for Dirac delta functions on $\mathbb{R}$ and $\mathbb{R}^n$. In the context of differential geometry, one may of course ask whether there is a proper, coordinate-independent description of such Dirac delta expressions. Demanding that they can be integrated to yield a finite value indicates that they should be expressed as differential forms. However, this description can only be formal, since in order to yield a finite integral despite vanishing almost everywhere, they would have to be infinite somewhere. A suitable definition is the following.

Definition 25.9.1 (Dirac distribution)**.** Let M be a manifold of dimension $\dim M = m$ and S an embedded submanifold of dimension $\dim S = s$ with inclusion map $\iota : S \to M$. The *Dirac distribution* of S is the formal $(m-s)$-form δ_S such that

$$\int_M \omega \wedge \delta_S = \int_S \iota^* \omega \tag{25.9.2}$$

for all s-forms $\omega \in \Omega^s(M)$.

We see that the right hand side is well-defined since $\iota^*\omega \in \Omega^s(S)$ is a s-form on S. The left hand side has the form of an integral of a m-form; however, this definition is purely formal. To see how it is related to the usual definition in coordinates, we first study the most simple example of a zero-dimensional submanifold S.

Example 25.9.1. Let M be a manifold, $S = \{p\} \subset M$ the zero-dimensional submanifold containing a single point $p \in M$ and (U, ϕ) a chart such that $p \in U$. ▶. . . ◀

25.10 Integration along fibers

So far we have studied the case of integrating n-forms on n-dimensional domains, the result of which is a real number. Sometimes, however, one encounters the case that one would like to integrate "only along some directions", and be left with a function, or even a differential form of lower rank, which still depends "on the remaining directions". We will now make these statements more rigorous. We first provide a definition.

Definition 25.10.1 (Integration along fibers)**.** Let $\pi : E \to B$ be a fiber bundle with fiber dimension k. The *integration along the fibers* of E is defined as the unique linear prescription $\pi_* : \Omega^n(E) \to \Omega^{n-k}(B)$ such that for every $\sigma \in \Omega^{n-k}(B)$, $\tau \in \Omega^k(E)$ and $p \in B$ holds

$$\pi_*(\pi^*\sigma \wedge \tau)(p) = \sigma(p) \int_{\pi^{-1}(p)} \tau\,. \tag{25.10.1}$$

A few remarks are necessary. First note that in order for the integral in the definition to be defined, $\operatorname{supp} \tau \cap \pi^{-1}(x)$ must be compact. Differential forms on E which satisfy this condition are said to have *compact support in the vertical direction*. As a consequence, also the integration along fibers is defined only for forms which have compact support in the vertical direction, which we shall therefore assume without explicitly mentioning it. Further, it is sufficient to define the integration along fibers for differential forms of the form $\pi^*\sigma \wedge \tau$ as given above, since these span the space of all differential n-forms with $n \leq k$ on E, which can be seen, for example, by considering a local trivialization on $U \subset B$ and all σ which form a local coordinate basis of $\Omega^{n-k}(B)$, as well as all τ obtained from k-forms on the fiber F obtained via the local trivialization; we will show this explicitly using local induced coordinates below.

In the literature also another, equivalent definition of integration along fibers can be found, which is related to our definition as follows.

Theorem 25.10.1. *Let $\pi : E \to B$ be a fiber bundle with fiber dimension k. For every $\omega \in \Omega^n(E)$ with $n \leq k$, $\pi_*\omega \in \Omega^{n-k}(B)$ is the unique $(n-k)$-form such that for all vector fields $Y_1, \ldots, Y_{n-k} \in \mathrm{Vect}(E)$ and $X_1, \ldots, X_{n-k} \in \mathrm{Vect}(B)$ satisfying*

$$X_i \circ \pi = \pi_* \circ Y_i \tag{25.10.2}$$

and $p \in B$ holds

$$\pi_*\omega(X_1, \ldots, X_{n-k})(p) = \int_{\pi^{-1}(p)} \omega(Y_1, \ldots, Y_{n-k}, \bullet)\,. \tag{25.10.3}$$

Proof. First note that a $(n-k)$-form on B is uniquely determined by its action on $n-k$ vector fields at each point. Hence, we only have to check that the action of the fiber integral given in definition 25.10.1 satisfies the equation given here. Further, we can restrict ourselves to differential forms of the form $\omega = \pi^*\sigma \wedge \tau$ with $\sigma \in \Omega^{n-k}(B)$ and $\tau \in \Omega^k(E)$, since the integration along fibers is linear by definition, the equation given here is linear in ω, and these forms span $\Omega^n(E)$. By (25.10.1) we have

$$\pi_*(\pi^*\sigma \wedge \tau)(X_1, \ldots, X_{n-k})(p) = \sigma(X_1, \ldots, X_{n-k})(p) \int_{\pi^{-1}(p)} \tau\,, \tag{25.10.4}$$

while (25.10.3) yields

$$\pi_*(\pi^*\sigma \wedge \tau)(X_1, \ldots, X_{n-k})(p) = \int_{\pi^{-1}(p)} (\pi^*\sigma \wedge \tau)(Y_1, \ldots, Y_{n-k}, \bullet)\,. \tag{25.10.5}$$

For the integral given here we find

$$\int_{\pi^{-1}(p)} (\pi^*\sigma \wedge \tau)(Y_1, \ldots, Y_{n-k}, \bullet) = \int_{\pi^{-1}(p)} (\pi^*\sigma)(Y_1, \ldots, Y_{n-k}) \wedge \tau\,, \tag{25.10.6}$$

since all terms which we have omitted under the integral are of the form

$$(\pi^*\sigma)(Y_{i_1}, \ldots, Y_{i_m}, \bullet) \wedge \tau(Y_{i_{m+1}}, \ldots, Y_{i_{n-k}}, \bullet) \tag{25.10.7}$$

with $m < n-k$; note that the first factor is horizontal, and so its pullback to the fiber π^{-p}, and thus the integral over the fiber, vanishes. For the first factor under the integral we have

$$(\pi^*\sigma)(Y_1, \ldots, Y_{n-k}) = \pi^*\sigma(X_1, \ldots, X_{n-k})\,, \tag{25.10.8}$$

since the vector fields Y_i project to X_i. Here $\sigma(X_1, \ldots, X_{n-k})$ is a zero-form, i.e., a function, and so its pullback along π is constant along the fibers, and we can pull it out of the integral, so that

$$\int_{\pi^{-1}(p)} (\pi^*\sigma)(Y_1, \ldots, Y_{n-k}) \wedge \tau = \sigma(X_1, \ldots, X_{n-k})(p) \int_{\pi^{-1}(p)} \tau\,, \tag{25.10.9}$$

which agrees with the form given by (25.10.1). ■

The fiber integral has a few helpful properties, which can be used in calculations, and which we show below. The first relation we show is that it commutes with the exterior derivative.

Theorem 25.10.2. *Integration along fibers commutes with exterior differentiation, i.e.,*

$$\pi_*(\mathrm{d}\omega) = \mathrm{d}(\pi_*\omega) \tag{25.10.10}$$

for all $\omega \in \Omega^n(E)$ with $n \geq k$.

Proof. ►. . . ◄ ■

Another helpful relation concerns the exterior product with the pullback of a differential form on the base manifold. This is also known as the projection formula.

Theorem 25.10.3. *For all $\omega \in \Omega^n(E)$ with $n \geq k$ and $\alpha \in \Omega^m(B)$ holds*

$$\pi_*(\pi^*\alpha \wedge \omega) = \alpha \wedge \pi_*\omega\,. \tag{25.10.11}$$

Proof. Again we consider only differential forms of the form $\omega = \pi^*\sigma \wedge \tau$, and conclude on the general case by linearity. Then we have

$$\begin{aligned}\pi_*(\pi^*\alpha \wedge \pi^*\sigma \wedge \tau)(p) &= \pi_*[\pi^*(\alpha \wedge \sigma) \wedge \tau](p)\\ &= (\alpha \wedge \sigma)(p)\int_{\pi^{-1}(p)} \tau\\ &= [\alpha \wedge \pi_*(\pi^*\sigma \wedge \tau)](p)\end{aligned} \tag{25.10.12}$$

for all $p \in B$. ■

If the forms under consideration are of maximal rank, we can apply the projection formula to obtain the following important result.

Theorem 25.10.4. *For all $\omega \in \Omega^n(E)$ with $n \geq k$ and $\sigma \in \Omega^{b-n+k}(B)$, where $b = \dim B$, holds*

$$\int_E (\pi^*\sigma \wedge \omega) = \int_M (\sigma \wedge \pi_*\omega)\,. \tag{25.10.13}$$

Proof. ►. . . ◄ ■

In the particular case $n = b + k = \dim E$ and $\sigma = 1$, this formula represents the intuitive idea that a differential form of maximal rank on the total space of a fiber bundle can be "integrated along the fibers first, and then over the base manifold".

Chapter 26

Connections

26.1 Horizontal distributions

One of the most important notions in differential geometry is that of a *connection* on a fiber bundle. The most general type we discuss here is called an *Ehresmann connection*. There are many different, but equivalent ways to define Ehresmann connections. We will start with a definition that is most intuitive and pictorial (although the least practical for calculations).

Definition 26.1.1 (Horizontal distribution)**.** Let $\pi : E \to M$ be a fiber bundle. A *horizontal distribution* on E is an assignment $e \mapsto H_eE$ of a *horizontal tangent space* $H_eE \subset T_eE$ to every $e \in E$ such that $T_eE = V_eE \oplus H_eE$ and for every $e \in E$ there exists a neighborhood U_e on which the horizontal tangent spaces are spanned by $n = \dim M$ vector fields $(X_1, \ldots, X_n)$.

Recall that on a fiber bundle $\pi : E \to M$ we have for every $e \in E$ a canonically defined vertical tangent space $V_eE = \ker \pi_* \subset T_eE$, and that the vertical tangent spaces constitute the fibers of a vector bundle - the vertical tangent bundle VE. A horizontal distribution assigns to each $e \in E$ a complement of the vertical tangent space V_eE. The condition that these horizontal vector spaces are locally spanned by vector fields can be understood geometrically as a requirement that this assignment is smooth. This means in particular that the union of the horizontal vector spaces forms a manifold HE, which is the total space of a vector bundle over E, called the horizontal tangent bundle. Note that while the vertical tangent bundle VE is canonically defined over the total space of every fiber bundle, a horizontal bundle is *not* canonically given and thus defines an additional structure.

An important observation is that the splitting $T_eE = V_eE \oplus H_eE$ implies that every tangent vector $w \in T_eE$ is uniquely decomposed in the form $w = w_V + w_H$, where $w_V \in V_eE$ is the vertical part and $w_H \in H_eE$ is the horizontal part. We thus have uniquely defined projections $w \mapsto w_V$ and $w \mapsto w_H$, which extend to maps $TE \to VE$ and $TE \to HE$ on all of the tangent bundle TE. One easily checks that these maps are vector bundle homomorphisms, and that we can identify TE with the Whitney sum $VE \oplus HE$.

Note that definition 26.1.1 already implies that the dimension of the horizontal tangent spaces is given by $\dim H_eE = \dim M$. This can be seen from the fact that the dimension of the vertical tangent spaces is given by the fiber dimension, and so

$$\dim H_eE = \dim T_eE - \dim V_eE = \dim E - (\dim E - \dim M) = \dim M \,. \tag{26.1.1}$$

Recall that $\dim M$ is also the same as the dimension of the tangent spaces T_xM for $x \in M$.

This already hints towards the existence of a relation between these vector spaces. Indeed, such a relation exists, which we state as follows.

Theorem 26.1.1. *A connection on a fiber bundle $\pi : E \to M$ induces a vector space isomorphism between H_eE and $T_{\pi(e)}M$ for every $e \in E$.*

Proof. Consider the projector $w \mapsto w_H$ onto the horizontal distribution HE. It follows from the unique decomposition $w = w_V + w_H$ that the kernel of this projector is the vertical tangent bundle $VE = \ker \pi_*$. This means that for all $e \in E$ there is a vector space isomorphism

$$H_eE \cong T_eE/V_eE \cong T_{\pi(e)}M\,, \tag{26.1.2}$$

which relates the elements

$$\begin{aligned} H_eE \ni w_H \sim [w_H] &= \{w' \in T_eE, w' - w_H \in V_eE\} \in T_eE/V_eE \\ &= \{w' \in T_eE, \pi_*(w') = \pi_*(w_H)\} \in T_eE/V_eE \\ &\sim \pi_*(w_H) \in T_{\pi(e)}M \\ &= \pi_*(w) \in T_{\pi(e)}M\,, \end{aligned} \tag{26.1.3}$$

To see that the restriction of π_* to H_eE is indeed a vector space isomorphism from H_eE to $T_{\pi(e)}M$, note that for $w, w' \in T_eE$ we have $\pi_*(w) = \pi_*(w')$ if and only if $w - w'$ is vertical, and hence $w_H = w'_H$, so that $\pi_*(w)$ is uniquely determined by the horizontal part w_H and vice versa. ■

To further illustrate these concepts, we introduce coordinates (x^μ) on a trivializing neighborhood $U \subset M$ and (y^a) on the fiber space of the bundle $\pi : E \to M$, so that we have coordinates (x^μ, y^a) on E and the projection π simply discards the second part of these coordinates. We can then write a tangent vector $w \in T_eE$ in the form

$$w = u^\mu \frac{\partial}{\partial x^\mu} + v^a \frac{\partial}{\partial y^a} = u^\mu \partial_\mu + v^a \bar{\partial}_a\,, \tag{26.1.4}$$

from which we obtain coordinates (u^μ, v^a) on T_eE, (v^a) on V_eE and thus (x^μ, y^a, u^μ, v^a) on TE and (x^μ, y^a, v^a) on VE.

To understand the horizontal tangent bundle in these coordinates, first note that the differential π_* of the bundle projection π takes the form

$$\pi_*\left(u^\mu \partial_\mu + v^a \bar{\partial}_a\right) = u^\mu \partial_\mu\,, \tag{26.1.5}$$

where ∂_μ on the left hand side denotes one part of a coordinate basis of T_eE, while on the right hand side it denotes a coordinate basis of $T_{\pi(e)}M$. Using theorem 26.1.1, we thus see that the horizontal part w_H is fully determined by the components u^μ. We may therefore write the horizontal projector in coordinates in the form

$$w_H = \left(u^\mu \partial_\mu + v^a \bar{\partial}_a\right)_H = u^\mu \delta_\mu\,, \tag{26.1.6}$$

where δ_μ is the basis of H_eE which satisfies $\pi_*(\delta_\mu) = \partial_\mu$. Using the fact that $T_eE = V_eE \oplus H_eE$, we realize that

$$(\delta_\mu, \bar{\partial}_a) \tag{26.1.7}$$

is a basis of T_eE, which we call the *adapted basis* induced by the connection. We will construct explicit coordinate expressions of this adapted basis and its relation to the coordinate basis, and hence a coordinate representation of the connection, in the following sections.

Recall from section 19.3 that the (canonically defined) vertical tangent bundle VE gives rise to a horizontal cotangent bundle H^*E. If we have a horizontal distribution HE, then we can analogously also define a vertical cotangent bundle. We will do so as follows.

Definition 26.1.2 (Vertical cotangent vector)**.** Let $\pi : E \to M$ be a fiber bundle, $HE \subset TE$ a horizontal distribution and $e \in E$. A covector $\alpha \in T_e^*E$ is called *vertical* if and only if $\langle v, \alpha \rangle = 0$ for all horizontal vectors $v \in H_eE$. The space V_e^*E of all vertical covectors at e is called the *vertical cotangent space* over e.

▶Discuss split of cotangent bundle◀

26.2 Connection forms

More practical definitions of a connection can be obtained by constructing certain maps, which can be derived from the unique decomposition $w = w_V + w_H$ for every tangent vector $w \in T_eE$, where $w_V \in V_eE$ is the vertical part and $w_H \in H_eE$ is the horizontal part. In the following it makes sense to focus on the vertical projector, since its codomain VE is canonically defined. We can define such a projector as follows.

Definition 26.2.1 (Connection form)**.** Let $\pi : E \to M$ be a fiber bundle. A *connection form* on E is a vector bundle homomorphism $\theta : TE \to VE$ covering the identity map id_E on E and restricting to the identity map on VE, i.e., $\theta|_{VE} = \mathrm{id}_{VE}$.

This definition requires a few explanations. First, note that following theorem 4.6.2 one may equivalently view θ as a section of the homomorphism bundle $\mathrm{Hom}(TE, VE)$, which according to theorem 4.6.1 is identified with the bundle $T^*E \otimes VE$; the latter justifies the name "connection form", as it can be seen as a one-form taking values in the vertical tangent bundle.

Since $\theta : TE \to VE$ covers the identity, we have $\pi \circ \theta = \pi$. Further, θ is a vector bundle homomorphism, which means that each restriction $\theta|_e : T_eE \to V_eE$ is linear. Finally, it is in fact a projection onto VE, since $\theta(w) \in VE$ for all $w \in TE$ and θ restricts to the identity on VE, so that $\theta \circ \theta = \theta$. With these properties, the following is straightforward.

Theorem 26.2.1. *For every fiber bundle $\pi : E \to M$ there is a one-to-one correspondence between horizontal distributions and connection forms on E.*

Proof. Given a connection form θ on E, the kernel of θ is a horizontal distribution. Conversely, given a horizontal distribution, we can uniquely split every vector $w \in TE$ in the form $w = w_V + w_H$, where $w_V \in VE$ is vertical and $w_H \in HE$ is horizontal. Then $\theta : w \mapsto w_V$ defines a connection form. ■

Recall the coordinates (x^μ, y^a, u^μ, v^a) we introduced on TE in the previous section. In these coordinates a vector bundle homomorphism $\theta : TE \to TE$ must be linear in u, v, and can be expressed in the form

$$\theta(x, y, u, v) = (u^\nu \theta^\mu_\nu(x, y) + v^a \theta^\mu_a(x, y))\, \partial_\mu + \left(u^\mu \theta^a_\mu(x, y) + v^b \theta^a_b(x, y)\right) \bar{\partial}_a \in T_{(x,y)}E\,, \quad (26.2.1)$$

so that it is determined by the component functions $\theta^\mu_\nu, \theta^a_\mu, \theta^\mu_a, \theta^a_b$. Demanding that the image is vertical implies $\theta^\mu_\nu \equiv 0$ and $\theta^\mu_a \equiv 0$, so that we have

$$\theta(x, y, u, v) = \left(u^\mu \theta^a_\mu(x, y) + v^b \theta^a_b(x, y)\right) \bar{\partial}_a \in V_{(x,y)}E\,. \quad (26.2.2)$$

Further demanding that θ is the identity on vertical vectors implies $\theta^a_b = \delta^a_b$, so that for a connection form we find the coordinate expression

$$\theta(x, y, u, v) = \left(u^\mu \theta^a_\mu(x, y) + v^a\right) \bar{\partial}_a \in V_{(x,y)}E\,. \quad (26.2.3)$$

A connection form is thus uniquely determined by the coordinate functions $\theta^a_\mu(x, y)$. Note that their index structure is reminiscent of a tensor field with mixed (covariant and contravariant) components. This is not a coincidence, as we shall see now.

Recall from theorem 4.6.2 that a vector bundle homomorphism covering the identity can also be regarded as a section of the corresponding homomorphism bundle, which in this case is $\mathrm{Hom}(TE, VE)$, and which is isomorphic to $VE \otimes T^*E$ following theorem 4.6.1. Adopting this point of view, we may write the connection form θ as the tensor field

$$\theta = \bar{\partial}_a \otimes (\theta^a_\mu \mathrm{d}x^\mu + \mathrm{d}y^a)\,, \tag{26.2.4}$$

whose $\mathrm{d}y^a$-part follows from the fact that θ acts as the identity on vertical vectors, and where we now omitted the argument (x, y), as we express the tensor field, not its value at a single point. Such a tensor field can be regarded as a one-form which takes values in the vertical tangent bundle - hence the name *connection form*.

A word of warning should be issued here. Naively one might think of omitting the $\mathrm{d}y^a$-part, as it does not depend on θ, and only consider a form $\theta^a_\mu \bar{\partial}_a \otimes \mathrm{d}x^\mu$. This implies that the coordinate expression θ^a_μ *looks like* coordinates for a map $(u^\mu \partial_\mu \mapsto u^\mu \theta^a_\mu \bar{\partial}_a) \in \mathrm{Hom}(T_xM, V_eE)$, but this is misleading. The reason for this coordinate expression is simply that while introducing coordinates on E we have fixed a local trivialization. If we choose a different trivialization, these components will not transform as components of a vector space, but as components of an *affine* space, which is what they really are. To see this, consider two connection forms θ, θ'. Given any vertical tangent vector $v \in VE$, their difference satisfies

$$(\theta' - \theta)(v) = \theta'(v) - \theta(v) = v - v = 0\,. \tag{26.2.5}$$

In coordinates, we have

$$\theta' - \theta = \bar{\partial}_a \otimes (\theta'^a_\mu \mathrm{d}x^\mu + \mathrm{d}y^a) - \bar{\partial}_a \otimes (\theta^a_\mu \mathrm{d}x^\mu + \mathrm{d}y^a) = (\theta'^a_\mu - \theta^a_\mu)\bar{\partial}_a \otimes \mathrm{d}x^\mu\,, \tag{26.2.6}$$

Hence, we find that $\theta' - \theta$ is a section of $VE \otimes H^*E$, which is a vector bundle. This suggests that we can consider connections as sections of an affine bundle modeled over this bundle. We will find that this is indeed the case in section 26.3, where we explicitly study this bundle.

To further illustrate these constructions, it is helpful to introduce a new basis $(\mathrm{d}x^\mu, \delta y^a)$ of the cotangent bundle T^*E, where $\mathrm{d}x^\mu$ is a basis of the horizontal cotangent bundle H^*E and

$$\delta y^a = \mathrm{d}y^a + \theta^a_\mu \mathrm{d}x^\mu \quad \Leftrightarrow \quad \mathrm{d}y^a = \delta y^a - \theta^a_\mu \mathrm{d}x^\mu\,. \tag{26.2.7}$$

In this basis the connection form simply reads

$$\theta = \bar{\partial}_a \otimes \delta y^a\,. \tag{26.2.8}$$

The dual basis $(\delta_\mu, \bar{\partial}_a)$ of the tangent bundle TE, which is defined such that

$$\langle \delta_\mu, \mathrm{d}x^\nu \rangle = \delta^\nu_\mu\,, \quad \langle \delta_\mu, \delta y^a \rangle = 0\,, \quad \langle \bar{\partial}_a, \mathrm{d}x^\nu \rangle = 0\,, \quad \langle \bar{\partial}_a, \delta y^b \rangle = \delta^b_a\,, \tag{26.2.9}$$

is given by

$$\delta_\mu = \partial_\mu - \theta^a_\mu \bar{\partial}_a \quad \Leftrightarrow \quad \partial_\mu = \delta_\mu + \theta^a_\mu \bar{\partial}_a\,. \tag{26.2.10}$$

It immediately follows that the vector fields δ_μ lie in the kernel of θ, and so they span the horizontal tangent bundle HE. Hence, this basis respects the split $TE = VE \oplus HE$ of the tangent bundle imposed by the connection. This becomes even more clear if we understand θ as a vertical projector. Consequently, we can also define a horizontal projector

$$\begin{aligned}\delta - \theta &= \partial_\mu \otimes \mathrm{d}x^\mu + \bar{\partial}_a \otimes \mathrm{d}y^a - \bar{\partial}_a \otimes \delta y^a \\ &= \partial_\mu \otimes \mathrm{d}x^\mu - \theta^a_\mu \bar{\partial}_a \otimes \mathrm{d}x^\mu \\ &= \delta_\mu \otimes \mathrm{d}x^\mu\,,\end{aligned} \tag{26.2.11}$$

where

$$\delta = \partial_\mu \otimes \mathrm{d}x^\mu + \bar{\partial}_a \otimes \mathrm{d}y^a = \delta_\mu \otimes \mathrm{d}x^\mu + \bar{\partial}_a \otimes \delta y^a \tag{26.2.12}$$

denotes the unit section 5.5.1, which can be expressed either in the coordinate basis, or in the basis adapted to the connection. Finally, it follows that in this basis the isomorphism between H_eE and $T_{\pi(e)}M$ simply relates $u^\mu\delta_\mu \in H_eE$ and $u^\mu\partial_\mu \in T_{\pi(e)}M$. This shows that $(\delta_\mu, \bar{\partial}_a)$ is simply the adapted basis (26.1.7) we have introduced before.

26.3 Ehresmann connections and jet bundle sections

As we have seen before, there are different possibilities how to describe an Ehresmann connection on a fiber bundle. We have also seen that connections form an affine space and not a vector space. It thus seems helpful to introduce another description and to model them as sections of an *affine* bundle. We have already encountered affine bundles when we discussed jet bundles. We will thus use the following geometric object as our main definition and - since it apparently lacks a particular name in the literature - simply call it an Ehresmann connection.

Definition 26.3.1 (Ehresmann connection)**.** Let $\pi : E \to M$ be a fiber bundle. An *Ehresmann connection* is a section $\omega : E \to J^1(E)$ of the jet bundle $\pi_{1,0} : J^1(E) \to E$.

To better understand the geometric meaning of this definition, recall that an element of $J^1(E)$ is an equivalence class of local sections around a point $x \in M$, i.e., maps $\sigma : U_\sigma \to E$ with $x \in U_\sigma$ for an open subset $U_\sigma \subset M$, where two local sections σ, τ are considered equivalent if for all curves $\gamma \in C^\infty(\mathbb{R}, U_\sigma \cap U_\tau)$ with $\gamma(0) = x$ and all functions $f \in C^\infty(E, \mathbb{R})$ holds

$$(f \circ \sigma \circ \gamma)(0) = (f \circ \tau \circ \gamma)(0) \quad \text{and} \quad (f \circ \sigma \circ \gamma)'(0) = (f \circ \tau \circ \gamma)'(0)\,. \tag{26.3.1}$$

The first condition simply translates to $\sigma(x) = \tau(x)$, while the second condition can be written as $\sigma_*(u) = \tau_*(u)$ for all $u \in T_xM$. The equivalence class of σ, for which we introduced the notation $j^1_x\sigma$, is thus fully characterized by the following data:

- the point $\pi_1(j^1_x\sigma) = x \in M$,
- the image $\pi_{1,0}(j^1_x\sigma) = \sigma(x) \in E_x = \pi^{-1}(x)$,
- a linear map $\sigma_*|_x : T_xM \to T_{\sigma(x)}E$ such that $\pi_* \circ \sigma_*|_x = \mathrm{id}_{T_xM}$.

In the given case we are interested in sections ω of the bundle $\pi_{1,0} : J^1(E) \to E$. By definition of the section we have $\pi_{1,0} \circ \omega = \mathrm{id}_E$. For all $e \in E$ thus follows that $\omega(e) \in \pi_{1,0}^{-1}(e)$, so that $\omega(e)$ must be of the form $j^1_x\sigma$ with $\sigma(x) = e$ and $x = \pi(e)$. This requirement already uniquely fixes the first two items from the list above, so that in order to specify a section ω we only need to supply the last item. To see how much freedom we have for choosing this item, we consider two different jets $j^1_x\sigma, j^1_x\tau$ (where this time σ and τ should be sections of $\pi : E \to M$ which are *not* in the same equivalence class, i.e., define different jets). Since we are dealing with linear maps, we can take the difference

$$0 = (\pi_* \circ \sigma_*|_x) - (\pi_* \circ \tau_*|_x) = \pi_* \circ (\sigma_*|_x - \tau_*|_x)\,, \tag{26.3.2}$$

which shows that the image of $\sigma_*|_x - \tau_*|_x$ must be contained in the vertical tangent space V_eE. In other words, for any $u \in T_xM$, the image $\sigma_*|_x(u) - \tau_*|_x(u)$ is vertical. Note, however, that $\sigma_*|_x(u)$ is *not* vertical, so that it is *not* sufficient to specify a vertical vector only. The reason for this is that the condition $\pi_* \circ \sigma_*|_x = \mathrm{id}_{T_xM}$ specifies an *affine* space, i.e., the difference of any two such linear maps $\sigma_*|_x, \tau_*|_x$ lies in the vector space $\mathrm{Hom}(T_xM, V_eE)$, but the space of linear maps satisfying this condition is not a vector space. In fact, we have shown this in theorem 21.6.4, where now we have a special case: the jet bundle $\pi_{1,0} : J^1(E) \to E$ is an affine bundle modeled over $VE \otimes \pi^*T^*M$. The fibers of this bundle are, of course, just $V_eE \otimes T^*_xM \cong \mathrm{Hom}(T_xM, V_eE)$ with $x = \pi(e)$.

It further follows from these considerations that the linear map $\sigma_*|_x$ must have maximal rank, i.e., it must be injective, and so it establishes a vector space isomorphism between T_xM and its image in $T_{\sigma(x)}E$. One may already expect that this image is the horizontal tangent space of a horizontal distribution, or equivalently the kernel of a connection form on E. The following theorem shows that this intuition is indeed correct.

Theorem 26.3.1. *For every fiber bundle $\pi : E \to M$ there is a one-to-one correspondence between Ehresmann connections and connection forms on E.*

Proof. We have seen that an Ehresmann connection assigns to each $e \in E$ with $\pi(e) = x$ a jet $j^1_x\sigma_e$ with $\pi_1(j^1_x\sigma_e) = x$ and $\pi_{1,0}(j^1_x\sigma_e) = \sigma_e(x) = e$ such that $\pi_* \circ \sigma_{e*}|_x = \mathrm{id}_{T_xM}$, and that the latter is the only ingredient that differs between different Ehresmann connections. Given this jet we can define for each $e \in E$ a linear function

$$\begin{array}{cccc} \theta_e & : & T_eE & \to & V_eE \\ & & w & \mapsto & w - \sigma_{e*}(\pi_*(w)) \end{array} . \tag{26.3.3}$$

It is clear that $\theta_e(w) \in V_eE$, since

$$\pi_*(\theta_e(w)) = \pi_*(w) - \pi_*(\sigma_{e*}(\pi_*(w))) = \pi_*(w) - \pi_*(w) = 0 . \tag{26.3.4}$$

Further, for $w \in V_eE$, we have $\pi_*(w) = 0$, and thus $\theta_e(w) = w$. Together with the linearity it follows that θ_e is a projection onto V_eE. Finally, since we have such a function θ_e for all $e \in E$, they constitute a map $\theta : TE \to VE$ which covers the identity on E. One easily checks that θ is a connection form.

Conversely, let θ be a connection form on E, i.e., a vector bundle homomorphism $\theta : TE \to VE$ covering the identity map id_E on E and restricting to the identity map on VE. For each $e \in E$ it thus defines a projection $\theta|_e : T_eE \to V_eE$. Let σ_e be a local section of the bundle $\pi : E \to M$ around $x = \pi(e)$ such that $\sigma_e(x) = e$ and $\theta(\sigma_{e*}(u)) = 0$ for all $u \in T_xM$. The latter condition means that $\sigma_{e*}(u)$ lies in the kernel of the projection $\theta|_e$. This completely fixes σ_{e*}, since for every local section σ_e we also have $\pi_* \circ \sigma_{e*} = \mathrm{id}_{T_xM}$. The set of all such local sections σ_e is thus simply the jet $j^1_x\sigma_e$. The jets for each $e \in E$ finally define a section $\omega : E \to J^1(E), e \mapsto j^1_x\sigma_e$ of the jet bundle, and thus an Ehresmann connection. ■

It follows from the construction above that

$$\sigma_{e*}(\pi_*(w)) = w - \theta_e(w) = w - w_V = w_H \in H_eE \tag{26.3.5}$$

is simply the horizontal part of w. By comparing with theorem 26.1.1 we thus see that the maps $\sigma_{e*} : T_xM \to H_eE$ and $\pi_* : H_eE \to T_xM$ simply realize the isomorphism between these two vector spaces, which is induced by the connection.

To further illustrate this construction, we introduce coordinates (x^μ) on a trivializing neighborhood $U \subset M$ and (y^a) on the fiber space of the bundle $\pi : E \to M$, so that we have coordinates (x^μ, y^a) on E and the projection π simply discards the second part of these coordinates. We can denote the coordinates on the first jet space by (x^μ, y^a, y^a_μ). In these coordinates a section of the bundle $\pi_{1,0} : J^1(E) \to E$ is thus expressed by an assignment

$$\omega : (x^\mu, y^a) \mapsto (x^\mu, y^a, y^a_\mu(x, y)) . \tag{26.3.6}$$

Now let $\sigma_e : U \to E$ with $x = \pi(e) \in U \subset M$ be a local section such that $\sigma_e(x) = e$ and $j^1_x\sigma_e = \omega(e)$. Writing this section as an assignment $(x^\mu) \mapsto (x^\mu, y^a(x))$, we see that the latter condition on its first jet implies that

$$\partial_\mu y^a(x) = y^a_\mu(x, y(x)) . \tag{26.3.7}$$

The pushforward $\sigma_{e*}(u)$ of a vector $u = u^\mu \partial_\mu \in T_x M$ is thus given by

$$\sigma_{e*}(u) = u^\mu(\partial_\mu x^\nu \partial_\nu + \partial_\mu y^a \bar{\partial}_a) = u^\mu(\partial_\mu + y^a_\mu(x,y)\bar{\partial}_a)\,. \tag{26.3.8}$$

For $w = u^\mu \partial_\mu + v^a \bar{\partial}_a \in T_e E$ we then use the definition (26.3.3) in theorem 26.3.1 to obtain the connection form θ_e as

$$\begin{aligned} \theta_e(w) &= w - \sigma_{e*}(\pi_*(w)) \\ &= u^\mu \partial_\mu + v^a \bar{\partial}_a - u^\mu(\partial_\mu + y^a_\mu(x,y)\bar{\partial}_a) \\ &= (v^a - u^\mu y^a_\mu(x,y))\bar{\partial}_a\,. \end{aligned} \tag{26.3.9}$$

By comparing this result with the expression (26.2.3) it thus follows that the coordinate expression for the connection form θ is simply given by $\theta^a_\mu(x,y) = -y^a_\mu(x,y)$.

Another possibility to see this relation between the coordinate expressions of the Ehresmann connection ω and the connection form θ is by using the induced relation (26.3.5) between $T_x M$ and $H_e E$. Recall that we may use the coordinate expression of θ to construct a basis (26.2.10) of $T_e E$ which respects the split into horizontal and vertical parts, where δ_μ are the horizontal basis vectors. It follows from their definition that a horizontal vector $w = u^\mu \delta_\mu$ is projected to $\pi_*(w) = u^\mu \partial_\mu = u \in T_x M$. The inverse map must therefore satisfy

$$\sigma_{e*}(u) = u^\mu \delta_\mu = u^\mu(\partial_\mu - \theta^a_\mu(x,y)\bar{\partial}_a)\,. \tag{26.3.10}$$

Comparing this to the coordinate expression for the pushforward of a section shows again that $y^a_\mu(x,y) = \partial_\mu y^a(x) = -\theta^a_\mu(x,y)$.

In an intuitive, geometric picture we can thus understand an Ehresmann connection as an assignment that takes a point e in E with coordinates (x^μ, y^a), which could come from evaluating a section $\sigma_e : M \to E$ at the point with coordinates (x^μ), and it assigns to it a jet, which determines "partial derivatives of the coordinates y^a with respect to the coordinates x^μ", such that the resulting tangent vectors at e to the graph of the section σ_e in E are horizontal.

26.4 Horizontal lift map

We have seen in section 26.1 that a connection on a fiber bundle $\pi : E \to M$, represented by the choice of horizontal tangent spaces $H_e E$, induces a vector space isomorphism $H_e E \cong T_{\pi(e)} M$ for every $e \in E$. We now make use of these vector space isomorphisms in order to lift geometric objects from M to E. For this purpose, it is useful to form them into a map, whose existence is guaranteed by the following statement.

Theorem 26.4.1. *Let $\pi : E \to M$ be a fiber bundle equipped with a horizontal distribution HE and $\chi : TE \to E$ the tangent bundle projection. The map*

$$\begin{array}{cccc} (\chi, \pi_*) & : & HE & \to & \pi^*TM \\ & & w & \mapsto & (\chi(w), \pi_*(w)) \end{array} \tag{26.4.1}$$

is a vector bundle isomorphism covering the identity on E.

Proof. First, it is clear that the map given above is a smooth map from HE to π^*TM. These two manifolds are constructed as total spaces of fiber bundles over E, with the projection given by the diagram

$$\begin{array}{ccc} HE & \xrightarrow{(\chi,\pi_*)} & \pi^*TM \\ & {\scriptstyle \chi} \searrow \quad \swarrow {\scriptstyle \mathrm{pr}_1} & \\ & E & \end{array}\,. \tag{26.4.2}$$

This diagram obviously commutes, so that (χ, π_*) is a bundle morphism. Given $e \in E$ and $w \in H_eE$, we have

$$(\chi(w), \pi_*(w)) = (e, \pi_*(w)) \tag{26.4.3}$$

with $\pi_*(w) \in T_{\pi(e)}M$. Since the vector space structure on the fibers of π^*TM is inherited from the fibers of TM, and π_* is a vector bundle homomorphism, hence linear on every fiber, also (χ, π^*) is linear on every fiber, and thus a vector bundle homomorphism. We then once again use the fact that we can identify the fibers of π^*TM and TM over points $e \in E$ and $p = \pi(e) \in M$ as

$$\pi_e^*TM \ni (e, v) \sim v \in T_{\pi(e)}M\,, \tag{26.4.4}$$

and that this identification is a vector space isomorphism. Now recall from theorem 26.1.1 that $\pi_*|_{H_eE} : H_eE \to T_{\pi(e)}M$ is a vector space isomorphism. Hence, also the composition

$$H_eE \xrightarrow{\pi_*} T_{\pi(e)}M \xrightarrow{(e,\bullet)} \pi_e^*TM \tag{26.4.5}$$

is a vector space isomorphism. Therefore, we indeed have a vector bundle isomorphism. ■

We remark that the map we considered above is, in fact, defined independently of the connection as a map

$$\begin{array}{cccc} (\chi, \pi_*) & : & TE & \to & \pi^*TM \\ & & w & \mapsto & (\chi(w), \pi_*(w)) \end{array} \tag{26.4.6}$$

on all of TE. However, this map is not an isomorphism, but only a homomorphism, whose kernel is the (also canonically defined) vertical tangent bundle VE. Only if we restrict this map to a horizontal bundle, which complements the vertical bundle, its restriction becomes an isomorphism. In the following, it will turn out to be useful to consider the inverse of this isomorphism, which we define as follows, essentially following [KSM93, sec. 9].

Definition 26.4.1 (Horizontal lift map)**.** Let $\pi : E \to M$ be a fiber bundle equipped with a horizontal distribution HE and $\chi : TE \to E$ the tangent bundle projection. The vector bundle isomorphism $\eta = (\chi, \pi_*)^{-1} : \pi^*TM \to HE$ is called the *horizontal lift map.*

It should be clear that by specifying η, we can reconstruct HE as its image. In fact, it turns out that horizontal lift maps provide yet another possibility to specify connections, which we have already encountered, though in a different form. We formalize this in the following statement.

Theorem 26.4.2. *Let $\pi : E \to M$ be a fiber bundle and $\chi : TE \to E$ the tangent bundle projection. Then there exists a one-to-one correspondence between Ehresmann connections $\omega : E \to J^1(E)$ and vector bundle homomorphisms $\eta : \pi^*TM \to TE$ covering the identity on E and satisfying $\pi_* \circ \eta = \mathrm{pr}_2$, where $\mathrm{pr}_2 : \pi^*TM \to TM$ is the projection onto the second factor of the fibered product $\pi^*TM \cong E \times_M TM$.*

Proof. For $e \in E$, let $\pi(e) = x$ and $\omega(e) = j_x^1\sigma$ for some representative $\sigma \in \Gamma_x(E)$. Recall that $\omega(e)$ is uniquely specified by the map $\sigma_* : T_xM \to T_eE$. One easily checks that the relation $\sigma_*(v) = \eta(e, v)$ establishes the claimed one-to-one correspondence, i.e., that η defined from ω with this prescription is a horizontal lift map, and that a horizontal lift map η conversely defines an Ehresmann connection ω. ■

We have now derived four equivalent ways to specify a connection: as a horizontal distribution, as a connection form, as a jet bundle section or as a horizontal lift map. These are summarized in figure 26.1.

We illustrate the definition of the horizontal lift map using coordinates (x^μ, y^a, u^μ), where (x^μ) are coordinates on M which specify the common base point of the points with coordinates

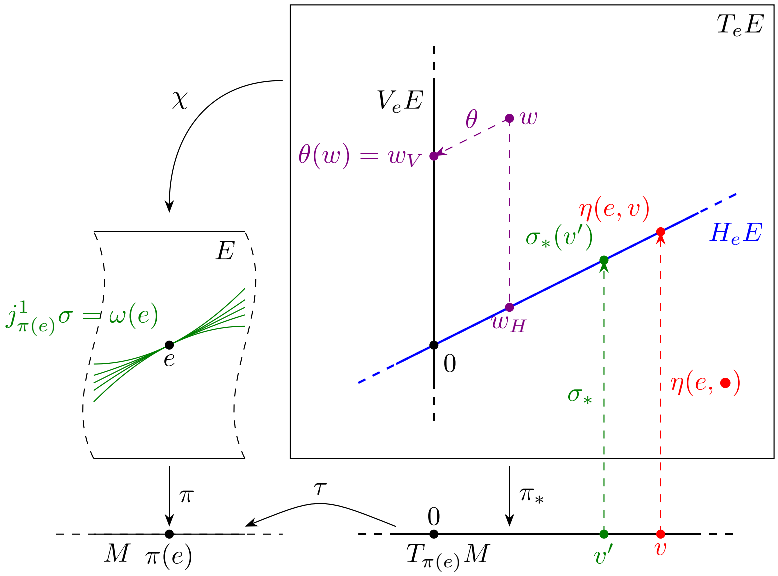


Figure 26.1: Four different, but equivalent ways to specify a connection.

(x^μ, y^a) on E and (x^μ, u^μ) on TM, where the latter are defined via the coordinate basis ∂_μ of TM. Recall that we can use the adapted basis (26.1.7) to write a horizontal tangent vector at $e \in E$ as $w_H = u^\mu \delta_\mu \in H_eE$, and so we can label the elements of HE by the same set of coordinates (x^μ, y^a, u^μ). This is justified, since we have already seen that $\pi_*(w_H) = u^\mu \partial_\mu \in T_{\pi(e)}M$. The horizontal lift map is the inverse of this map, and so it assigns to every $(e, v) \in \pi^*TM$, where $e \in E$ and $v = u^\mu \partial_\mu \in T_{\pi(e)}M$, the corresponding element $u^\mu \delta_\mu \in H_eE$. Here we see how the horizontal lift map constitutes an isomorphism of these vector bundles.

26.5 Frame bundle reduction

We now come to yet another description of connections, which once again turns out to be equivalent to the descriptions we have discussed before, and which helps to link the descriptions in terms of horizontal distributions and jet bundle sections. The starting point of our discussion is the tangent frame bundle $F(TE)$ over the total space of the bundle $\pi : E \to M$. If we denote the dimension of M by $\dim M = m$ and of E by $\dim E = m + n$, then TE is of rank $m + n$, and so $F(TE)$ is a principal $\mathrm{GL}(m+n, \mathbb{R})$-bundle. Now recall from section 19.2 that TE has a canonically defined subbundle, namely the vertical tangent bundle VE of rank n. This canonical structure allows us to define the following bundle.

Definition 26.5.1 (Adapted frame bundle)**.** Let $\pi : E \to M$ be a fiber bundle with $\dim M = m$ and $\dim E = m + n$. For $e \in E$, we call an *adapted frame* a frame $p : \mathbb{R}^{m+n} \cong \mathbb{R}^m \oplus \mathbb{R}^n \to T_eE$ if $p|_{\{0\}\oplus\mathbb{R}^n} : \{0\} \oplus \mathbb{R}^n \to T_eE$ is a frame of V_eE, i.e., p restricts to a bijective linear map onto V_eE in the last n components. The space of all adapted frames is called the *adapted frame bundle* $F_\pi(TE)$.

Similarly to the general linear frame bundle, one may expect that also the associated frame bundle is a principal bundle. This we show next.

Theorem 26.5.1. *The adapted frame bundle $F_\pi(TE)$ is a principal bundle with structure group*

$$G = \left\{ \begin{pmatrix} A & 0 \\ L & B \end{pmatrix}, A \in \mathrm{GL}(m,\mathbb{R}), B \in \mathrm{GL}(n,\mathbb{R}), L \in \mathrm{M}_{m,n}(\mathbb{R}) \right\}$$
$$\cong \mathrm{GL}(m,\mathbb{R}) \ltimes \mathrm{M}_{n,m}(\mathbb{R}) \rtimes \mathrm{GL}(n,\mathbb{R}) \subset \mathrm{GL}(m+n,\mathbb{R})\,. \quad (26.5.1)$$

Proof. We first show that G acts from the right on $F_\pi(TE)$ via $p \cdot g = p \circ g$. For this purpose, let $(v,w) \in \mathbb{R}^m \oplus \mathbb{R}^n$, $e \in E$, $p \in F_{\pi\,e}(TE)$ and $g = \left(\begin{smallmatrix} A & 0 \\ L & B \end{smallmatrix}\right) \in G$. Then we find another frame

$$p'(v,w) = (p \circ g)(v,w) = p(Av, Lv + Bw)\,. \quad (26.5.2)$$

In particular, for $v = 0$, we have $p'(0,w) = p(0,Bw) \in V_eE$, since p is an adapted frame. This mapping $\mathbb{R}^n \to V_eE, w \mapsto p(0,Bw)$ is bijective, since $B \in \mathrm{GL}(n,\mathbb{R})$, and so also p' is an adapted frame. To show that this action is free and transitive, let $p,p' \in F_{\pi\,e}(TE)$. Then there exists a unique $g \in \mathrm{GL}(m+n,\mathbb{R})$ such that $g = p^{-1} \circ p'$. If we apply this element to $(0,w)$, we have that $p'(0,w) \in V_eE$, and so $p^{-1}(p'(0,w))$ must again be of the form $(0,w')$. Hence, g must be of the form $\left(\begin{smallmatrix} A & 0 \\ L & B \end{smallmatrix}\right)$, and thus $g \in G$. ■

It is instructive to take a closer look at the group G which appears here. Clearly, this is a Lie group of dimension $m^2 + n^2 + mn$. Its group operation is given by matrix multiplication, and thus takes the form

$$\begin{pmatrix} A & 0 \\ L & B \end{pmatrix} \cdot \begin{pmatrix} A' & 0 \\ L' & B' \end{pmatrix} = \begin{pmatrix} AA' & 0 \\ LA' + BL' & BB' \end{pmatrix}, \quad (26.5.3)$$

and so the inverse is given by

$$\begin{pmatrix} A & 0 \\ L & B \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & 0 \\ -B^{-1}LA^{-1} & B^{-1} \end{pmatrix}. \quad (26.5.4)$$

If the bundle $\pi : E \to M$ is further equipped with a connection, and hence a horizontal distribution, the tangent spaces split in the form $T_eE = H_eE \oplus V_eE$. In this case one can further restrict the tangent bundle. We define this restriction as follows.

Definition 26.5.2 (Split frame bundle)**.** Let $\pi : E \to M$ be a fiber bundle with $\dim M = m$ and $\dim E = m+n$, and HE a horizontal distribution, corresponding to a connection ω. For $e \in E$, we call a *split frame* a frame $p : \mathbb{R}^{m+n} \cong \mathbb{R}^m \oplus \mathbb{R}^n \to T_eE$ if $p|_{\mathbb{R}^m \oplus \{0\}} : \mathbb{R}^m \oplus \{0\} \to T_eE$ is a frame of H_eE and $p|_{\{0\} \oplus \mathbb{R}^n} : \{0\} \oplus \mathbb{R}^n \to T_eE$ is a frame of V_eE, i.e., p restricts to a bijective linear map onto H_eE in the first m components and onto V_eE in the last n components. The space of all split frames is called the *split frame bundle* $F_\pi^\omega(TE)$.

Now the following is straightforward.

Theorem 26.5.2. *The split frame bundle $F_\pi^\omega(TE)$ is a principal bundle with structure group*

$$H = \left\{ \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}, A \in \mathrm{GL}(m,\mathbb{R}), B \in \mathrm{GL}(n,\mathbb{R}) \right\}$$
$$\cong \mathrm{GL}(m,\mathbb{R}) \times \mathrm{GL}(n,\mathbb{R}) \subset \mathrm{GL}(m+n,\mathbb{R})\,. \quad (26.5.5)$$

Proof. It is clear from the definition that a split frame at $e \in E$ can be seen as a pair of frames of H_eE and V_eE, and so $F_\pi^\omega(TE) \cong F(HE) \times_E F(VE)$. These bundles are principal bundles with structure groups $\mathrm{GL}(m,\mathbb{R})$ and $\mathrm{GL}(n,\mathbb{R})$, respectively. Hence, following theorem 20.1.5, $F_\pi^\omega(TE)$ is a principal bundle with structure group H given above. ■

Hence, we have seen that a connection allows us to reduce the adapted frame bundle to a split frame bundle. This suggests that a connection induces a reduction of the structure group. In order to show that this is indeed the case, and that even the converse relation holds, we recall from theorem 20.6.1 that if H is a closed subgroup of G, as it is the case here, then a H-reduction of a principal G-bundle is equivalent to a section of the associated coset bundle with fiber G/H. It is thus helpful to first clarify the structure of the coset space G/H. Note that for each $g = \left(\begin{smallmatrix} A & 0 \\ L & B \end{smallmatrix}\right) \in G$ we can find a unique $h = \left(\begin{smallmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{smallmatrix}\right) \in H$ such that

$$gh = \begin{pmatrix} \mathbb{1}_m & 0 \\ LA^{-1} & \mathbb{1}_n \end{pmatrix} \in gH\,. \tag{26.5.6}$$

Hence, we have a unique representative for every $gH \in G/H$ for each $LA^{-1} = K \in \mathrm{M}_{n,m}(\mathbb{R})$, and thus a bijection

$$\begin{array}{rcccc} \tilde{\bullet} & : & \mathrm{M}_{n,m}(\mathbb{R}) & \to & G/H \\ & & K & \mapsto & \tilde{K} = \begin{pmatrix} \mathbb{1}_m & 0 \\ K & \mathbb{1}_n \end{pmatrix} H \end{array} . \tag{26.5.7}$$

Next, we study the action $\rho : G \times G/H \to G/H$ by left multiplication, which is essential in the construction of the associated bundle. For this purpose, we calculate

$$\begin{pmatrix} A & 0 \\ L & B \end{pmatrix} \cdot \begin{pmatrix} \mathbb{1}_m & 0 \\ K & \mathbb{1}_n \end{pmatrix} \cdot \begin{pmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{pmatrix} = \begin{pmatrix} \mathbb{1}_m & 0 \\ (L+BK)A^{-1} & \mathbb{1}_n \end{pmatrix} . \tag{26.5.8}$$

We see that ρ acts by affine transformations $K \mapsto (L+BK)A^{-1}$ on the space $G/H \cong \mathrm{M}_{n,m}(\mathbb{R})$, i.e., it preserves the affine structure of this space, and so it is an affine representation of G. This affine representation induces a linear representation $\vec{\rho}$ on the underlying vector space $\overrightarrow{G/H} \cong \mathrm{M}_{n,m}(\mathbb{R})$ constituted by differences of elements of G/H, which is given by

$$\begin{pmatrix} A & 0 \\ L & B \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ \varkappa & 0 \end{pmatrix} \cdot \begin{pmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ B\varkappa A^{-1} & 0 \end{pmatrix} \tag{26.5.9}$$

for $\varkappa \in \mathrm{M}_{n,m}(\mathbb{R})$. Writing the corresponding vector space isomorphism as

$$\begin{array}{rcccc} \vec{\bullet} & : & \mathrm{M}_{n,m}(\mathbb{R}) & \to & \overrightarrow{G/H} \\ & & \varkappa & \mapsto & \vec{\varkappa} = \begin{pmatrix} 0 & 0 \\ \varkappa & 0 \end{pmatrix} H \end{array} , \tag{26.5.10}$$

We see that $\tilde{\bullet}$ is an affine isomorphism covering the linear isomorphism $\vec{\bullet}$. Further, following theorem 20.5.1, we find that $F_\pi(TE) \times_\rho G/H$ is an affine bundle modeled over the vector bundle $F_\pi(TE) \times_{\vec{\rho}} \overrightarrow{G/H}$. Recalling that in definition 26.3.1 we have defined an Ehresmann connection as a sections of $J^1(E)$, which is also an affine bundle over E, modeled over the vector bundle $VE \otimes \pi^*(T^*M)$, one may thus suspect that these bundles are related to each other. We will show this now in several steps. In the first step, we need to relate the underlying vector bundles, since the affine bundles will be built upon them. Hence, we start by proving the following.

Theorem 26.5.3. *The bundle $VE \otimes \pi^*(T^*M)$ is canonically isomorphic to the associated vector bundle $F_\pi(TE) \times_{\vec{\rho}} \overrightarrow{G/H}$.*

Proof. We consider an equivalence class $[p, \vec{\varkappa}] \in F_\pi(TE) \times_{\vec{\rho}} \overrightarrow{G/H}$, where $p \in F_\pi(TE)$ is an adapted frame at $e \in E$ and $\varkappa \in \mathrm{M}_{n,m}(\mathbb{R})$. This allows us to define a linear function

$$\begin{array}{rcccc} \Phi(p,\varkappa) & : & T_eE & \to & V_eE \\ & & u & \mapsto & p\left(\begin{pmatrix} 0 & 0 \\ -\varkappa & \mathbb{1}_n \end{pmatrix} \cdot p^{-1}(u)\right) \end{array} . \tag{26.5.11}$$

We see that $\Phi(p,\varkappa)(u)$ is vertical for all $u \in T_eE$, and so we can identify $\Phi(p,\varkappa)$ with a homomorphism $\Phi(p,\varkappa) \in \mathrm{Hom}(T_eE, V_eE) \cong V_eE \otimes T_e^*E$. Further, if $u \in V_eE$ is already vertical, and thus of the form $p(0,w)$ for some $w \in \mathbb{R}^n$, we have $\Phi(p,\varkappa)(u) = 0$. Hence, $\Phi(p,\varkappa)$ vanishes on vertical vectors, and so $\Phi(p,\varkappa) \in V_eE \otimes H_e^*E$. Recalling from theorem 19.3.2 that $H^*E \cong \pi^*(T^*M)$, we have thus constructed an element of the desired bundle $VE \otimes \pi^*(T^*M)$. To show that this is independent of the choice of the representative of $[p,\vec{\varkappa}]$, consider $[p',\vec{\varkappa}']$ with $p' = p \cdot g$ and $\vec{\varkappa}' = \vec{\rho}(g^{-1},\vec{\varkappa})$, so that

$$\varkappa' = B^{-1}\varkappa A\,. \tag{26.5.12}$$

Then we have

$$\begin{aligned}
\Phi(p',\varkappa')(u) &= p'\left(\begin{pmatrix} 0 & 0 \\ -\varkappa' & 0 \end{pmatrix} \cdot p'^{-1}(u)\right) \\
&= p\left(\begin{pmatrix} A & 0 \\ L & B \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ -B^{-1}\varkappa A & 0 \end{pmatrix} \cdot \begin{pmatrix} A^{-1} & 0 \\ -B^{-1}LA^{-1} & B^{-1} \end{pmatrix} \cdot p^{-1}(u)\right) \\
&= p\left(\begin{pmatrix} 0 & 0 \\ -\varkappa A & B \end{pmatrix} \cdot \begin{pmatrix} A^{-1} & 0 \\ -B^{-1}LA^{-1} & B^{-1} \end{pmatrix} \cdot p^{-1}(u)\right) \\
&= p\left(\begin{pmatrix} 0 & 0 \\ -\varkappa & 0 \end{pmatrix} \cdot p^{-1}(u)\right) \\
&= \Phi(p,\varkappa)(u)\,.
\end{aligned} \tag{26.5.13}$$

Since $\Phi(p,\varkappa)$ depends only on the equivalence class $[p,\vec{\varkappa}]$, we can interpret Φ as a map $\Phi : F_\pi \times_\rho \overrightarrow{G/H} \to VE \otimes \pi^*(T^*M)$. This map is smooth by construction, which involves only smooth maps. Further, it preserves the base point, and so it is a bundle morphism covering the identity on E. Also it is by definition linear in $\varkappa$, and thus linear on each fiber, since the vector space structure on each fiber is inherited from $\overrightarrow{G/H} \cong \mathrm{M}_{n,m}(\mathbb{R})$, and therefore a vector bundle homomorphism. To show that it is a vector bundle isomorphism, let $\varphi \in V_eE \otimes H_e^*E \subset V_eE \otimes T_e^*E \cong \mathrm{Hom}(T_eE, V_eE)$ a linear map which vanishes on V_eE. Choosing an adapted frame $p \in F_{\pi\,e}(TE)$, we can construct a linear map

$$\begin{array}{rccc} \vartheta(p,\varphi) & : & \mathbb{R}^{m+n} & \to & \mathbb{R}^{m+n} \\ & & (v,w) & \mapsto & (p^{-1} \circ \varphi \circ p)(v,w) \end{array}\,. \tag{26.5.14}$$

Writing this map as a matrix $\vartheta(p,\varphi) = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, we see that $a = b = 0$, since $\varphi(u) \in V_eE$ is vertical for all $u \in T_eE$, and the adapted frame p^{-1} maps vertical vectors to the subspace $\{0\} \oplus \mathbb{R}^n$. Further, if $u \in V_eE$ is already vertical, we have $\varphi(u) = 0$, and so $d = 0$. Setting $\varkappa = -c$ one thus finds the desired element $[p,\vec{\varkappa}] \in F_\pi(TE) \times_{\vec{\rho}} \overrightarrow{G/H}$. To check that this is independent of the choice of p, one essentially follows the same steps as given before in this proof. One easily checks that this construction yields the inverse of the bundle morphism constructed above, proving that it is indeed a vector bundle isomorphism. ∎

Hence, we have shown that the underlying vector bundles of the affine bundles $J^1(E)$ and $F_\pi(TE) \times_\rho G/H$ are isomorphic as vector bundles. This allows us to proceed with the affine bundles, which is what we do next.

Theorem 26.5.4. *The first jet bundle $\pi_{1,0} : J^1(E) \to E$ is canonically isomorphic to the associated coset bundle $F_\pi(TE) \times_\rho G/H$, where G and H are given in theorems 26.5.1 and 26.5.2, and ρ is the action of G given by left multiplication.*

Proof. In order to construct this isomorphism, we first aim to find a description for the equivalence class $[p,\tilde{K}] \in F_\pi(TE) \times_\rho G/H$, where $p \in F_\pi(TE)$ is an adapted frame at $e \in E$ and $K \in \mathrm{M}_{n,m}(\mathbb{R})$. For this purpose, consider the linear function

$$\begin{array}{rccc} \Theta(p,K) & : & T_eE & \to & V_eE \\ & & u & \mapsto & p\left(\begin{pmatrix} 0 & 0 \\ -K & \mathbb{1}_n \end{pmatrix} \cdot p^{-1}(u)\right) \end{array}\,. \tag{26.5.15}$$

Clearly, the element $\Theta(p,K)(u)$ constructed above is vertical. Also if $u \in V_eE$ is already vertical, and thus of the form $u = p(0,w)$ for some $w \in \mathbb{R}^n$, we have

$$\Theta(p,K)(u) = p\left(\begin{pmatrix} 0 & 0 \\ -K & \mathbb{1}_n \end{pmatrix} \cdot \begin{pmatrix} 0 \\ w \end{pmatrix}\right) = p(0,w) = u\,, \tag{26.5.16}$$

so that $\Theta(p,K)$ restricts to the identity on V_eE. To check that $\Theta(p,K)$ is independent of the representative, let $p' = p \cdot g$ and $\tilde{K}' = \rho(g^{-1}, \tilde{K})$, so that

$$K' = (-B^{-1}LA^{-1} + B^{-1}K)A = B^{-1}(KA - L)\,. \tag{26.5.17}$$

By direct calculation we then find

$$\begin{aligned}
\Theta(p',K')(u) &= p'\left(\begin{pmatrix} 0 & 0 \\ -K' & \mathbb{1}_n \end{pmatrix} \cdot p'^{-1}(u)\right) \\
&= p\left(\begin{pmatrix} A & 0 \\ L & B \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ B^{-1}(L-KA) & \mathbb{1}_n \end{pmatrix} \cdot \begin{pmatrix} A^{-1} & 0 \\ -B^{-1}LA^{-1} & B^{-1} \end{pmatrix} \cdot p^{-1}(u)\right) \\
&= p\left(\begin{pmatrix} 0 & 0 \\ L-KA & B \end{pmatrix} \cdot \begin{pmatrix} A^{-1} & 0 \\ -B^{-1}LA^{-1} & B^{-1} \end{pmatrix} \cdot p^{-1}(u)\right) \\
&= p\left(\begin{pmatrix} 0 & 0 \\ -K & \mathbb{1}_n \end{pmatrix} \cdot p^{-1}(u)\right) \\
&= \Theta(p,K)(u)\,.
\end{aligned} \tag{26.5.18}$$

Hence, we see that Θ assigns to each equivalence class $[p,\tilde{K}] \in F_\pi(TE) \times_\rho G/H$ a linear projection $\Theta(p,\tilde{K}) : T_eE \to V_eE$. Following theorem 26.3.1, this uniquely determines a jet $j^1_{\pi(e)}\sigma$ with $\sigma(\pi(x)) = e$ and $\Theta(p,\tilde{K}) \circ \sigma_* = 0$. Hence, we have constructed a map from $F_\pi(TE) \times_\rho G/H$ to $J^1(E)$. This map is smooth by construction, as can be verified by constructing the local trivializations, and preserves the base point, and so it is a bundle morphism.

To show that the bundle morphism we constructed above is an isomorphism, we explicitly construct its inverse. For this purpose, we make use of theorem 26.3.1 again, and represent an element of $J^1(E)$ by a linear projection $\theta_e : T_eE \to V_eE$. Choosing an adapted frame $p \in F_{\pi\,e}(TE)$, we can construct a linear map

$$\begin{array}{rccc} \kappa(p,\theta_e) & : & \mathbb{R}^{m+n} & \to & \mathbb{R}^{m+n} \\ & & (v,w) & \mapsto & (p^{-1} \circ \theta_e \circ p)(v,w) \end{array}. \tag{26.5.19}$$

Writing this map as a matrix $\kappa(p,\theta_e) = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, we see that $a = b = 0$, since $\theta_e(u) \in V_eE$ is vertical for all $u \in T_eE$, and the adapted frame p^{-1} maps vertical vectors to the subspace $\{0\} \oplus \mathbb{R}^n$. Further, if $u \in V_eE$ is already vertical, we have $\theta_e(u) = u$, and so $d = \mathbb{1}_n$. Setting $K = -c$ one thus finds the desired element $[p,\tilde{K}] \in F_\pi(TE) \times_\rho G/H$. To check that this is independent of the choice of p, one essentially follows the same steps as given before in this proof. One easily checks that this construction yields the inverse of the bundle morphism constructed above, proving that it is indeed a bundle isomorphism.

Finally, we show that the bundle isomorphism we constructed restricts to an affine map on every fiber. For this purpose, recall that we can represent an element $\vec{\varkappa} \in \overrightarrow{G/H}$ by $\varkappa \in \mathrm{M}_{n,m}(\mathbb{R})$. One easily checks that

$$\Theta(p, K + \varkappa) = \Theta(p,K) + \Phi(p,\varkappa)\,. \tag{26.5.20}$$

Hence, we have constructed an affine bundle isomorphism relating $F_\pi(TE) \times_\rho G/H$ and $\pi_{1,0} : J^1(E) \to E$, whose linear derivative is constructed in the proof of theorem 26.5.3. ■

With this result at hand, the following conclusion is straightforward.

Theorem 26.5.5. *There is a one-to-one correspondence between connections ω on a fiber bundle $\pi : E \to M$ and $(\mathrm{GL}(m,\mathbb{R}) \times \mathrm{GL}(n,\mathbb{R}))$-reductions of its adapted frame bundle $F_\pi(TE)$.*

Proof. Following theorem 20.6.1, there is a one-to-one correspondence between H-reductions of the adapted frame bundle $F_\pi(TE)$ and sections of the coset bundle $F_\pi(TE) \times_\rho G/H$. This coset bundle is isomorphic to $J^1(E)$ as shown in theorem 26.5.4, whose sections are the connections on $\pi : E \to M$.

Alternatively, one can proof the equivalence by expressing a connection in terms of a horizontal distribution. Every horizontal distribution HE defines a split $TE = HE \oplus VE$ of the tangent bundle, and thus gives rise to a split frame bundle, which defines a H-reduction of the adapted frame bundle. Conversely, every H-reduction of the adapted frame bundle yields a split frame bundle, whose horizontal components span a unique horizontal tangent space in each fiber, and thus define a horizontal distribution. ■

26.6 Horizontal vector fields

The horizontal lift map is useful to provide explicit formulas which allow us to lift further geometric objects from the base manifold M to the total space E. The first such object we discuss here is a vector field, for which we define the following notion.

Definition 26.6.1 (Horizontal lift of a vector field)**.** Let $\pi : E \to M$ be a fiber bundle equipped with a horizontal distribution HE and $X \in \mathrm{Vect}(M)$ a vector field. The *horizontal lift* of X is the unique horizontal vector field $\hat{X} \in \Gamma(HE)$ such that $\pi_* \circ \hat{X} = X \circ \pi$.

In this definition we have made use of the fact that for every $e \in E$ the differential $\pi_* : T_eE \to T_{\pi(e)}M$ restricts to a vector space isomorphism between H_eE and $T_{\pi(e)}M$, so that we can define $\hat{X}(e)$ to be the unique preimage of $X(\pi(e))$ under this isomorphism. We can write this preimage in terms of the horizontal lift map, using the following statement.

Theorem 26.6.1. *Let $\pi : E \to M$ be a fiber bundle equipped with a horizontal lift map $\eta : \pi^*TM \to HE$ and $X \in \mathrm{Vect}(M)$ a vector field. Then the horizontal lift $\hat{X}$ of X is given by $\hat{X}(e) = \eta(e, X(\pi(e)))$ for all $e \in E$.*

Proof. By definition of a vector field, X satisfies

$$(\tau \circ X \circ \pi)(e) = \pi(e)\,, \tag{26.6.1}$$

where $\tau : TM \to M$ is the tangent bundle projection. Hence, $(e, X(\pi(e))) \in \pi^*TM$. Further, the images under η are horizontal, and η covers the identity on E. Hence, $\eta(e, X(\pi(e))) \in H_eE$, and so $e \mapsto \eta(e, X(\pi(e)))$ defines a vector field on E. Finally, by definition of η, this vector field satisfies

$$\pi_*(\eta(e, X(\pi(e))) = X(\pi(e))\,, \tag{26.6.2}$$

and so it agrees with the unique vector field $\hat{X}$. ■

One can now prove a few helpful statements.

Theorem 26.6.2. *Let $X, Y \in \mathrm{Vect}(M)$ be vector fields and $f \in C^\infty(M, \mathbb{R})$ a function. Then the horizontal lift $\hat{\bullet} : \mathrm{Vect}(M) \to \mathrm{Vect}(E)$ on a bundle $\pi : E \to M$ equipped with a connection satisfies:*

$$\widehat{X + Y} = \hat{X} + \hat{Y}\,, \tag{26.6.3a}$$

$$\widehat{fX} = (f \circ \pi)\hat{X}\,, \tag{26.6.3b}$$

$$\widehat{[X, Y]} = [\hat{X}, \hat{Y}]_H\,. \tag{26.6.3c}$$

Proof. The first two relations follow immediately from the formula given in theorem 26.6.1. For the last formula, one uses the fact that the horizontal lift is the unique vector field on the total space E which projects to the original vector field on the base manifold M. $[\hat{X},\hat{Y}]_H$ is obviously horizontal by definition. Further, we use the fact that a tangent vector is fully determined by its action on a function. For every $f \in C^\infty(M,\mathbb{R})$ and $e \in E$ we have

$$\begin{aligned} \pi_*([\hat{X},\hat{Y}](e))(f) &= [\hat{X},\hat{Y}](e)(f\circ\pi) \\ &= (\hat{X}\hat{Y}(f\circ\pi))(e) - (\hat{Y}\hat{X}(f\circ\pi))(e) \\ &= (XYf)(\pi(e)) - (YXf)(\pi(e)) \\ &= [X,Y](\pi(e))(f)\,, \end{aligned} \tag{26.6.4}$$

and thus $\pi_* \circ [\hat{X},\hat{Y}] = [X,Y]\circ\pi$. ■

Note that the commutator of two horizontal vector fields is not necessarily horizontal. The importance of this will become clear in section 26.10.

26.7 Horizontal curves

Another object we may lift using a connection is a curve. Here we define the following notion.

Definition 26.7.1 (Horizontal lift of a curve)**.** Let $\pi : E \to M$ be a fiber bundle equipped with a horizontal distribution HE and $\gamma : (a,b) \to M$ a curve. A curve $\hat{\gamma} : (a',b') \to E$ is called a *horizontal lift* of γ on $(a',b') \subseteq (a,b)$ if $\gamma = \pi\circ\hat{\gamma}$ and $\dot{\hat{\gamma}}(t) \in H_{\hat{\gamma}(t)}E$ for all $t \in (a',b')$.

Note that in contrast to the horizontal lift of a vector field, that of a curve is not unique. In general, there exist many different horizontal lifts of the same curve. However, we can uniquely select one of them, at least locally by specifying a point on one fiber which it is supposed to pass. This can be stated as follows.

Theorem 26.7.1. *Let $\pi : E \to M$ be a fiber bundle equipped with an Ehresmann connection $\omega : E \to J^1(E)$, $\gamma : \mathbb{R} \to M$ a curve and $e \in E_{\gamma(0)}$. Then there exists $\epsilon > 0$ and a unique horizontal lift $\hat{\gamma} : (-\epsilon,\epsilon)$ of γ on $(-\epsilon,\epsilon)$ such that $\hat{\gamma}(0) = e$.*

Proof. For this proof it is helpful to use the notion of pullback bundles. Let $\gamma^*\pi : \gamma^*E \to \mathbb{R}$ be the pullback of $\pi : E \to M$ along γ. Recall that the elements of γ^*E are pairs $(t,e) \in \mathbb{R}\times E$ such that $\gamma(t) = \pi(e)$. Let $\mathrm{pr}_2 : \gamma^*E \to E, (t,e) \mapsto e$ denote the projection onto the second factor. For every $(t,e) \in \gamma^*E$, let $j^1_{\gamma(t)}\sigma = \omega(e)$ be the jet that is defined by the connection ω and $\sigma \in \Gamma_{\gamma(t)}(E)$ a local section with domain $U \subset M$, which is representative of this jet. Note that $\sigma(\gamma(t)) = e$ and $\sigma_*(\dot{\gamma}(t)) \in H_eE$ is horizontal by construction. Further, we can define a local curve

$$\begin{array}{cccc} \gamma_\sigma & : & \gamma^{-1}(U) & \to & \gamma^*E \\ & & \tilde{t} & \mapsto & (\tilde{t},\sigma(\gamma(\tilde{t}))) \end{array}. \tag{26.7.1}$$

This is well-defined, since we have

$$\gamma(\tilde{t}) = \pi(\sigma(\gamma(\tilde{t}))) \tag{26.7.2}$$

for all $\tilde{t} \in \gamma^{-1}(U)$. Further, let $X(t,e) = \dot{\gamma}_\sigma(t) \in T\gamma^*E$ the tangent vector of this curve at $\tilde{t} = t$. Note that we have $\gamma_\sigma(t) = (t,e)$ and thus $X(t,e) \in T_{(t,e)}E$. Further, $X(t,e)$ depends only on the first jet $j^1_t\gamma_\sigma$, and hence only on $j^1_{\gamma(t)}\sigma = \omega(e)$, but not on the choice of the representative σ. Doing this for all $(t, e) \in \gamma^*E$, we have defined a vector field

$$\begin{array}{cccccc} X & : & \gamma^*E & \to & T\gamma^*E \\ & & (t,e) & \mapsto & \eta(e,\dot{\gamma}(t)) \end{array}, \tag{26.7.3}$$

such that $\mathrm{pr}_{2*}(X(t, e))$ is horizontal and

$$\pi_*(\mathrm{pr}_{2*}(X(t,e))) = \dot{\gamma}(t) \in T_{\gamma(t)}M = T_{\pi(e)}M\,. \tag{26.7.4}$$

For every $e \in E_{\gamma(0)}$, there exists $\epsilon > 0$ and an integral curve $\Gamma : (-\epsilon, \epsilon) \to \gamma^*E$ of X such that $\Gamma(0) = (0, e)$ and $\dot{\Gamma} = X \circ \Gamma$. Finally, define $\hat{\gamma} = \mathrm{pr}_2 \circ \Gamma : (-\epsilon, \epsilon) \to E$. One finds that $\hat{\gamma}$ is a horizontal lift of γ with $\hat{\gamma}(0) = e$. ■

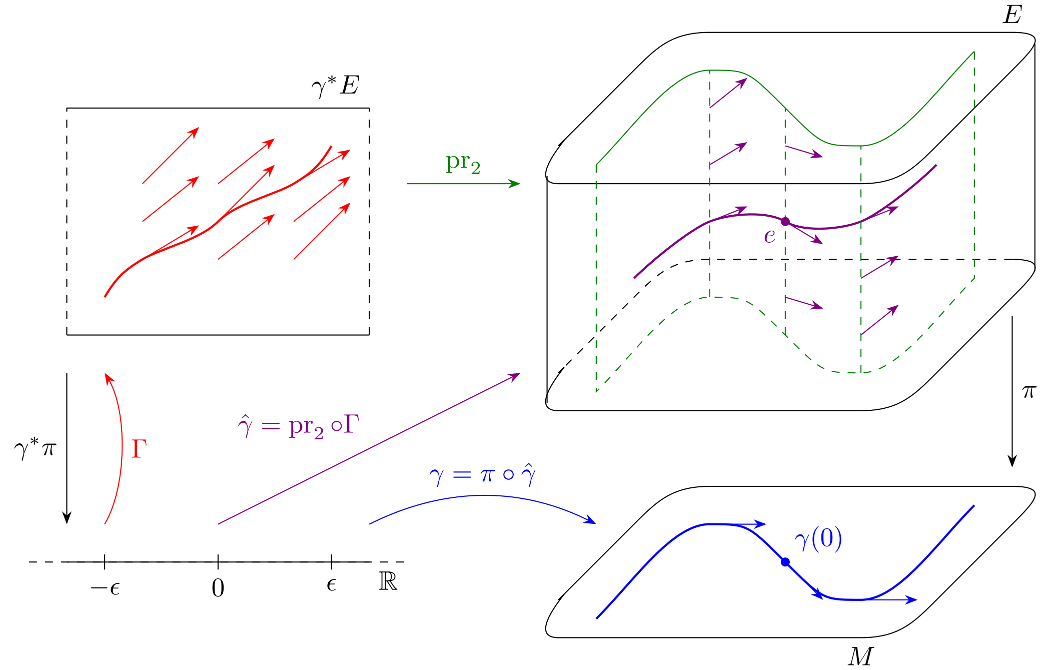


Figure 26.2: Illustration of the horizontal lift of a curve as the image of an integral curve on the pullback bundle.

The construction is illustrated in figure 26.2. Note that, in general, the horizontal lift can be constructed only locally. This can be seen from the proof above, since the vector field X we introduced may not be complete, and thus only locally admit integral curves. In fact, we may define a notion which is similar to that of a complete vector field as follows.

Definition 26.7.2 (Complete connection)**.** Let $\pi : E \to M$ be a fiber bundle. A connection on E is called *complete* if and only if for every curve $\gamma : (a, b) \to M$, $t \in (a, b)$ and $e \in E_{\gamma(t)}$ there exists a horizontal lift $\hat{\gamma} : (a, b) \to E$ of γ, such that $\hat{\gamma}(t) = e$ and $\hat{\gamma}$ has the same domain as γ.

To see that not every connection is complete, consider the following counterexample.

Example 26.7.1. Let $M = \mathbb{R}$ and $E = M \times (0, 1)$ the trivial bundle. We use a global chart given by the canonical embedding $E \to \mathbb{R}^2$ to define coordinates $(x, y) \in \mathbb{R} \times (0, 1)$, where x is a coordinate on M. We specify the connection on E by the choice of the horizontal basis vectors, which we define as $\delta_x = \partial_x + \bar{\partial}_y$. Then consider the curve $\gamma = \mathrm{id}_M : \mathbb{R} \to M$ and let $e = (x, y) \in E_x$ for some $x \in \mathbb{R}$. We find that the horizontal lift $\hat{\gamma}$ of γ passing through e can be defined only on the domain $(x - y, x - y + 1)$ and is given by $t \mapsto (t, t - x + y)$. See also example 16.1.1, where essentially the same argument holds.

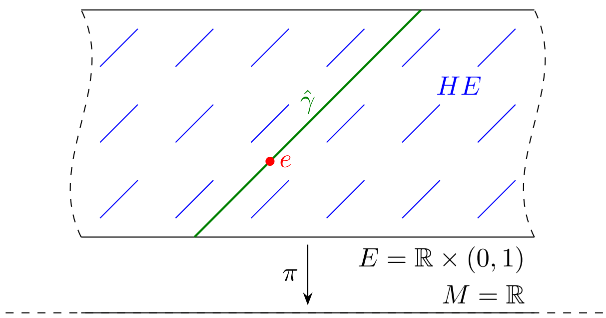


Figure 26.3: The connection defined in example 26.7.1 is not complete, since every horizontal lift can be defined only on a finite domain.

26.8 Parallel transport

If one has a complete connection, one can define the following notion, which plays an important role in different areas of physics:

Definition 26.8.1 (Parallel transport)**.** Let $\pi : E \to M$ be a fiber bundle equipped with a complete connection. For every curve $\gamma : [0, 1] \to M$ the *parallel transport* is the map $\mathcal{P}_\gamma : E_{\gamma(0)} \to E_{\gamma(1)}$ such that $\mathcal{P}_\gamma(e) = \hat{\gamma}_e(1)$, where $\hat{\gamma}_e : [0, 1] \to E$ is the unique horizontal lift of γ with $\hat{\gamma}_e(0) = e$.

Theorem 26.8.1. *Let $\pi : E \to M$ be a fiber bundle equipped with a complete connection, $\gamma : [0, 1] \to M$ a curve and $\bar{\gamma} : [0, 1] \to M, t \mapsto \gamma(1 - t)$ the reversed curve. Then $\mathcal{P}_{\bar{\gamma}} = \mathcal{P}_\gamma^{-1}$.*

Proof. ►. . . ◄ ■

Theorem 26.8.2. *Let $\pi : E \to M$ be a fiber bundle equipped with a complete connection. For every curve $\gamma : [0, 1] \to M$ the parallel transport $\mathcal{P}_\gamma : E_{\gamma(0)} \to E_{\gamma(1)}$ is a diffeomorphism.*

Proof. ►. . . ◄ ■

26.9 Integral sections

If we have a section, whose partial derivatives at each point agree with those determined by the connection, then we give it a particular name, using the following definition.

Definition 26.9.1 (Integral section). Let $\pi : E \to M$ be a fiber bundle and $\omega : E \to J^1(E)$ an Ehresmann connection. A local section $\sigma : U \to E$ on $U \subset M$ is called an *integral section* of ω if $j^1\sigma = \omega \circ \sigma$.

For the specially chosen section σ_e we demanded that the aforementioned tangent vectors at e are horizontal, i.e., that $\sigma_{e*}(u)$ is horizontal for $u \in T_{\pi(e)}M$. For an integral section this must hold at *every* point. It is important to know that not every Ehresmann connection admits integral sections. There are conditions when an Ehresmann connection is *integrable*, at least locally, which we will encounter later.

26.10 Curvature

We have seen in the first sections of this chapter that there exist different, equivalent ways to specify a connection on a general fiber bundle $\pi : E \to M$. An essential part which these definitions have in common is the fact that they provide a unique prescription how to decompose an arbitrary vector $w \in TE$ over some point $e = \pi(w) \in E$ into horizontal and vertical parts, $w_H \in H_eE$ and $w_V \in V_eE$. Applying this prescription pointwise to a vector field $X \in \text{Vect}(E)$, we obtain its horizontal and vertical parts, $X_H \in \Gamma(HE)$ and $X_V \in \Gamma(VE)$. Following theorem 19.2.4, the commutator $[X_V, Y_V]$ of the vertical parts of any vector fields $X, Y \in \text{Vect}(E)$ is again vertical. The commutator of the horizontal parts, however, is in general not horizontal, but may have a vertical part. This fact plays an important role in the theory of connections, and deserves its own definition.

Definition 26.10.1 (Curvature of a general connection). Let $\pi : E \to M$ be a fiber bundle equipped with a connection, whose horizontal and vertical projectors we denote by $\bullet_H$ and $\bullet_V$. The *curvature form* of this connection is the unique vertical-valued two-form $R \in \Gamma(\Lambda^2 T^*E \otimes VE)$ such that for all $X, Y \in \text{Vect}(E)$ holds

$$R(X, Y) = -[X_H, Y_H]_V \,. \tag{26.10.1}$$

In the definition above we have claimed that the curvature defines a vertical-valued two-form. In order for this to hold true, it must satisfy a number of properties, which we will show next.

Theorem 26.10.1. *For all $X, Y, Z \in \text{Vect}(E)$ and $f \in C^\infty(M, \mathbb{R})$, the curvature R satisfies:*

$$R(Y, X) = -R(X, Y) \,, \tag{26.10.2a}$$

$$R(X + Y, Z) = R(X, Z) + R(Y, Z) \,, \tag{26.10.2b}$$

$$R(fX, Y) = fR(X, Y) \,. \tag{26.10.2c}$$

Proof. The first property follows immediately from the fact that the Lie bracket of vector fields is antisymmetric. Similarly, the second property follows from the linearity of the Lie bracket and the horizontal and vertical projectors. Finally, for the third property we calculate

$$\begin{aligned} R(fX, Y) &= [fX_H, Y_H]_V \\ &= (f[X_H, Y_H] - (Y_H f)X_H)_V \\ &= f[X_H, Y_H]_V \\ &= fR(X, Y) \,, \end{aligned} \tag{26.10.3}$$

using the fact that X_H is horizontal, and so its vertical part vanishes. ∎

It is instructive to calculate the curvature using the coordinates and the horizontal-vertical basis we introduced before. Consider two horizontal vector fields $X = X_H = X^\mu\delta_\mu$ and $Y = Y_H = Y^\mu\delta_\mu$. We already omitted their vertical part, since it will not contribute to the curvature (26.10.1). By direct calculation we then find

$$\begin{aligned} R(X,Y) &= -[X_H, Y_H]_V \\ &= -[X^\mu\delta_\mu, Y^\nu\delta_\nu]_V \\ &= -[X^\mu(\partial_\mu - \theta^a_\mu\bar{\partial}_a), Y^\nu(\partial_\nu - \theta^b_\nu\bar{\partial}_b)]_V \\ &= -[X^\mu(\partial_\mu - \theta^a_\mu\bar{\partial}_a)Y^\nu(\partial_\nu - \theta^b_\nu\bar{\partial}_b) - X^\mu Y^\nu(\partial_\mu - \theta^a_\mu\bar{\partial}_a)\theta^b_\nu\bar{\partial}_b \\ &\quad - Y^\nu(\partial_\nu - \theta^b_\nu\bar{\partial}_b)X^\mu(\partial_\mu - \theta^a_\mu\bar{\partial}_a) + Y^\nu X^\mu(\partial_\nu - \theta^b_\nu\bar{\partial}_b)\theta^a_\mu\bar{\partial}_a]_V \\ &= -\left(X^\mu\delta_\mu Y^\nu\delta_\nu - X^\mu Y^\nu\delta_\mu\theta^b_\nu\bar{\partial}_b - Y^\nu\delta_\nu X^\mu\delta_\mu + Y^\nu X^\mu\delta_\nu\theta^a_\mu\bar{\partial}_a\right)_V \\ &= X^\mu Y^\nu(\delta_\mu\theta^a_\nu - \delta_\nu\theta^a_\mu)\bar{\partial}_a\,. \end{aligned} \tag{26.10.4}$$

This shows that R is indeed tensorial, since the result does not depend on the derivatives of the vector field components. Hence, we can write

$$R = \frac{1}{2}R^a{}_{\mu\nu}\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \otimes \bar{\partial}_a = \frac{1}{2}(\delta_\mu\theta^a_\nu - \delta_\nu\theta^a_\mu)\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \otimes \bar{\partial}_a\,, \tag{26.10.5}$$

and the components are given by the formula

$$R^a{}_{\mu\nu} = \delta_\mu\theta^a_\nu - \delta_\nu\theta^a_\mu = \partial_\mu\theta^a_\nu - \partial_\nu\theta^a_\mu + \theta^b_\nu\bar{\partial}_b\theta^a_\mu - \theta^b_\mu\bar{\partial}_b\theta^a_\nu\,. \tag{26.10.6}$$

Since the curvature form is a particular type of vector valued form, one may ask whether it is related to the operations on graded derivations we discussed in chapter 17. Indeed, recalling that we may interpret a connection form as a vertical-valued one-form $\theta \in \Gamma(T^*E \otimes VE)$, we find the following relation.

Theorem 26.10.2. *The curvature form R of a connection form θ is given by its Nijenhuis tensor*

$$R = -N_\theta = -\frac{1}{2}[\![\theta, \theta]\!]\,. \tag{26.10.7}$$

Proof. Using theorem 17.6.6, we find by direct calculation

$$\begin{aligned} N_\theta(X,Y) &= [\theta X, \theta Y] + \theta^2[X,Y] - \theta([\theta X, Y] + [X, \theta Y]) \\ &= [X_V, Y_V] + [X,Y]_V - [X_V, Y]_V + [X, Y_V]_V \\ &= [X_V, Y_V]_V + [X_H + X_V, Y_H + Y_V]_V - [X_V, Y_H + Y_V]_V - [X_H + X_V, Y_V]_V \\ &= [X_H, Y_H]_V\,, \end{aligned} \tag{26.10.8}$$

where we made use of the fact that θ is the projection to the vertical part and that the vertical distribution is integrable, so that $[X_V, Y_V]_H = 0$. ■

We remark that in the literature one also finds the curvature to be defined with the opposite sign. Here we use this sign convention in order to be consistent with the sign convention for the curvature of principal connections introduced in section 27.3 and that of linear connections in section 28.12.

From the definition of the curvature directly follows an important relation.

Theorem 26.10.3 (Bianchi identity)**.** *The curvature form R of a connection form θ satisfies* $[\![\theta, R]\!] = 0$.

Proof. From the graded Jacobi identity (17.6.5) and symmetry (17.6.4) immediately follows

$$[\![\theta, [\![\theta, \theta]\!]]\!] = 0\,. \tag{26.10.9}$$

■

Definition 26.10.2 (Flat connection)**.** A connection is called *flat* if and only if its curvature vanishes.

26.11 Canonical flat connection

Definition 26.11.1 (Canonical flat connection)**.** Let (E, M, π, F) with $E = M \times F$ and $\pi = \mathrm{pr}_1$ be a trivial fiber bundle. The *canonical flat connection* on (E, M, π, F) is the connection which assigns to each (m, f) in $M \times F$ the jet $j^1_m \sigma_f$, where $\sigma_f : M \to M \times F, m \mapsto (m, f)$ is the constant section.

Theorem 26.11.1. *The canonical flat connection is flat.*

Proof. ►. . . ◄ ■

26.12 Fibered product connection

So far we have discussed the definition and properties of a connection on a given fiber bundle. In the following, we will discuss the question whether and how one may obtain a connection on a constructed bundle from connections defined on its constituents. In the previous section we have already seen that on a trivial bundle, a canonical connection is defined by the product manifold structure. We will now discuss the following construction.

Definition 26.12.1 (Fibered product connection)**.** Let (E_1, M, π_1, F_1) and (E_2, M, π_2, F_2) be fiber bundles equipped with Ehresmann connections $\omega_1 : E_1 \to J^1(E_1)$ and $\omega_2 : E_2 \to J^1(E_2)$. The *fibered product connection* is the connection $\omega_1 \times_M \omega_2 : E_1 \times_M E_2 \to J^1(E_1 \times_M E_2)$ that assigns to $(e_1, e_2) \in E_1 \times_M E_2$ with $\pi_1(e_1) = \pi_2(e_1) = x \in M$ the jet $j^1_x(\sigma_1, \sigma_2)$, where $\omega_1(e_1) = j^1_x \sigma_1$ and $\omega_1(e_2) = j^1_x \sigma_2$.

26.13 Pullback connection

A construction which is particularly useful is the following, that allows to construct a connection on a pullback bundle. We start with its definition.

Definition 26.13.1 (Pullback connection)**.** Let M be a manifold, (E, B, π, F) a fiber bundle, $\omega : E \to J^1(E)$ an Ehresmann connection and $\psi : M \to B$ a map. The *pullback connection* $\psi^*\omega : \psi^*E \to J^1(\psi^*E)$ is the connection on the pullback bundle $(\psi^*E, M, \psi^*\pi, F)$ that assigns to each $(m, e) \in \psi^*E$ the jet $j^1_m(\mathrm{id}_M, \sigma \circ \psi)$, where $\omega(e) = j^1_{\psi(m)}\sigma$.

It is clear that $j^1_m(\sigma \circ \psi)$ is a jet at $m = \psi^*\pi(m, e)$. Further, it does not depend on the choice of the representative $\sigma \in \Gamma_{\psi(m)}(E)$, since it is given by the jet of a composition. This shows that $\psi^*\omega$ indeed defines an Ehresmann connection on the pullback bundle.

Chapter 27

Principal connections

27.1 Connections on principal bundles

Recall that a principal G-bundle $\pi : P \to M$ is equipped with a right action of a Lie group G which is fiber preserving and free and transitive on the fibers. For $p \in P$ and $g \in G$ we can write this action in the form $R_g(p) = p \cdot g$. This right action also induces a right action on the space $\Gamma(P)$ of (local) sections given by $R_g(\sigma) = R_g \circ \sigma$ for $\sigma \in \Gamma(P)$. To see that this is indeed a right action and not a left action, one can explicitly calculate

$$\begin{aligned} R_{gh}(\sigma)(x) &= (R_{gh} \circ \sigma)(x) \\ &= R_{gh}(\sigma(x)) \\ &= \sigma(x) \cdot (gh) \\ &= \sigma(x) \cdot g \cdot h \\ &= R_g(\sigma(x)) \cdot h \\ &= R_h(R_g(\sigma(x))) \\ &= (R_h \circ R_g \circ \sigma)(x) \\ &= R_h(R_g(\sigma))(x)\,. \end{aligned} \tag{27.1.1}$$

Note that since R_g is a fiber preserving diffeomorphism for all $g \in G$, the r-jets for any $r \in \mathbb{N}$ of the images $R_g(\sigma), R_g(\tau)$ of two sections $\sigma, \tau \in \Gamma(P)$ agree if any only if σ and τ have the same r-jets. For all $x \in M$ thus holds

$$j^r_x R_g(\sigma) = j^r_x R_g(\tau) \quad \Leftrightarrow \quad j^r_x \sigma = j^r_x \tau\,. \tag{27.1.2}$$

This defines a right action on the jet spaces $J^r(P)$ given by $R_g(j^r_x\sigma) = j^r_x R_g(\sigma)$. Making use of this right action on $J^1(P)$ we can now define the following.

Definition 27.1.1 (Principal Ehresmann connection)**.** Let $\pi : P \to M$ be a principal G-bundle with Lie group G. A *principal Ehresmann connection* on P is a G-equivariant section of the jet bundle $\pi_{1,0} : J^1(P) \to P$.

In addition to the definition of a general Ehresmann connection we thus have the condition that the section $\omega : P \to J^1(P)$ must be G-equivariant. To study the consequences of this condition, recall that an Ehresmann connection assigns to every $p \in P$ a jet $\omega(p) = j^1_x\sigma_p$ with $x = \pi(p)$, where σ_p is a local section of $\pi : P \to M$ around x such that $\sigma_p(x) = p$. The condition of equivariance then takes the form

$$j^1_x\sigma_{p\cdot g} = \omega(p \cdot g) = \omega(R_g(p)) = R_g(\omega(p)) = R_g(j^1_x\sigma_p) = j^1_x(R_g \circ \sigma_p)\,. \tag{27.1.3}$$

Of course also principal Ehresmann connections, defined as sections of the jet bundle following section 26.3, can be expressed using connection forms, as in section 26.2, which leads to a more commonly used description. Recall that a connection form on a bundle $\pi : P \to M$ is a vector bundle homomorphism $\theta : TP \to VP$ covering the identity map on P and restricting to the identity on VP. Since both TP and VP carry right actions by the Lie group G, which are given by the differential R_{g*} of the right action on P, there is a well-defined notion of G-equivariant connection forms. The following statement should thus not be a big surprise.

Theorem 27.1.1. *For every principal G-bundle $\pi : P \to M$ with Lie group G there is a one-to-one correspondence between principal Ehresmann connections and G-equivariant connection forms on P.*

Proof. We have already proven that for general fiber bundles there is a one-to-one correspondence between Ehresmann connections ω and connection forms θ. We now have to show that ω is a principal Ehresmann connection if and only if θ is G-equivariant. We will thus start with a principal Ehresmann connection ω, which assigns to $p \in P$ with $\pi(p) = x$ the jet $\omega(p) = j^1_x\sigma_p$. This defines the connection form θ_p at p as $w \mapsto w - \sigma_{p*}(\pi_*(w))$, as shown for general Ehresmann connections. To see that θ is equivariant, we check that

$$\begin{aligned} \theta_{p\cdot g}(R_{g*}(w)) &= R_{g*}(w) - \sigma_{p\cdot g*}(\pi_*(R_{g*}(w))) \\ &= R_{g*}(w) - (R_g \circ \sigma_p)_*(\pi_*(w)) \\ &= R_{g*}(w - \sigma_{p*}(\pi_*(w))) \\ &= R_{g*}(\theta_p(w))\,. \end{aligned} \tag{27.1.4}$$

Thus, θ is equivariant. We also see from the derivation above that if ω is not a principal Ehresmann connection, then θ is not equivariant. ■

We can thus describe any principal Ehresmann connection in terms of a G-equivariant connection form. However, it is more common to replace the target space VP of the connection form by the Lie algebra $\mathfrak{g}$. This is possible, since the fundamental vector fields establish a linear isomorphism between $\mathfrak{g}$ and every vertical tangent space V_pP. One thus often uses the following definition for a connection on a principal bundle.

Definition 27.1.2 (Principal G-connection). Let $\pi : P \to M$ be a principal G-bundle with Lie group G. A *principal G-connection* on P is a $\mathfrak{g}$-valued one-form $\vartheta \in \Omega^1(P, \mathfrak{g})$ on P such that:

- ϑ is G-equivariant: $\vartheta = \mathrm{Ad}_g \circ R^*_g(\vartheta)$ for all $g \in G$.
- For all $X \in \mathfrak{g}$ and $p \in P$ the fundamental vector field $\tilde{X}$ yields $\iota_{\tilde{X}}\vartheta(p) = X$.

This definition requires a few explanations. The space $\Omega^1(P, \mathfrak{g})$ of Lie algebra valued one-forms is simply the tensor product space $\Omega^1(P) \otimes \mathfrak{g}$. The map $\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g})$ is the adjoint representation defined in definition 15.9.1. With this definition, we can now come to the following statement about principal G-connections.

Theorem 27.1.2. *For every principal G-bundle $\pi : P \to M$ with Lie group G there is a one-to-one correspondence between principal Ehresmann connections and principal G-connections on P.*

Proof. For every $p \in P$ there exists a vector space isomorphism $\tilde{\bullet}|_p : \mathfrak{g} \to V_pP$ defined by the fundamental vector fields. Via this isomorphism there exist isomorphisms between the following spaces:

$$\mathrm{Hom}(T_pP, V_pP) \cong \mathrm{Hom}(T_pP, \mathfrak{g}) \cong T^*_pP \otimes \mathfrak{g}\,. \tag{27.1.5}$$

Thus, a vector bundle homomorphism $\theta : TP \to VP$ covering the identity on E uniquely determines a section $\vartheta \in \Omega^1(P, \mathfrak{g})$. It is easy to see that θ restricts to the identity on VP if and only if $\iota_{\tilde{X}}\vartheta(p) = X$ for all $X \in \mathfrak{g}$ and $p \in P$. Further, it follows from the definition of the adjoint representation that θ is G-equivariant if and only if ϑ is G-equivariant. ■

Further recalling section 26.1, we can also describe any connection by a horizontal distribution, and so one may ask what is the distinctive property of a principal connection in this picture. We formulate it as follows.

Theorem 27.1.3. *For every principal G-bundle $\pi : P \to M$ with Lie group G there is a one-to-one correspondence between principal Ehresmann connections and horizontal distributions HP which satisfy $R_{g*}(w) \in HP$ for every $w \in HP$ and $g \in G$.*

Proof. We can express the principal Ehresmann connection via its equivariant connection form $\theta : TP \to VP$. Following theorem 26.2.1, it defines a horizontal distribution as its kernel, $HP = \ker\theta$. Let $p \in P$, $w \in H_pP$ and $g \in G$. From $\theta_p(w) = 0$ and the equivariance of θ then follows

$$0 = R_{g*}(\theta_p(w)) = \theta_{p\cdot g}(R_{g*}(w)), \tag{27.1.6}$$

and thus $R_{g*}(w) \in HP$.

Conversely, let HP be a horizontal distribution satisfying $R_{g*}(w) \in HP$ for every $w \in HP$ and $g \in G$. For $w \in TP$, we denote $w_V \in VP$ and $w_H \in HP$ its vertical and horizontal parts. Again using theorem 26.2.1, the assignment $\theta : TP \to VP, w \mapsto \theta(w) = w_V$ defines a connection form. To check its equivariance, note first that R_g preserves the fibers of P by definition, $\pi \circ R_g = \pi$, and hence $\pi_* \circ R_{g*} = \pi_*$, so that

$$0 = \pi_*(w_V) = \pi_*(R_{g*}(w_V)), \tag{27.1.7}$$

and so $R_{g*}(w_V) \in VP$. Further, $R_{g*}(w_H) \in HP$ by assumption. Since $R_{g*}(w_V) + R_{g*}(w_H) = R_{g*}(w)$ by linearity of R_{g*} and the vertical and horizontal part of a vector $R_{g*}(w)$ are uniquely defined, we have

$$\theta_{p\cdot g}(R_{g*}(w)) = R_{g*}(w)_V = R_{g*}(w_V) = R_{g*}(\theta_p(w)), \tag{27.1.8}$$

proving that θ is equivariant. ■

Finally, we have also described connections as horizontal lift maps in section 26.4. Also this formulation allows us to distinguish principal connections.

Theorem 27.1.4. *For every principal G-bundle $\pi : P \to M$ with Lie group G there is a one-to-one correspondence between principal Ehresmann connections and vector bundle homomorphisms $\eta : \pi^*TM \to TP$ covering the identity on P and satisfying $\pi_* \circ \eta = \mathrm{pr}_2$ and $\eta \circ (R_g, \mathrm{id}_{TM}) = R_{g*} \circ \eta$, where $\mathrm{pr}_2 : \pi^*TM \to TM$ is the projection onto the second factor of $\pi^*TM \cong P \times_M TM$.*

Proof. Given a principal Ehresmann connection $\omega : P \to J^1(P)$, following theorem 26.4.2, we have a horizontal lift map $\eta : \pi^*TM \to TP$ defined by $\eta(p, v) = \sigma_{p*}(v)$, where $\sigma_p \in \Gamma_x(P)$ with $x = \pi(p)$ and $\sigma_p(x) = p$ and $j^1_x\sigma_p = \omega(p)$ is a representative for the jet $\omega(p)$. Since ω is a principal connection, its representative at $p \cdot g$ with $g \in G$ can be chosen to be $\sigma_{p\cdot g} = R_g \circ \sigma_p$, and so for the horizontal lift holds

$$\eta(p \cdot g, v) = \sigma_{(p\cdot g)*}(v) = R_{g*}(\sigma_{p*}(v)) = R_{g*}(\eta(p, v)), \tag{27.1.9}$$

so that

$$\eta \circ (R_g, \mathrm{id}_{TM}) = R_{g*} \circ \eta. \tag{27.1.10}$$

To show the converse direction, it is most simple to represent the Ehresmann connection on P induced by η by its horizontal distribution HP, which is obtained as the image of the map η. Recalling that $\eta : \pi^*TM \to HP$ is a vector bundle isomorphism covering the identity on P, we can identify every horizontal vector $w \in HP$ with $(p, v) = (\chi(w), \pi_*(w)) \in \pi^*TM$, where $\chi : TP \to P$ is the tangent bundle projection, and $w = \eta(p, v)$. Using the equivariance of η we have

$$R_{g*}(w) = R_{g*}(\eta(p, v)) = \eta(p \cdot g, v) \in HP\,, \tag{27.1.11}$$

i.e., $R_{g*}(w) \in HP$ for all $w \in HP$, proving that HP is the horizontal distribution of a principal Ehresmann connection. ∎

It is helpful to illustrate the constructions shown in this section using coordinates. This can be done most easily in the case that G is a matrix group. Hence, we will consider the following example.

Example 27.1.1 (Principal connections for matrix groups). Let $G \subset M_{n,n}(\mathbb{R})$ be a matrix group and $\pi : P \to M$ a principal G-bundle. We will use the matrix components $(g^a{}_b)$ as coordinates on G (imposing suitable restrictions on them, in order to represent only those matrices that lie in G). In order to construct coordinates on P, we pick a local trivialization $\phi : \pi^{-1}(U) \to U \times G$ on an open set $U \subset M$. Introducing coordinates (x^μ) on U, we can write the coordinates on $\pi^{-1}(U)$ which are induced by the trivialization ϕ and the coordinates $(g^a{}_b)$ on G in the form $(x^\mu, p^a{}_b)$. These are the coordinates we will use in this illustration.

We start by illustrating definition 27.1.2 of a principal G-connection $\vartheta \in \Omega^1(P, \mathfrak{g})$. In order to derive a coordinate expression for ϑ, we first investigate the second condition on evaluating ϑ on the fundamental vector field $\tilde{X} \in \Gamma(VP)$ defined by $X \in \mathfrak{g}$. Using the coordinate expression we derived in example 20.1.2 we find that this condition reads

$$X^a{}_b = \vartheta^a{}_b(p)\left(p^c{}_d X^d{}_e \frac{\partial}{\partial p^c{}_e}\right) = p^c{}_d X^d{}_e \vartheta^a{}_b(p)\left(\frac{\partial}{\partial p^c{}_e}\right), \tag{27.1.12}$$

so that

$$\vartheta^a{}_b(p)\left(\frac{\partial}{\partial p^c{}_d}\right) = (p^{-1})^a{}_c \delta^d_b\,. \tag{27.1.13}$$

This completely determines the action of ϑ on vertical tangent vectors. Hence, ϑ must be of the form

$$\vartheta^a{}_b(p) = (p^{-1})^a{}_c \mathrm{d}p^c{}_b + \vartheta^a{}_{b\mu}(p)\mathrm{d}x^\mu\,. \tag{27.1.14}$$

We then make use of the first condition in definition 27.1.2, which states that ϑ must be equivariant, and can be expressed in coordinates as

$$\begin{aligned}
\vartheta^a{}_b(p) &= \mathrm{Ad}_g((R_g^*\vartheta)(p))^a{}_b\\
&= g^a{}_c(g^{-1})^d{}_b(R_g^*\vartheta)^c{}_d(p)\\
&= g^a{}_c(g^{-1})^d{}_b\left[(p'^{-1})^c{}_e\frac{\partial p'^e{}_d}{\partial p^f{}_g}\mathrm{d}p^f{}_g + \vartheta^c{}_{d\mu}(p')\frac{\partial x'^\mu}{\partial x^\nu}\mathrm{d}x^\nu\right]\\
&= g^a{}_c(g^{-1})^d{}_b\left[(g^{-1}p^{-1})^c{}_e\frac{\partial (pg)^e{}_d}{\partial p^f{}_g}\mathrm{d}p^f{}_g + \vartheta^c{}_{d\mu}(p')\frac{\partial x^\mu}{\partial x^\nu}\mathrm{d}x^\nu\right]\\
&= g^a{}_c(g^{-1})^d{}_b\left[(g^{-1})^c{}_h(p^{-1})^h{}_e\delta^e_f g^g{}_d\mathrm{d}p^f{}_g + \vartheta^c{}_{d\mu}(p')\mathrm{d}x^\mu\right]\\
&= (p^{-1})^a{}_c\mathrm{d}p^c{}_b + g^a{}_c(g^{-1})^d{}_b\vartheta^c{}_{d\mu}(p')\mathrm{d}x^\mu\,,
\end{aligned} \tag{27.1.15}$$

where $p' = p \cdot g$ and $x' = x$. This means that once $\vartheta(p)$ is specified for some $p \in P$, then also $\vartheta(p \cdot g)$ is determined for all $g \in G$. Making use of the local trivialization ϕ, we can pick an element over each fiber $\pi^{-1}(x)$ with $x \in U$ by taking $p' = \phi^{-1}(x, e) \in \pi^{-1}(x)$, where $e \in G$ is the unit element. In the coordinates chosen on $\pi^{-1}(U)$ this element is expressed by $(x^\mu, \delta^a{}_b)$. We then define the coordinate expression

$$\vartheta^a{}_{b\mu}(\phi^{-1}(x, e)) = \Gamma^a{}_{b\mu}(x)\,. \tag{27.1.16}$$

For an arbitrary element $p = \phi^{-1}(x, g) = \phi^{-1}(x, e) \cdot g$ we then have

$$\vartheta^a{}_{b\mu}(p) = (p^{-1})^a{}_c p^d{}_b \Gamma^c{}_{d\mu}(x)\,, \tag{27.1.17}$$

so that

$$\vartheta^a{}_b(p) = (p^{-1})^a{}_c \left[\mathrm{d}p^c{}_b + p^d{}_b \Gamma^c{}_{d\mu}(x)\mathrm{d}x^\mu\right]\,. \tag{27.1.18}$$

The principal G-connection ϑ is thus fully determined by the coordinate functions $\Gamma^a{}_{b\mu}(x)$, which are called the *connection coefficients* with respect to the chosen coordinates.

From the coordinate expression of ϑ one can now also derive the connection form θ, which is obtained by taking the fundamental vector field of $\vartheta(p)$ and evaluating it at p. This leads to the expression

$$\theta_p = \frac{\partial}{\partial p^a{}_b} \otimes \left[\mathrm{d}p^a{}_b + p^d{}_b \Gamma^a{}_{d\mu}(x)\mathrm{d}x^\mu\right]\,. \tag{27.1.19}$$

Evaluating this at a tangent vector $w \in T_pP$ yields

$$\theta_p(w) = \theta_p\left(u^\mu \frac{\partial}{\partial x^\mu} + v^a{}_b \frac{\partial}{\partial p^a{}_b}\right) = \left[p^d{}_b \Gamma^a{}_{d\mu}(x) u^\mu + v^a{}_b\right] \frac{\partial}{\partial p^a{}_b}\,. \tag{27.1.20}$$

To see that this connection form is equivariant, let $p' = p \cdot g$ and consider the pushforward

$$R_{g*}(w) = u^\mu \frac{\partial}{\partial x^\mu} + v^a{}_c g^c{}_b \frac{\partial}{\partial p'^a{}_b} \in T_{p'}P \tag{27.1.21}$$

and evaluate

$$\begin{aligned}
\theta_{p\cdot g}(R_{g*}(w)) &= \left[p'^d{}_b \Gamma^a{}_{d\mu}(x) u^\mu + v^a{}_c g^c{}_b\right] \frac{\partial}{\partial p'^a{}_b} \\
&= \left[p^d{}_c \Gamma^a{}_{d\mu}(x) u^\mu + v^a{}_c\right] g^c{}_b \frac{\partial}{\partial p'^a{}_b} \\
&= R_{g*}\left(\left[p^d{}_b \Gamma^a{}_{d\mu}(x) u^\mu + v^a{}_b\right] \frac{\partial}{\partial p^a{}_b}\right) \\
&= R_{g*}(\theta_p(w))\,.
\end{aligned} \tag{27.1.22}$$

Thus, θ is indeed an equivariant connection form.

From the connection form it is now easy to read off the corresponding Ehresmann connection $\omega : P \to J^1(P)$. For this purpose we introduce coordinates $(x^\mu, p^a{}_b, p^a{}_{b\mu})$ on the first jet bundle $J^1(P)$. From the coordinate expression of θ one immediately reads off

$$\omega : (x^\mu, p^a{}_b) \mapsto (x^\mu, p^a{}_b, -p^c{}_b \Gamma^a{}_{c\mu}(x))\,. \tag{27.1.23}$$

One easily sees that $\omega \circ R_g = R_g \circ \omega$ from

$$\omega : (x^\mu, p^a{}_c g^c{}_b) \mapsto (x^\mu, p^a{}_c g^c{}_b, -p^d{}_c g^c{}_b \Gamma^a{}_{d\mu}(x))\,, \tag{27.1.24}$$

which shows that ω is equivariant, and hence a principal Ehresmann connection.

Further, we can now derive the horizontal lift map $\eta : \pi^*TM \to TP$. Given $p \in P$ with $\pi(p) = x \in M$ and $u \in T_xM$, we can write

$$\eta(p, u) = u^\mu \left(\frac{\partial}{\partial x^\mu} - p^c{}_b \Gamma^a{}_{c\mu}(x) \frac{\partial}{\partial p^a{}_b}\right)\,, \tag{27.1.25}$$

and one easily checks that $\theta_p(\eta(p, u)) = 0$. Here equivariance can be seen by calculating

$$R_{g*}(\eta(p, u)) = u^\mu \left(\frac{\partial}{\partial x^\mu} - p^c{}_d g^d{}_b \Gamma^a{}_{c\mu}(x) \frac{\partial}{\partial p'^a{}_b}\right) = u^\mu \left(\frac{\partial}{\partial x^\mu} - p'^c{}_b \Gamma^a{}_{c\mu}(x) \frac{\partial}{\partial p'^a{}_b}\right) = \eta(p', u) \tag{27.1.26}$$

for $p' = p \cdot g$.

From the horizontal lift one now finds that a basis of TP which realizes the split $VP \oplus HP$ is given by the basis vector fields

$$\delta_\mu = \frac{\partial}{\partial x^\mu} - p^c{}_b \Gamma^a{}_{c\mu}(x) \frac{\partial}{\partial p^a{}_b} , \quad \bar{\partial}_a{}^b = \frac{\partial}{\partial p^a{}_b} , \tag{27.1.27}$$

while the dual basis of T^*P reads

$$\mathrm{d}x^\mu , \quad \delta p^a{}_b = \mathrm{d}p^a{}_b + p^d{}_b \Gamma^a{}_{d\mu}(x) \mathrm{d}x^\mu , \tag{27.1.28}$$

as can most easily seen from the coordinate expression for θ.

We see from this example that the connection is fully determined by the coefficients $\Gamma^a{}_{b\mu}(x)$. Obviously, these depend on the choice of the local trivialization ϕ and the coordinates (x^μ) on the base manifold. If we express the base point $x \in M$ in new coordinates $(\tilde{x}^\mu)$, the coordinate expression of θ changes to

$$\begin{aligned} \theta_p &= \frac{\partial}{\partial p^a{}_b} \otimes \left[\mathrm{d}p^a{}_b + p^d{}_b \Gamma^a{}_{d\mu}(x) \mathrm{d}x^\mu\right] \\ &= \frac{\partial}{\partial p^a{}_b} \otimes \left[\mathrm{d}p^a{}_b + p^d{}_b \Gamma^a{}_{d\nu}(x) \frac{\partial x^\nu}{\partial \tilde{x}^\mu} \mathrm{d}\tilde{x}^\mu\right] \\ &= \frac{\partial}{\partial p^a{}_b} \otimes \left[\mathrm{d}p^a{}_b + p^d{}_b \tilde{\Gamma}^a{}_{d\mu}(x) \mathrm{d}\tilde{x}^\mu\right] , \end{aligned} \tag{27.1.29}$$

and so the connection coefficients transform like the components of a one-form,

$$\tilde{\Gamma}^a{}_{d\mu}(x) = \Gamma^a{}_{d\nu}(x) \frac{\partial x^\nu}{\partial \tilde{x}^\mu} . \tag{27.1.30}$$

For the change of the local trivialization, we consider another trivialization $\tilde{\phi} : \pi^{-1}(U) \to U \times G$ defined such that $p = \tilde{\phi}^{-1}(x, g) = \tilde{\phi}^{-1}(x, e) \cdot \tilde{g}$, and we use again the same coordinates on G. To relate this trivialization to the previously used trivialization ϕ, we define a map $h : U \to G$ such that

$$\tilde{\phi}^{-1}(x, e) \cdot h(x) = \phi^{-1}(x, e) . \tag{27.1.31}$$

Note that such a unique element $h(x)$ always exists, since G acts transitively and freely on the fibers of P, and that this definition depends only on x. For a general element $p \in \pi^{-1}(U)$, we then have

$$\tilde{\phi}^{-1}(x, e) \cdot \tilde{g} = p = \phi^{-1}(x, e) \cdot g = (\tilde{\phi}^{-1}(x, e) \cdot h(x)) \cdot g = \tilde{\phi}^{-1}(x, e) \cdot (h(x)g) , \tag{27.1.32}$$

and so we have $\tilde{g} = h(x)g$. Writing this function in components $h^a{}_b(x)$, we obtain the new coordinates

$$(\tilde{x}^\mu, \tilde{p}^a{}_b) = (x^\mu, h^a{}_c(x) p^c{}_b) \tag{27.1.33}$$

on $\pi^{-1}(U)$. Hence, we have

$$\mathrm{d}\tilde{x}^\mu = \mathrm{d}x^\mu , \quad \mathrm{d}\tilde{p}^a{}_b = h^a{}_c(x) \mathrm{d}p^c{}_b + \partial_\mu h^a{}_c(x) p^c{}_b \mathrm{d}x^\mu , \tag{27.1.34}$$

as well as

$$\frac{\partial}{\partial x^\mu} = \frac{\partial}{\partial \tilde{x}^\mu} + \partial_\mu h^a{}_c(x) p^c{}_b \frac{\partial}{\partial \tilde{p}^a{}_b} , \quad \frac{\partial}{\partial p^a{}_b} = h^c{}_a(x) \frac{\partial}{\partial \tilde{p}^c{}_b} . \tag{27.1.35}$$

From these we extract the new coordinate expression for θ as

$$\begin{aligned} \theta_p &= \frac{\partial}{\partial p^a{}_b} \otimes \left[\mathrm{d}p^a{}_b + p^d{}_b \Gamma^a{}_{d\mu}(x) \mathrm{d}x^\mu\right] \\ &= h^c{}_a(x) \frac{\partial}{\partial \tilde{p}^c{}_b} \otimes \left[(h^{-1})^a{}_d(x) \left(\mathrm{d}\tilde{p}^d{}_b - p^e{}_b \partial_\mu h^d{}_e(x) \mathrm{d}\tilde{x}^\mu\right) + p^d{}_b \Gamma^a{}_{d\mu}(x) \mathrm{d}\tilde{x}^\mu\right] \\ &= \frac{\partial}{\partial \tilde{p}^c{}_b} \otimes \left[\mathrm{d}\tilde{p}^c{}_b - p^e{}_b \partial_\mu h^c{}_e(x) \mathrm{d}\tilde{x}^\mu + h^c{}_a(x) p^d{}_b \Gamma^a{}_{d\mu}(x) \mathrm{d}\tilde{x}^\mu\right] \\ &= \frac{\partial}{\partial \tilde{p}^c{}_b} \otimes \left[\mathrm{d}\tilde{p}^c{}_b - (h^{-1})^e{}_d(x) \tilde{p}^d{}_b \partial_\mu h^c{}_e(x) \mathrm{d}\tilde{x}^\mu + h^c{}_a(x) (h^{-1})^d{}_e(x) \tilde{p}^e{}_b \Gamma^a{}_{d\mu}(x) \mathrm{d}\tilde{x}^\mu\right] \\ &= \frac{\partial}{\partial \tilde{p}^a{}_b} \otimes \left\{\mathrm{d}\tilde{p}^a{}_b + \tilde{p}^c{}_b \left[h^a{}_e(x) (h^{-1})^d{}_c(x) \Gamma^e{}_{d\mu}(x) - (h^{-1})^d{}_c(x) \partial_\mu h^a{}_d(x)\right] \mathrm{d}\tilde{x}^\mu\right\} \\ &= \frac{\partial}{\partial \tilde{p}^a{}_b} \otimes \left[\mathrm{d}\tilde{p}^a{}_b + \tilde{p}^c{}_b \tilde{\Gamma}^a{}_{c\mu}(x) \mathrm{d}\tilde{x}^\mu\right] . \end{aligned} \tag{27.1.36}$$

Hence, we see that the connection coefficients transform as

$$\tilde{\Gamma}^a{}_{c\mu} = h^a{}_e(h^{-1})^d{}_c\Gamma^e{}_{d\mu} - (h^{-1})^d{}_c\partial_\mu h^a{}_d \tag{27.1.37}$$

under a change of the local trivialization.

27.2 Exterior covariant derivative

The fact that a principal connection can be expressed in terms of a Lie algebra valued differential form allows to construct another operation on differential forms which take values in a representation space of the structure group, or its Lie algebra. For this purpose, we start by defining the space of differential forms we will act upon, as follows.

Definition 27.2.1 (Basic form)**.** Let $\pi : P \to M$ be a principal bundle with structure group G and $\rho : G \times F \to F$ a linear representation of G on a vector space F. A differential k-form $\alpha \in \Omega^k(P, F)$ is called *basic of type* ρ if:

1. α is horizontal, i.e., $\iota_X\alpha = 0$ for any vertical vector field X,
2. α is equivariant, i.e., $R_g^*\alpha = \rho_{g^{-1}} \circ \alpha$.

The space of basic forms of type ρ is denoted $\Omega^k_\rho(P, F)$.

The fact that basic forms are equivariant, and thus uniquely defined along the fibers of P by their value at any single point on each fiber, suggests that basic forms are closely related to certain vector valued forms on the base manifold M. Here we show that this is indeed the case.

Theorem 27.2.1. *There exists a one-to-one correspondence between basic forms $\alpha \in \Omega^k_\rho(P, F)$ and $P \times_\rho F$ valued k-forms $\sigma \in \Omega^k(M, P \times_\rho F)$ on M.*

Proof. Let $\sigma \in \Omega^k(M, P \times_\rho F)$. Then we define

$$\begin{array}{cccc} \alpha & : & P & \to & \Lambda^k T^*P \otimes F \\ & & p & \mapsto & [p]^{-1} \circ (\pi^*\sigma)_p \end{array}, \tag{27.2.1}$$

where $[p] : F \to P_{\pi(p)} \times_\rho F$ is the fiber diffeomorphism from definition 20.3.2. We see that α is horizontal, which follows from the fact that $\pi^*\sigma$ is horizontal by construction. To check that α is also equivariant, we calculate

$$\begin{aligned} (R_g^*\alpha)_p(X_1, \ldots, X_k) &= \alpha_{p\cdot g}(R_{g*} \cdot X_1, \ldots, R_{g*} \circ X_k) \\ &= [p \circ g]^{-1} \circ (\pi^*\sigma)_{p\cdot g}(R_{g*} \cdot X_1, \ldots, R_{g*} \circ X_k) \\ &= \rho_{g^{-1}} \circ [p]^{-1} \circ (\pi^*\sigma)_p(X_1, \ldots, X_k) \\ &= (\rho_{g^{-1}} \circ \alpha_p)(X_1, \ldots, X_k)\,. \end{aligned} \tag{27.2.2}$$

Finally, α is a smooth differential form, since it is constructed by a composition of smooth maps. Hence, α is a basic form of type ρ. Finally, to show that this mapping is bijective, note that for any $\alpha \in \Omega^k_\rho(P, F)$ we can define

$$\sigma_{\pi(p)}(\pi_*(v_1), \ldots, \pi_*(v_k)) = [p, \alpha_p(v_1, \ldots, v_k)] \tag{27.2.3}$$

for all $p \in P$ and $v_1, \ldots, v_k \in T_pP$. Note that the right hand side depends only on the pushforward $\pi_*(v_i)$ of the vectors v_i, since any element of the kernel of π_* is vertical by definition, and α is horizontal. Further, it does not depend on the choice of p, but only on the base point $\pi(p)$, as a consequence of the equivariance of α. It follows that $\sigma \in \Omega^k(M, P \times_\rho F)$. One easily checks that this establishes a one-to-one correspondence. ■

Note that we have already encountered the particular case $k = 0$ of the statement above in theorem 20.3.3. With the statement proven above, we have generalized this relation to basic forms on the total space of the principal bundle and to vector-valued forms on the base manifold, which are of arbitrary rank.

One may wonder whether the space of basic forms of type ρ is preserved under the exterior derivative, i.e., whether the exterior derivative of a basic form is again basic. First, note that the usual exterior derivative of an equivariant form of type ρ is again equivariant of type ρ, since the exterior derivative commutes with the pullback, and so we have

$$R_g^*\mathrm{d}\alpha = \mathrm{d}R_g^*\alpha = \mathrm{d}(\rho_{g^{-1}} \circ \alpha) = \rho_{g^{-1}} \circ \mathrm{d}\alpha\,, \tag{27.2.4}$$

where the last equality holds, since the exterior derivative acts only on the differential form part, while $\rho_{g^{-1}}$ acts only on the vector space part of α. However, if α is horizontal, then $\mathrm{d}\alpha$, in general, is not horizontal. To obtain a horizontal form, we may use the connection, which defines a projector onto the space of horizontal forms. Hence, we define as follows.

Definition 27.2.2 (Exterior covariant derivative)**.** Let $\alpha \in \Omega^k_\rho(P, F)$ be a k-form on the principal bundle $\pi : P \to M$ with values in F and HP the horizontal distribution of a principal G-connection ϑ on P. The *exterior covariant derivative* of α with respect to ϑ is the F-valued $k+1$-form $\mathrm{d}_\vartheta\alpha \in \Omega^{k+1}_\rho(P, F)$ defined such that

$$\mathrm{d}_\vartheta\alpha(X_0, \ldots, X_k) = \mathrm{d}\alpha((X_0)_H, \ldots, (X_k)_H) \tag{27.2.5}$$

for all vector fields $X_0, \ldots, X_k \in \mathrm{Vect}(P)$.

Obviously, $\mathrm{d}_\vartheta\alpha$ is horizontal, i.e., it vanishes on any vertical vector field, $\iota_X\alpha = 0$ for any vertical vector field X. We still need to check whether the exterior covariant derivative also maps equivariant forms to equivariant forms. We show this as follows.

Theorem 27.2.2. *The exterior covariant derivative of a basic form is again basic.*

Proof. It follows from the definition equation (27.2.5) that $\mathrm{d}_\vartheta\alpha$ vanishes on any vertical vector, and is therefore horizontal. To check the equivariance, we calculate

$$\begin{aligned}
(R_g^*\mathrm{d}_\vartheta\alpha)(X_0, \ldots, X_k) &= \mathrm{d}_\vartheta\alpha(R_{g*} \circ X_0, \ldots, R_{g*} \circ X_k)\\
&= \mathrm{d}\alpha((R_{g*} \circ X_0)_H, \ldots, (R_{g*} \circ X_k)_H)\\
&= \mathrm{d}\alpha(R_{g*} \circ (X_0)_H, \ldots, R_{g*} \circ (X_k)_H)\\
&= (R_g^*\mathrm{d}\alpha)((X_0)_H, \ldots, (X_k)_H)\\
&= (\mathrm{d}R_g^*\alpha)((X_0)_H, \ldots, (X_k)_H)\\
&= \mathrm{d}(\rho_{g^{-1}} \circ \alpha)((X_0)_H, \ldots, (X_k)_H)\\
&= (\rho_{g^{-1}} \circ \mathrm{d}\alpha)((X_0)_H, \ldots, (X_k)_H)\\
&= (\rho_{g^{-1}} \circ \mathrm{d}_\vartheta\alpha)(X_0, \ldots, X_k)\,,
\end{aligned} \tag{27.2.6}$$

and hence

$$R_g^*\mathrm{d}_\vartheta\alpha = \rho_{g^{-1}} \circ \mathrm{d}_\vartheta\alpha\,. \tag{27.2.7}$$

■

For practical purposes, it is more useful to have a formula for the exterior covariant derivative in terms of the principal G-connection $\vartheta \in \Omega^1(P, \mathfrak{g})$ instead of the horizontal distribution and corresponding projection. Indeed it is possible to find such an expression. We show that it has a simple form.

Theorem 27.2.3. *The exterior covariant derivative of a basic form of type ρ is given by*

$$\mathrm{d}_\vartheta\alpha = \mathrm{d}\alpha + \rho_*(\vartheta) \wedge \alpha\,, \tag{27.2.8}$$

where $\rho_* : \mathfrak{g} \to \mathfrak{gl}(F)$ is the representation of the Lie algebra $\mathfrak{g}$ on F induced by $\rho : G \to \mathrm{GL}(F)$.

Proof. By construction, the exterior covariant derivative $\mathrm{d}_\vartheta\alpha$ is the unique F-valued $k+1$-form which agrees with $\mathrm{d}\alpha$ on horizontal vectors and vanishes on any vertical vector. To check the former, recall that ϑ vanishes on horizontal vectors, $\vartheta(X_H) = 0$, and since ρ_* is linear, also $\rho_*(\vartheta(X_H)) = 0$. Hence, the second term on the right hand side vanishes if it is contracted with $k+1$ horizontal vectors, and only $\mathrm{d}\alpha$ remains. Further, to check that the right hand side vanishes for any vertical vector, let $X \in \mathfrak{g}$ and consider the fundamental vector field $\tilde{X}$. Then we have

$$\begin{aligned}\iota_{\tilde{X}}\mathrm{d}\alpha &= \iota_{\tilde{X}}\mathrm{d}\alpha + \mathrm{d}\iota_{\tilde{X}}\alpha\\ &= \mathcal{L}_{\tilde{X}}\alpha\\ &= -\rho_*(\vartheta) \wedge \alpha\,.\end{aligned} \tag{27.2.9}$$

Since the fundamental vector fields span the vertical tangent space, this holds for all vertical vectors. Hence, the left and right hand sides agree. ■

27.3 Curvature

In section 26.10 we have discussed the curvature of a connection on an arbitrary fiber bundle, given in terms of the Lie bracket of horizontal vector fields, and we have seen that it can also be expressed in terms of the Nijenhuis tensor of the connection form. We now aim to find a similar description for the curvature of principal connections, which constitute a special case of the more general case we discussed before. Recall that a principal G-connection is a one-form ϑ with values in a Lie algebra, and hence vector space $\mathfrak{g}$. It appears thus natural to apply the exterior covariant derivative d_ϑ from definition 27.2.2. We thus arrive at the following definition.

Definition 27.3.1 (Curvature form of a principal connection)**.** Let $\pi : P \to M$ be a principal bundle with structure group G and $\vartheta \in \Omega^1(P, \mathfrak{g})$ a principal G-connection on P. Its *curvature form* is the Lie algebra valued two-form

$$\Omega = \mathrm{d}_\vartheta\vartheta \in \Omega^2(P, \mathfrak{g})\,. \tag{27.3.1}$$

To see how this is related to the notion of curvature we previously defined, we prove the following.

Theorem 27.3.1. *The curvature form Ω of a principal G-connection ϑ is related to the curvature $R = -[\![\theta, \theta]\!]/2$ of the corresponding connection form θ by*

$$R(X, Y) = \widetilde{\Omega(X, Y)} \tag{27.3.2}$$

for all $X, Y \in \mathrm{Vect}(P)$.

Proof. By definition of the exterior covariant derivative one has

$$\begin{aligned}\Omega(X,Y) &= \mathrm{d}_\vartheta\vartheta(X,Y)\\ &= \mathrm{d}\vartheta(X_H,Y_H)\\ &= X_H(\vartheta(Y_H)) - Y_H(\vartheta(X_H)) - \vartheta([X_H,Y_H])\\ &= -\vartheta([X_H,Y_H])\,,\end{aligned} \tag{27.3.3}$$

where we further used the relation (9.4.5) and the fact that ϑ vanishes on horizontal vector fields. Now recall that the curvature of a general connection is equivalently defined as

$$R(X,Y) = -[X_H,Y_H]_V = -\widetilde{\vartheta([X_H,Y_H])} = \widetilde{\Omega(X,Y)} \tag{27.3.4}$$

in terms of the horizontal and vertical projectors, and we expressed the vertical part of $[X_H,Y_H]$ through the connection form and the fundamental vector fields. ■

Before we derive an alternative formula for the connection form, we prove an important property. We have see that for basic forms, the exterior covariant derivative is again basic. We cannot apply this argument here, since the connection form ϑ itself is *not* a basic form, as it is not horizontal, $\iota_X\vartheta \neq 0$ for vertical vector fields X. To show that Ω is nevertheless basic, we prove the following statement.

Theorem 27.3.2. *The curvature form Ω of a principal G-connection ϑ is a basic form of type* Ad.

Proof. It follows immediately from the definition that the curvature form is horizontal. Further taking into account that the principal G-connection ϑ is equivariant by definition, one can apply the same argument as in the proof of theorem 27.2.2 and calculate

$$\begin{aligned}(R_g^*\mathrm{d}_\vartheta\vartheta)(X_0,\ldots,X_k) &= \mathrm{d}_\vartheta\vartheta(R_{g*}\circ X_0,\ldots,R_{g*}\circ X_k)\\ &= \mathrm{d}\vartheta((R_{g*}\circ X_0)_H,\ldots,(R_{g*}\circ X_k)_H)\\ &= \mathrm{d}\vartheta(R_{g*}\circ (X_0)_H,\ldots,R_{g*}\circ (X_k)_H)\\ &= (R_g^*\mathrm{d}\vartheta)((X_0)_H,\ldots,(X_k)_H)\\ &= (\mathrm{d}R_g^*\vartheta)((X_0)_H,\ldots,(X_k)_H)\\ &= \mathrm{d}(\mathrm{Ad}_{g^{-1}}\circ\vartheta)((X_0)_H,\ldots,(X_k)_H)\\ &= (\mathrm{Ad}_{g^{-1}}\circ\mathrm{d}\vartheta)((X_0)_H,\ldots,(X_k)_H)\\ &= (\mathrm{Ad}_{g^{-1}}\circ\mathrm{d}_\vartheta\vartheta)(X_0,\ldots,X_k)\,,\end{aligned} \tag{27.3.5}$$

showing that Ω is equivariant, and thus basic of type Ad. ■

In order to provide a simple formula also for the curvature form, we prove the following formula.

Theorem 27.3.3. *Let $\pi : P \to M$ be a principal bundle with structure group G and $\vartheta \in \Omega^1(P,\mathfrak{g})$ a principal G-connection on P. Its curvature form $\Omega \in \Omega^2(P,\mathfrak{g})$ satisfies the* structure equation

$$\Omega = \mathrm{d}\vartheta + \frac{1}{2}[\vartheta\wedge\vartheta]\,. \tag{27.3.6}$$

Proof. ►...◄ ■

We further discuss a few properties of the connection form. Recall from theorem 26.10.3 that the connection of a general connection satisfies a particular identity, which holds also for a principal connection, and which we can formulate in terms of the connection form as follows.

Theorem 27.3.4 (Bianchi identity)**.** *The connection form Ω of a principal G-connection ϑ satisfies*

$$\mathrm{d}_\vartheta\Omega = 0\,. \tag{27.3.7}$$

Proof. The exterior derivative of the structure equation takes the form

$$\mathrm{d}\Omega = \mathrm{d}\mathrm{d}\vartheta + \frac{1}{2}\mathrm{d}[\vartheta \wedge \vartheta] = \frac{1}{2}\left([\mathrm{d}\vartheta \wedge \vartheta] - [\vartheta \wedge \mathrm{d}\vartheta]\right) = [\mathrm{d}\vartheta \wedge \vartheta]\,. \tag{27.3.8}$$

The exterior covariant derivative $\mathrm{d}_\vartheta\Omega$ is the horizontal part of $\mathrm{d}\Omega$. However, since ϑ vanishes on horizontal vectors, so does $[\mathrm{d}\vartheta \wedge \vartheta]$, and hence $\mathrm{d}_\vartheta\Omega = 0$. ■

Finally, we discuss how the curvature of a principal connection is related to the exterior covariant derivative of forms with values in a representation vector space. We find that a simple relation holds, which can be stated as follows.

Theorem 27.3.5. *The exterior covariant derivative d_ϑ and the curvature form Ω of a principal G-connection are related by*

$$\mathrm{d}_\vartheta \mathrm{d}_\vartheta \alpha = \rho_*(\Omega) \wedge \alpha \tag{27.3.9}$$

for any basic form α.

Proof. By direct calculation using theorem 27.2.3 we have

$$\begin{aligned}\mathrm{d}_\vartheta \mathrm{d}_\vartheta \alpha &= \mathrm{d}(\mathrm{d}\alpha + \rho_*(\vartheta) \wedge \alpha) + \rho_*(\vartheta) \wedge (\mathrm{d}\alpha + \rho_*(\vartheta) \wedge \alpha)\\ &= \mathrm{d}\rho_*(\vartheta) \wedge \alpha - \rho_*(\vartheta) \wedge \mathrm{d}\alpha + \rho_*(\vartheta) \wedge \mathrm{d}\alpha + \rho_*(\vartheta) \wedge \rho_*(\vartheta) \wedge \alpha\\ &= \rho_*(\mathrm{d}\vartheta) \wedge \alpha + \rho_*(\vartheta) \wedge \rho_*(\vartheta) \wedge \alpha\end{aligned} \tag{27.3.10}$$

For the second term, we use the fact that ρ_* is a Lie algebra homomorphism and that $\mathfrak{gl}(F)$ is a Lie algebra of vector space endomorphisms, whose Lie bracket is the commutator. For vector fields $X, Y \in \mathrm{Vect}(P)$ we thus have

$$\begin{aligned}(\rho_*(\vartheta) \wedge \rho_*(\vartheta))(X, Y) &= \rho_*(\vartheta(X)) \circ \rho_*(\vartheta(Y)) - \rho_*(\vartheta(Y)) \circ \rho_*(\vartheta(X))\\ &= [\rho_*(\vartheta(X)), \rho_*(\vartheta(Y))]\\ &= \rho_*([\vartheta(X), \vartheta(Y)])\\ &= \frac{1}{2}\rho_*([\vartheta \wedge \vartheta])(X, Y)\,.\end{aligned} \tag{27.3.11}$$

Hence, we find

$$\mathrm{d}_\vartheta \mathrm{d}_\vartheta \alpha = \rho_*\left(\mathrm{d}\vartheta + \frac{1}{2}[\vartheta \wedge \vartheta]\right) \wedge \alpha = \rho_*(\Omega) \wedge \alpha\,. \tag{27.3.12}$$

■

27.4 Horizontal lift

In section 26.7 we have seen that given a connection on a fiber bundle, it is always possible to locally lift a curve from the base manifold to a horizontal curve in the total space of the bundle, as guaranteed by theorem 26.7.1. However, as shown in example 26.7.1, this may not be the case globally, i.e., for the whole domain of the curve. Following definition 26.7.2 we call a connection complete if this is possible globally. Turning our view towards principal connections, we find the following remarkable result.

Theorem 27.4.1. *Every principal connection is complete.*

Proof. ▶. . . ◀ ■

Recall that by definition, principal connections are "rigid" in the sense that the horizontal space H_pP on a single point $p \in P$ determines the horizontal spaces on the whole fiber $p \cdot G$ by equivariance. Hence, this property is also inherited by the horizontal lift of curves and vector fields. For the horizontal lift of a curve, this can be formulated as follows.

Theorem 27.4.2. *Let $\pi : P \to M$ be a principal G-bundle equipped with a principal connection, $\gamma : \mathbb{R} \to M$ a curve, $p \in P_{\gamma(0)}$ and $g \in G$. Then the horizontal lifts $\hat{\gamma}, \hat{\gamma}' : \mathbb{R} \to P$ with $\hat{\gamma}(0) = p$ and $\hat{\gamma}'(0) = p' = p \cdot g$ are related by $\hat{\gamma}'(t) = \hat{\gamma}(t) \cdot g$ for all $t \in \mathbb{R}$.*

Proof. Since the horizontal lift of a curve γ is uniquely defined by one point in the total space of the bundle and the condition that it is everywhere horizontal, it is sufficient to check these two properties for the curve $\hat{\gamma}'$, using the fact that $\hat{\gamma}$ is a horizontal lift of γ with $\hat{\gamma}(0) = p$. Clearly, we have

$$\hat{\gamma}'(0) = \hat{\gamma}(0) \cdot g = p \cdot g = p' \,, \tag{27.4.1}$$

and so $\hat{\gamma}'$ passes through p' at 0. Further, to check whether $\hat{\gamma}'$ is horizontal, it is helpful to write $\hat{\gamma}' = R_g \circ \hat{\gamma}$, and thus $\dot{\hat{\gamma}}' = R_{g*} \circ \dot{\hat{\gamma}}$. We can then use the equivariant connection form θ to calculate

$$\theta_{\hat{\gamma}'(t)}(\dot{\hat{\gamma}}'(t)) = \theta_{\hat{\gamma}(t)\cdot g}(R_{g*}(\dot{\hat{\gamma}}(t))) = R_{g*}(\theta_{\hat{\gamma}(t)}(\dot{\hat{\gamma}}(t))) = 0 \,, \tag{27.4.2}$$

using the fact that $\hat{\gamma}$ is horizontal. Hence, also $\hat{\gamma}'$ is horizontal. ■

Recalling that the parallel transport is defined via horizontal curves, we can thus conclude as follows.

Theorem 27.4.3. *Let $\pi : P \to M$ be a principal G-bundle equipped with a principal connection, $\gamma : [a, b] \to M$ a curve and $g \in G$. Then the parallel transport $\mathcal{P}_\gamma : P_{\gamma(a)} \to P_{\gamma(b)}$ is equivariant, i.e., it satisfies $\mathcal{P}_\gamma \circ R_g = R_g \circ \mathcal{P}_\gamma$.*

Proof. For $p \in P_{\gamma(a)}$, let $\hat{\gamma}_p : [a, b] \to P$ denote the horizontal lift of γ with $\hat{\gamma}(a) = p$, so that $\mathcal{P}_\gamma(p) = \hat{\gamma}_p(b)$. Then we have

$$\mathcal{P}_\gamma(p \cdot g) = \hat{\gamma}_{p\cdot g}(b) = \hat{\gamma}_p(b) \cdot g = \mathcal{P}_\gamma(p) \cdot g \tag{27.4.3}$$

for all $g \in G$, and hence $\mathcal{P}_\gamma \circ R_g = R_g \circ \mathcal{P}_\gamma$. ■

An infinitesimal version of the statements above can be written as follows.

Theorem 27.4.4. *Let $\pi : P \to M$ be a principal G-bundle equipped with a principal connection and $X \in \mathrm{Vect}(M)$ a vector field on M. Then its horizontal lift $\hat{X} \in \Gamma(HP)$ is equivariant, i.e.,*

$$\hat{X} \circ R_g = R_{g*} \circ \hat{X} \tag{27.4.4}$$

for all $g \in G$.

Proof. Recall from theorem 26.4.2 that we can write the horizontal lift as $\hat{X}(p) = \eta(p, X(\pi(p)))$ for all $p \in P$, where η denotes the horizontal lift map. Using theorem 27.1.4 we then find

$$\hat{X}(p\cdot g) = \eta(p\cdot g, X(\pi(p\cdot g))) = \eta(R_g(p), X(\pi(p))) = R_{g*}(\eta(p, X(\pi(p)))) = R_{g*}(\hat{X}(p)) \,, \tag{27.4.5}$$

and thus $\hat{X} \circ R_g = R_{g*} \circ \hat{X}$. ■

27.5 Connections on associated bundles

In section 20.3 we have seen that from a principal G-bundle $\pi : P \to M$ and a left action $\rho : G \times F \to F$ one can construct the associated bundle $P \times_\rho F$. We now show that a connection on P also induces a connection on every associated bundle. This can be constructed in different, equivalent ways. Here we start from the interpretation of connections as jet bundle sections. This construction is based on the relation between local sections of the principal bundle and local sections of any associated bundle, which we will use as follows.

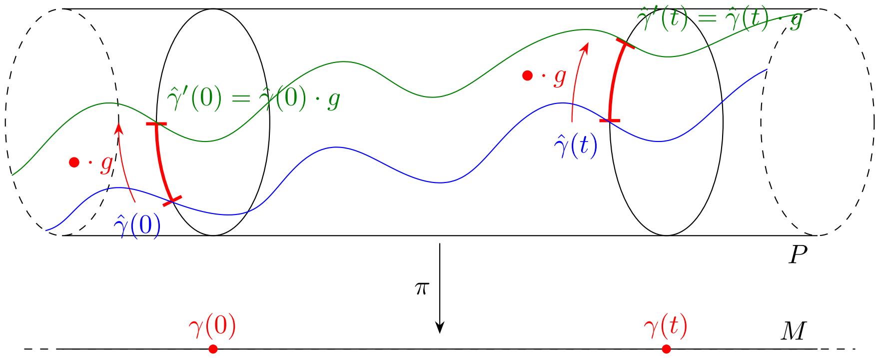


Figure 27.1: Equivariance of the horizontal lift on a principal bundle.

Theorem 27.5.1. *Let $\pi : P \to M$ be a principal G-bundle with Lie group G and $\pi_\rho : P \times_\rho F \to M$ an associated bundle with fiber F. A principal Ehresmann connection $\omega : P \to J^1(P)$ on P, which assigns to $p \in P$ with $\pi(p) = x \in M$ the jet $j^1_x\sigma_p$, induces a connection $\omega_\rho : P \times_\rho F \to J^1(P \times_\rho F)$, which assigns to $[p, f] \in P \times_\rho F$ the jet $j^1_x[\sigma_p, f]$.*

Proof. We first have to check that ω_ρ is well-defined. For this purpose, we have to check that it is independent of the representative (p, f) for $[p, f]$. Given another representative $(p \cdot g, \rho(g^{-1}, f))$ we find that

$$\omega_\rho([p \cdot g, \rho(g^{-1}, f)]) = j^1_x[\sigma_{p\cdot g}, \rho(g^{-1}, f)] = j^1_x[R_g \circ \sigma_p, \rho(g^{-1}, f)] = j^1_x[\sigma_p, f] = \omega_\rho([p, f]) \,, \tag{27.5.1}$$

so that this is indeed satisfied. Here we used the fact that ω is a principal Ehresmann connection, so that a representative $\sigma_{p\cdot g}$ for the jet $\omega(p \cdot g) = j^1_x\sigma_{p\cdot g}$ is given by $R_g \circ \sigma_p$. Further, ω_ρ is a section, since

$$\pi_{\rho\,1,0}(\omega_\rho([p, f])) = [\sigma_p, f](x) = [\sigma_p(x), f] = [p, f] \,. \tag{27.5.2}$$

This shows that ω_ρ is an Ehresmann connection on $P \times_\rho F$. ■

While this definition is most straightforward, it might not be the most intuitive. However, recalling that there are various closely related possibilities to express connections, one can easily obtain other expressions for the connection on $P \times_\rho F$. In the following, we will consider connections as horizontal lift maps.

Theorem 27.5.2. *Let $\pi : P \to M$ be a principal G-bundle with Lie group G, $\pi_\rho : P \times_\rho F \to M$ an associated bundle with fiber F and $\eta : \pi^*TM \to TP$ a principal horizontal lift map, i.e., a vector bundle homomorphism covering the identity on P and satisfying $\pi_* \circ \eta = \mathrm{pr}_2$ and $\eta \circ (R_g, \mathrm{id}_{TM}) = R_{g*} \circ \eta$ for all $g \in G$. Then the map $\eta_\rho : \pi_\rho^*TM \to T(P \times_\rho F)$ defined by $\eta([p, f], v) = \varphi_{\rho*}(\bar{\eta}_\rho((p, f), v))$, where $\varphi_\rho : P \times F \to P \times_\rho F, (p, f) \mapsto [p, f]$ and $\bar{\eta}_\rho : \mathrm{pr}_1^*\pi^*TM \to T(P \times F)$ is the unique map such that $\mathrm{pr}_{1*}(\bar{\eta}_\rho((p, f), v)) = \eta(p, v)$ and $\mathrm{pr}_{2*}(\bar{\eta}_\rho((p, f), v)) = 0$ for all $p \in P$, $f \in F$ and $v \in T_{\pi(p)}M$, is a horizontal lift map.*

Proof. We first check that the map $\bar{\eta}_\rho$ is well-defined. Let $p \in P$ and $f \in F$. Since $P \times F$ is a product manifold, there is a canonical decomposition $T_{(p,f)}(P \times F) \cong T_pP \oplus T_fF$ of the tangent space at (p, f), so that we can identify $w \in T_{(p,f)}(P \times F)$ with the pair $(\mathrm{pr}_{1*}(w), \mathrm{pr}_{2*}(w)) \in T_pP \oplus T_fF$. Hence, the prescription

$$\begin{array}{rccc} \bar{\eta}_\rho & : & \mathrm{pr}_1^*\pi^*TM & \to & T(P \times F) \\ & & ((p, f), v) & \mapsto & (\eta(p, v), 0) \in T_{(p,f)}(P \times F) \end{array} \tag{27.5.3}$$

indeed defines a map. This map is smooth since η is smooth. In order to further define

$$\begin{array}{rccc} \eta_\rho & : & \pi_\rho^*TM & \to & T(P \times_\rho F) \\ & & ([p, f], v) & \mapsto & \varphi_{\rho*}(\bar{\eta}_\rho((p, f), v)) \in T_{[p,f]}(P \times_\rho F) \end{array} , \tag{27.5.4}$$

we further need to check that the expression on the right hand side does not depend on the choice of the representative (p, f) for $[p, f]$. To see this, let $(p', f') = (p \cdot g, \rho(g^{-1}, f))$ for some $g \in G$. Then we have

$$\begin{aligned} \eta_\rho([p', f'], v) &= \varphi_{\rho *}(\bar{\eta}_\rho((p', f'), v)) \\ &= \varphi_{\rho *}(\eta(p', v), 0) \\ &= \varphi_{\rho *}(\eta(p \circ g, v), 0) \\ &= \varphi_{\rho *}(R_{g*}(\eta(p, v)), 0) \\ &= \varphi_{\rho *}(\eta(p, v), 0) \\ &= \eta_\rho([p, f], v)\,, \end{aligned} \tag{27.5.5}$$

where we used the fact that η is the horizontal lift map of a principal connection. Hence, we find that η_ρ is well-defined. Further, it is smooth, since $\bar{\eta}_\rho$ and φ_ρ are smooth. It is a bundle morphism covering the identity on $P \times_\rho F$, since

$$\eta_\rho([p, f], v) \in T_{\varphi_\rho(p,f)}(P \times_\rho F) = T_{[p,f]}(P \times_\rho F)\,. \tag{27.5.6}$$

All involved maps are linear in v, so that η_ρ is linear on every fiber and hence a vector bundle homomorphism. Finally, we have

$$\begin{aligned} \pi_{\rho *}(\eta_\rho([p, f], v)) &= \pi_{\rho *}(\varphi_{\rho *}(\bar{\eta}_\rho((p, f), v))) \\ &= \pi_*(\mathrm{pr}_{1*}(\bar{\eta}_\rho((p, f), v))) \\ &= \pi_*(\eta(p, v)) \\ &= v\,, \end{aligned} \tag{27.5.7}$$

where we used the relations $\pi \circ \mathrm{pr}_1 = \pi_\rho \circ \varphi_\rho$ for the maps defining the associated bundle and the fact that $\pi_* \circ \eta = \mathrm{pr}_2$. Hence, we also have $\pi_{\rho *} \circ \eta_\rho = \mathrm{pr}_2$, showing that η_ρ is a horizontal lift map. ■

We have now found two different prescriptions for constructing a connection on associated bundles from a principal connection, interpreted as a jet bundle section or as a horizontal lift map. Naturally the question arises whether these two constructions are equivalent, i.e., describe the same connection on the associated bundle. We now show that this is indeed the case.

Theorem 27.5.3. *The connections on $P \times_\rho F$ constructed in theorems 27.5.1 and 27.5.2 are identical.*

Proof. Let $p \in P$, $f \in F$, $x = \pi(p)$ and $v \in T_xM$. Given an Ehresmann connection $\omega : P \to J^1(P)$, let $\sigma \in \Gamma_x(P)$ be a representative of the jet $\omega(p) = j^1_x\sigma$. By definition of an Ehresmann connection, we have $\sigma(x) = p$. Further, the Ehresmann connection ω and horizontal lift map $\eta : \pi^*TM \to TP$ are related by $\eta(p, v) = \sigma_*(v) \in T_pP$ following theorem 26.4.2. Similarly, we can use theorem 27.5.1 to construct an Ehresmann connection $\omega_\rho : P \times_\rho F \to J^1(P \times_\rho F)$, which is related to ω such that a representative for $\omega_\rho([p, f]) = j^1_x\Sigma$ is given by $\Sigma : x \mapsto [\sigma(x), f]$. Hence,

$$\Sigma = \varphi_\rho \circ (\mathrm{id}_P, f) \circ \sigma\,. \tag{27.5.8}$$

Its differential is thus given by

$$\Sigma_* = \varphi_{\rho *} \circ (\sigma_*, 0)\,. \tag{27.5.9}$$

Thus, we find that

$$\Sigma_*(v) = \varphi_{\rho *}(\sigma_*(v), 0) = \varphi_{\rho *}(\eta(p, v), 0) = \eta_\rho([p, f], v)\,, \tag{27.5.10}$$

so that ω_ρ and η_ρ define the same connection on $P \times_\rho F$. ■

Using the results obtained so far, it is now straightforward to find also other expressions for the connection on $P \times_\rho F$. We now take a closer look at the most geometric interpretation in terms of horizontal distributions.

Theorem 27.5.4. *Let $\pi : P \to M$ be a principal G-bundle with Lie group G, HP a G-invariant horizontal distribution on P and $\pi_\rho : P \times_\rho F \to M$ an associated bundle with fiber F. Then there exists a unique horizontal distribution $H(P \times_\rho F)$ such that*

$$\varphi_{\rho*}^{-1}(H(P \times_\rho F)) = HP \oplus \ker \varphi_{\rho*} \subset T(P \times F)\,, \tag{27.5.11}$$

where $\varphi_\rho : P \times F \to P \times_\rho F, (p, f) \mapsto [p, f]$ denotes the quotient with respect to the action of G on $P \times F$.

Proof. Let $p \in P$ and $f \in F$. First note that the tangent space $T_{(p,f)}(P \times F)$ decomposes into direct sums

$$T_{(p,f)}(P \times F) = T_pP \oplus T_fF = V_pP \oplus H_pP \oplus T_fF\,, \tag{27.5.12}$$

where the first equality is the canonical split of the tangent bundle over a product manifold and the second equality is due to the connection on P, seen as a horizontal distribution. G acts from the right on $P \times F$ as

$$\begin{array}{rcccc} \varphi & : & (P \times F) \times G & \to & P \times F \\ & & ((p, f), g) & \mapsto & (p \cdot g, \rho(g^{-1}, f)) \end{array} . \tag{27.5.13}$$

Since the equivalence classes $[p, f]$ are the orbits with respect to this action, the kernel $\ker \varphi_{\rho*}$ is given by those tangent vectors which are tangent to the orbits, and hence

$$T_{(p,f)}(P \times F) \cap \ker \varphi_{\rho*} = \{\varphi_{(p,f)*}v, v \in T_eG\}\,. \tag{27.5.14}$$

Note in particular that this action leaves $\pi(p) \in M$ invariant, and so $\ker \varphi_{\rho*} \subseteq VP \oplus TF$, so that in particular $\ker \varphi_{\rho*} \cap HP = \{0\}$. Further, the action of G on P is free, from which follows $\ker \varphi_{\rho*} \cap TF = \{0\}$, and transitive, so that there exists a bijection between V_pP and $\ker \varphi_{\rho*}$. We can thus write the induced split of the tangent spaces as

$$T_{(p,f)}(P \times F) = V_pP \oplus H_pP \oplus T_fF = \ker \varphi_{\rho*} \oplus H_pP \oplus T_fF\,. \tag{27.5.15}$$

This implies that $\varphi_{\rho*} : T_{(p,f)}(P \times F) \to T_{[p,f]}(P \times_\rho F)$ bijectively maps $H_pP \oplus T_fF$ to $T_{[p,f]}(P \times_\rho F)$. Clearly, one finds that the image of T_fF is vertical with respect to π_ρ, so that the image of H_pP defines a horizontal complement. This decomposition is independent of the choice of the representative (p, f), since HP is invariant under the action of G on P. One easily checks that this defines a horizontal distribution on $P \times_\rho F$. ■

One may already expect that the horizontal distribution above defines the same connection as the previously defined notions. This is what we show next.

Theorem 27.5.5. *The horizontal distribution $H(P \times_\rho F)$ defined in theorem 27.5.4 represents the same connection as obtained in theorems 27.5.1 and 27.5.2.*

Proof. ►...◄ ■

The construction is illustrated in figure 27.2. Given a principal connection on P, the tangent space over $p \in P$ splits in the form $T_pP = V_pP \oplus H_pP$ into a canonically defined vertical subspace V_pP and a horizontal subspace H_pP. ►...◄

Another possibility to characterize a connection is in terms of its horizontal curves. It turns out that for an associated bundle, we can explicitly construct the horizontal lift of a curve from that defined on the principal bundle, as we show next.

Theorem 27.5.6. *Let $\pi : P \to M$ be a principal bundle with a principal connection, $\pi_\rho : P \times_\rho F \to M$ an associated bundle and $\gamma : \mathbb{R} \to M$ a curve. For every $(p, f) \in P_{\gamma(0)} \times F$, there is a unique horizontal lift $\hat{\gamma}_\rho : \mathbb{R} \to P \times_\rho F$ of γ such that $\hat{\gamma}_\rho(0) = [p, f]$ given by*

$$\begin{array}{rcccc} \hat{\gamma}_\rho & : & \mathbb{R} & \to & P \times_\rho F \\ & & t & \mapsto & [\hat{\gamma}(t), f] \end{array} , \tag{27.5.16}$$

where $\hat{\gamma} : \mathbb{R} \to P$ is the horizontal lift of γ such that $\hat{\gamma}(0) = p$.

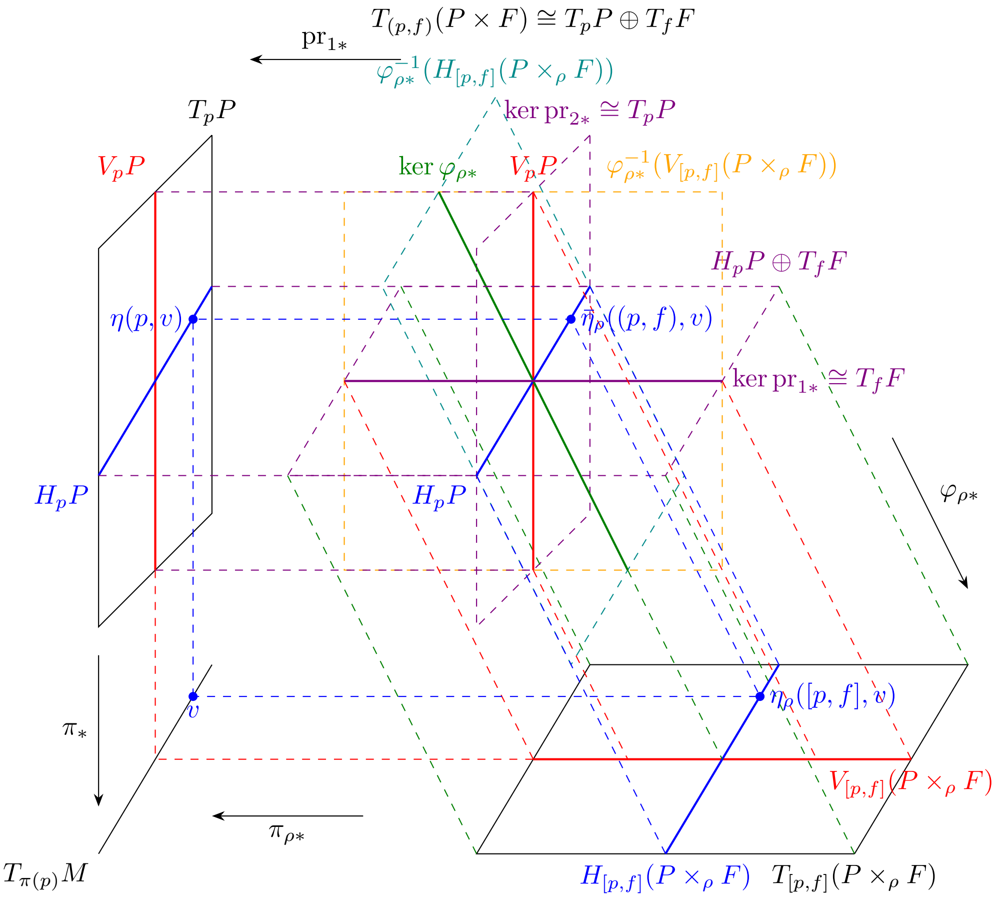


Figure 27.2: Construction of the associated bundle connection.

Proof. It is sufficient to show that $\hat{\gamma}_\rho$ is a horizontal lift; uniqueness follows from theorem 26.7.1. Using the fact that $\hat{\gamma}$ is a horizontal lift through p, one easily checks that

$$\hat{\gamma}_\rho(0) = [\hat{\gamma}(0), f] = [p, f] \tag{27.5.17}$$

and

$$\pi_\rho(\hat{\gamma}_\rho(t)) = \pi_\rho([\hat{\gamma}(t), f]) = \pi(\hat{\gamma}(t)) = \gamma(t)\,. \tag{27.5.18}$$

To show that $\dot{\hat{\gamma}}_\rho(t)$ is horizontal for all $t \in \mathbb{R}$, recall that we can write the horizontal lift map $\eta_\rho : \pi_\rho^* TM \to T(P \times_\rho F)$ as $\eta_\rho = \varphi_{\rho*} \circ (\eta, 0)$, where $\eta : \pi^* TM \to TP$ is the horizontal lift map on P and $\varphi_\rho : P \times F \to P \times_\rho F$ is the orbit projection. Further, we can write

$$\hat{\gamma}_\rho(t) = [\hat{\gamma}(t), f] = \varphi_\rho(\hat{\gamma}(t), f)\,, \tag{27.5.19}$$

and so

$$\dot{\hat{\gamma}}_\rho = \varphi_{\rho*}(\dot{\hat{\gamma}}(t), 0) = \varphi_{\rho*}(\eta(\hat{\gamma}(t), \dot{\gamma}(t)), 0) = \eta_\rho(\hat{\gamma}_\rho(t), \dot{\gamma}(t))\,, \tag{27.5.20}$$

so that $\hat{\gamma}_\rho$ is indeed horizontal. ∎

Now the following conclusion is straightforward.

Theorem 27.5.7. *Every connection on an associated bundle that is induced by a principal connection is complete.*

Proof. In theorem 27.4.1 we have shown that every curve on the base manifold M of a principal bundle $\pi : P \to M$ can be horizontally lifted to P on its entire domain. Further, in theorem 27.5.6 we have shown that this horizontal lift to P allows to construct a horizontal lift to any associated fiber bundle on the same domain. Hence, also the connection on the associated bundle is complete. ■

27.6 Extension of principal connections

In section 20.7 we studied the extension of principal bundles, which allows us to construct a principal G-bundle from a principal H-bundle and a Lie group homomorphism. This construction made us of the notion of associated bundles. In section 27.5 we showed that a principal connection on a principal bundles also gives rise to a connection on each associated bundle. We now explore how these two notions can be combined, and study the properties of the connection, which we define as follows.

Definition 27.6.1 (Extension of a principal connection)**.** Let $\chi : Q \to M$ be a principal fiber bundle with structure group H, $\omega : Q \to J^1(Q)$ a principal connection and $\lambda : H \to G$ a Lie group homomorphism. The connection $\omega_\rho : P \to J^1(P)$ on the λ-extension $\pi : P \to M$ defined by

$$P = Q \times_\rho G\,, \tag{27.6.1}$$

where $\rho : H \times G \to G$ is the left action defined by

$$\begin{array}{rcccc} \rho & : & H \times G & \to & G \\ & & (h, g) & \mapsto & \lambda(h)g \end{array}\,, \tag{27.6.2}$$

is called the λ-*extension* of ω.

In theorem 20.7.1 we have seen that the bundle $\pi : P \to M$ constructed above is a principal bundle. Hence, one may ask whether the connection constructed above is a principal connection. This will be shown next.

Theorem 27.6.1. *The connection ω_ρ given in definition 27.6.1 is a principal connection.*

Proof. Let $p = [q, g] \in P = Q \times_\rho G$ with $q \in Q$, $g \in G$ and $x = \chi(q) \in M$. The induced connection $\omega_\rho : P \to J^1(P)$ assigns to this element the jet $j^1_x[\sigma_q, g]$, where $j^1_x\sigma_q = \omega(q) \in J^1(Q)$. At any other element $p \cdot \tilde{g}$ of the same fiber with $\tilde{g} \in G$ we have

$$\omega_\rho(p \cdot \tilde{g}) = j^1_x[\sigma_q, g\tilde{g}] = j^1_x([\sigma_q, g] \cdot \tilde{g}) = j^1_x([\sigma_q, g] \circ R_{\tilde{g}}) = \omega_\rho(p) \cdot \tilde{g}\,, \tag{27.6.3}$$

and so ω_ρ is equivariant, and thus a principal connection. ■

We have already learned in the previous sections that principal connections are "rigid" in the sense that the specifying the connection at one point of a fiber determines it also at any other point of the same fiber. In the case of a λ-extension, there are points on $P = Q \times_\rho G$ which are of the form $[q, e]$ with $q \in Q$, and so it is sufficient to specify the connection at these points, since any other point along the fibers of P can be reached via the right translation. This can be used to find other characterizations of the extension of a connection. The following statement makes use of the formulation of connections in terms of horizontal lift maps.

Theorem 27.6.2. *The horizontal lift map $\eta_\rho : \pi^*_\rho TM \to T(Q \times_\rho G)$ of the λ-extension in definition 20.7.1 is the unique horizontal lift map related to $\eta : \pi^*TM \to TQ$ by $\eta_\rho(f(q), u) = f_*(\eta(q, u))$ for all $(q, u) \in \pi^*TM$, where $f : Q \to Q \times_\rho G, q \mapsto [q, e]$.*

Proof. Let $q \in Q$ with $\pi(q) = x \in M$ and $u \in T_xM$, and recall that $\eta(q,u) = \sigma_*(u)$, where $\sigma \in \Gamma_x(Q)$ is a representative of the jet $j^1_x\sigma = \omega(q)$. Similarly, the horizontal lift map on $Q \times_\rho Q$ is defined by $\eta_\rho(f(q),u) = \sigma_{\rho*}(u)$, where $\sigma_\rho \in \Gamma_x(Q \times_\rho G$ is a representative of the jet $\omega_\rho(f(q)) = \omega_\rho([q,e]) = j^1_x[\sigma,e]$ by definition 20.7.1; hence, we can set $\sigma_\rho = [\sigma,e] = f \circ \sigma$, so that $\sigma_{\rho*} = f_* \circ \sigma_*$ and thus

$$\eta_\rho(f(q),u) = \sigma_{\rho*}(u) = f_*(\sigma_*(u)) = f_*(\eta(q,u)). \tag{27.6.4}$$

This determines η_ρ on all other elements that are of the form $(f(q),u) = ([q,e],u)$. Since ω_ρ is a principal connection as shown in theorem 27.6.1, the horizontal lift map on all other elements $(f(q)\cdot g,u)$ with $g \in G$ is determined by theorem 27.1.4 as $\eta_\rho(f(q)\cdot g,u) = R_{g*}(\eta_\rho(f(q),u))$, and so η_ρ is uniquely defined. ■

Having constructed the horizontal lift map on the λ-extension, it is now easy to also relate the horizontal distributions, using the fact that these are simply the images of the horizontal lift maps.

Theorem 27.6.3. *The horizontal distributions HQ and $H(Q\times_\rho G)$ are related such that for all $q \in Q$, $w \in T_{f(q)}(Q \times_\rho Q)$ is horizontal if and only if there exists $v \in H_qQ$ such that $w = f_*(v)$, where $f : Q \to Q \times_\rho G, q \mapsto [q,e]$.*

Proof. For each $q \in Q$ with $\pi(q) = x \in M$, the horizontal tangent space is given by

$$H_qQ = \{\eta(q,u), u \in T_xM\}\,. \tag{27.6.5}$$

Hence, an element $v \in T_qQ$ is horizontal if and only if there exists $u \in T_xM$ such that $v = \eta(q,u)$. Similarly, $w \in T_{f(q)}(Q \times_\rho G)$ is horizontal if and only if there exists $u \in T_xM$ such that

$$w = \eta_\rho(f(q),u) = f_*(\eta(q,u))\,, \tag{27.6.6}$$

and thus if and only if there exists $v \in H_qQ$ with $w = f_*(v)$. ■

Finally, we also formulate the relation in terms of principal connection forms, which is the most commonly used formulation of principal connections. In this case, we find the following statement.

Theorem 27.6.4. *The principal G-connection ϑ_ρ on the λ-extension $\chi_\rho : Q \times_\rho G \to M$ is the unique principal G-connection satisfying $f^*\vartheta_\rho = \lambda_* \circ \vartheta$, where $f : Q \to Q \times_\rho G, q \mapsto [q,e]$.*

Proof. Let $q \in Q$. We make use of the fact that $(f^*\vartheta_\rho)_q : T_qQ \to \mathfrak{g}$ and $\lambda_* \circ \vartheta_q : T_qQ \to \mathfrak{g}$ are linear and that $T_qQ = H_qQ \oplus V_qQ$:

1. If $v \in H_qQ$ is horizontal, then by definition we have $\vartheta_q(v) = 0$, and thus also $\lambda_*(\vartheta_q(v)) = 0$. Following theorem 27.6.3, $f_*(v) \in H_{f(q)}(Q \times_\rho G)$ is horizontal, and thus

$$(f^*\vartheta_\rho)_q = \vartheta_{\rho f(q)}(f_*(v)) = 0\,. \tag{27.6.7}$$

2. If $v \in V_qQ$ is vertical, it can be written as $v = \tilde{X}(q)$, where $\tilde{X} \in \mathrm{Vect}(Q)$ is the fundamental vector field of $X = \vartheta_q(v) \in \mathfrak{h}$, and hence $\lambda_*(\vartheta_q(v)) = \lambda_*(X)$. Since f is a bundle morphism covering the identity on M, also $f_*(v) \in T_{f(q)}(Q \times_\rho G)$ is vertical, and so it can be written as $f_*(v) = \tilde{Y}(f(q))$, where $\tilde{Y} \in \mathrm{Vect}(Q \times_\rho G)$ is the fundamental vector field of $Y = \vartheta_{\rho f(q)}(f_*(v)) \in \mathfrak{g}$. Finally, let $\phi : \mathbb{R} \to H$ be the one-parameter group generated by X, so that $\tilde{X}(q)$ is the tangent vector of $t \mapsto q \cdot \phi(t)$ at $t = 0$. Then $f_*(\tilde{X}(q))$ is the tangent vector to the curve

$$t \mapsto f(q \cdot \phi(t)) = [q \cdot \phi(t), e] = [q, \lambda(\phi(t))] \tag{27.6.8}$$

at $t = 0$, where $\lambda \circ \phi : \mathbb{R} \to G$ is the one-parameter group generated by $\lambda_*(X) \in \mathfrak{g}$. Using $f_*(\tilde{X}(q)) = \tilde{Y}(f(q))$ and the uniqueness of the fundamental vector fields, we thus have $\lambda_*(X) = Y$.

By linearity, it follows that

$$(f^*\vartheta_\rho)_q(v) = \lambda_*(\vartheta_q(v)) \tag{27.6.9}$$

for all $v \in T_qQ$. Since this holds for all $q \in Q$, and ϑ_ρ is uniquely determined by its kernel, and hence the horizontal distribution defined by theorem 27.6.3, this uniquely determines ϑ_ρ. ■

27.7 Reduction of principal connections

Definition 27.7.1 (Reduction of a principal connection)**.** Let $\pi : P \to M$ be a principal bundle with structure group G, $\vartheta \in \Omega^1(P, \mathfrak{g})$ a principal G-connection, $\lambda : H \to G$ a Lie group homomorphism and a λ-reduction given by a principal H-bundle $\chi : Q \to M$ and a map $f : Q \to P$. A principal H-connection $\tilde{\vartheta}$ is called λ-reduction if and only if $f^*\vartheta = \lambda_* \circ \tilde{\vartheta}$.

27.8 Holonomy

In section 27.4 we have seen that every principal connection is complete, so that every curve on the base space can be lifted entirely into the total space of the bundle. Further, we have seen that the parallel transport along any curve is an equivariant map. We now take a closer look at the parallel transport along closed paths. This allows us to define the following notion.

Definition 27.8.1 (Holonomy)**.** Let $\pi : P \to M$ be a principal G-bundle equipped with a principal connection, $x \in M$, $p \in P_x$ and $\gamma : [0, 1] \to M$ with $\gamma(0) = \gamma(1) = x$ a closed curve. The unique element $\mathrm{hol}^\omega_p(\gamma) \in G$ for which

$$\mathcal{P}_\gamma(p) = p \cdot \mathrm{hol}^\omega_p(\gamma)\,, \tag{27.8.1}$$

is called the *holonomy* of γ with respect to p.

Note that we have included both the curve γ and the initial point p in the notation of the holonomy. From section 27.4 we know that the horizontal lift, and thus the parallel transport, on principal bundles is very "rigid" due to the equivariance of the connection, and so we expect that also the dependence of the holonomy on the initial point is fixed by equivariance. We now show that this indeed the case.

Theorem 27.8.1. *For any point $p \in P$, curve $\gamma : [0, 1] \to M$ with $\gamma(0) = \gamma(1) = \pi(p)$ and $g \in G$, the holonomy satisfies*

$$\mathrm{hol}^\omega_{p\cdot g}(\gamma) = g^{-1}\,\mathrm{hol}^\omega_p(\gamma) g\,. \tag{27.8.2}$$

Proof. From the equivariance of the parallel transport follows

$$p \cdot g \cdot \mathrm{hol}^\omega_{p\cdot g}(\gamma) = \mathcal{P}_\gamma(p \cdot g) = \mathcal{P}_\gamma(p) \cdot g = p \cdot \mathrm{hol}^\omega_p(\gamma) \cdot g\,. \tag{27.8.3}$$

■

The notion of holonomy is illustrated in figure 27.3.

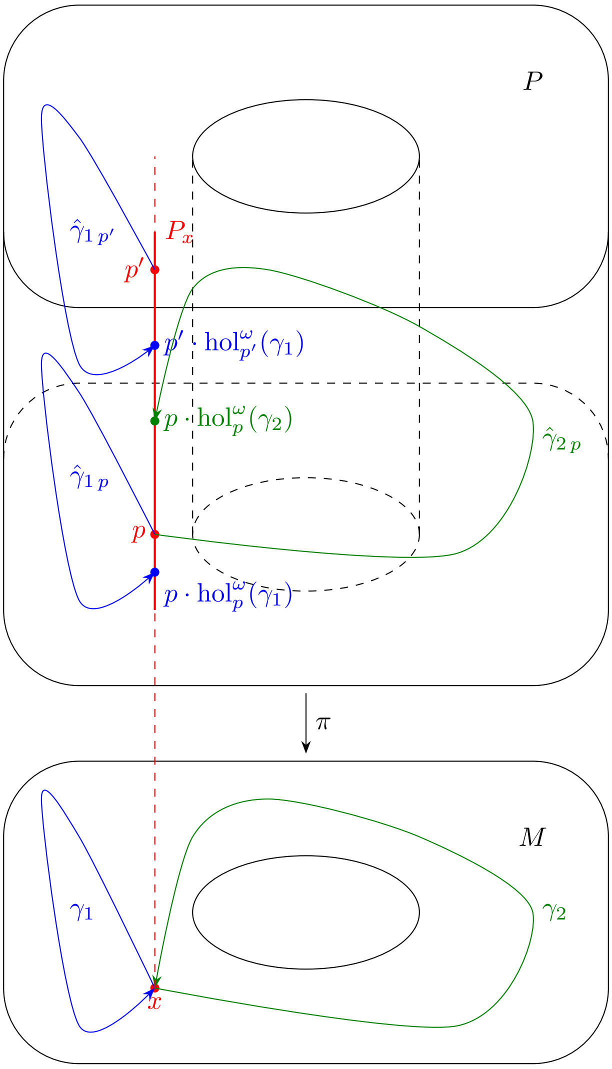


Figure 27.3: Illustration of holonomy for different curves on the base manifold and initial points in the fiber over the common base point.

Definition 27.8.2 (Holonomy group)**.** Let $\pi : P \to M$ be a principal bundle equipped with a principal connection ω. For $p \in P$, the *holonomy group* is given by

$$\mathrm{Hol}_p(\omega) = \{\mathrm{hol}^{\omega}_p(\gamma), \gamma \in \mathcal{C}_{\pi(p)}(M)\}\,, \tag{27.8.4}$$

where $\mathcal{C}_x(M)$ for $x \in M$ denotes the set of curves $\gamma : [0,1] \to M$ with $\gamma(0) = \gamma(1) = x$, while its *reduced holonomy group* is

$$\mathrm{Hol}^0_p(\omega) = \{\mathrm{hol}^{\omega}_p(\gamma), \gamma \in \mathcal{C}^0_{\pi(p)}(M)\}\,, \tag{27.8.5}$$

where $\mathcal{C}^0_x(M)$ for $x \in M$ denotes the set of zero-homotopic curves $\gamma : [0,1] \to M$ with $\gamma(0) = \gamma(1) = x$.

Theorem 27.8.2. *For each point $p \in P$ of a principal G-bundle $\pi : P \to M$ with connection ω, the holonomy group $\mathrm{Hol}_p(\omega)$ is a normal Lie subgroup of G and $\mathrm{Hol}^0_p(\omega)$ is its connected* *component containing the unit element.*

Proof. ▶. . . ◀ ■

Theorem 27.8.3. *For each point $p \in P$ of a principal G-bundle $\pi : P \to M$ with connection ω and $v, w \in T_pP$, the curvature satisfies*

$$\Omega_p(v, w) \in \mathrm{Lie}(\mathrm{Hol}_p(\omega)) \,. \tag{27.8.6}$$

Proof. ▶. . . ◀ ■

Chapter 28

Linear connections

28.1 Connections on vector bundles

An important special case we consider is that of connections on a vector bundle $\pi : E \to M$. In this case every fiber of the bundle carries the structure of a vector space. As we have shown in theorem 21.6.6, also the jet bundles $\pi_r : J^r(E) \to M$ are vector bundles, and the projections $\pi_{r,0} : J^r(E) \to E$ are vector bundle homomorphisms covering the identity on M. This allows us to define the following notion.

Definition 28.1.1 (Linear Ehresmann connection). Let $\pi : E \to M$ be a vector bundle. A *linear Ehresmann connection* on E is a vector bundle homomorphism $\omega : E \to J^1(E)$ such that $\pi_{1,0} \circ \omega = \mathrm{id}_E$.

This definition essentially consists of two parts. Being a map $\omega : E \to J^1(E)$ with $\pi_{1,0} \circ \omega = \mathrm{id}_E$ means that a linear Ehresmann connection is a section of the bundle $\pi_{1,0} : J^1(E) \to E$, and thus an Ehresmann connection. In addition, the restrictions $\omega|_x : E_x \to J^1_x(E)$ must be vector space homomorphisms for all $x \in M$.

To illustrate this definition, let (x^μ, y^a) be local coordinates on E as in the previous section, where in addition we demand that the coordinates (y^a) on the fiber space F correspond to a basis $(e_1, \ldots, e_f)$ of F, where $f = \dim F$ and $y = y^a e_a$. Recall that a general Ehresmann connection on a fiber bundle is uniquely determined by a set $y^a_\mu(x, y)$ of coordinate functions. For a linear Ehresmann connection these must be of the form $y^a_\mu(x, y) = y^a{}_{b\mu}(x) y^b$.

28.2 Koszul connections

On vector bundles one conventionally uses a different description for connections, which is given as follows.

Definition 28.2.1 (Koszul connection). Let $\pi : E \to M$ be a vector bundle. A *Koszul connection* on E is an $\mathbb{R}$-linear function $\nabla : \Gamma(E) \to \Gamma(T^*M \otimes E)$ such that $\nabla(f\epsilon) = f\nabla\epsilon + \mathrm{d}f \otimes \epsilon$ for all $\epsilon \in \Gamma(E)$ and $f \in C^\infty(M, \mathbb{R})$.

We illustrate this definition using the same coordinates as above. In these coordinates a section $\epsilon \in \Gamma(E)$ is expressed in the form $y = y^a e_a$, where (e_a) is a basis of E and y^a are smooth functions. A Koszul connection assigns to ϵ a section $\nabla\epsilon \in \Gamma(T^*M \otimes E)$, whose coordinate expression follows from the Leibniz rule, which states that

$$\nabla\epsilon = \partial_\mu y^a \mathrm{d}x^\mu \otimes e_a + y^a \nabla e_a \,. \tag{28.2.1}$$

We can express ∇e_a in the basis $\mathrm{d}x^\mu \otimes e_a$ in the form $\nabla e_a = \omega^b{}_{a\mu} \mathrm{d}x^\mu \otimes e_b$ to finally obtain

$$\begin{aligned} \nabla\epsilon &= \left[\partial_\mu y^a + \omega^a{}_{b\mu} y^b\right] \mathrm{d}x^\mu \otimes e_a \\ &= \nabla_\mu y^a \mathrm{d}x^\mu \otimes e_a \\ &= y^a{}_{;\mu} \mathrm{d}x^\mu \otimes e_a \,, \end{aligned} \tag{28.2.2}$$

where we introduced the *semicolon notation*.

Koszul connections are very helpful as they can be used to define a number of operations on vector bundles. Although they are defined rather differently from Ehresmann connections, they are closely related. One may already get this impression from the coordinate expressions $y^a{}_{b\mu}$ and $\omega^a{}_{b\mu}$, which carry the same index structure. More formally, we formulate it as follows.

Theorem 28.2.1. *For every vector bundle $\pi : E \to M$ there is a one-to-one correspondence between linear Ehresmann connections and Koszul connections on E.*

Proof. Let $\epsilon : M \to E$ be a section. For $x \in M$ it defines a point $e = \epsilon(x) \in E$ and a linear map $\epsilon_*|_x : T_xM \to T_eE$ with $\pi_* \circ \epsilon_*|_x = \mathrm{id}_{T_xM}$. Also a linear Ehresmann connection $\omega : E \to J^1(E)$ defines a linear map $\sigma_{e*}|_x : T_xM \to T_eE$ with $\pi_* \circ \sigma_{e*}|_x = \mathrm{id}_{T_xM}$ through the jet $\omega(e) = j^1_x\sigma_e$. Their difference $\nabla^\omega\epsilon|_x = \epsilon_*|_x - \sigma_{e*}|_x$ therefore defines a linear map from T_xM to V_eE. Hence,

$$\nabla^\omega\epsilon|_x \in \mathrm{Hom}(T_xM, V_eE) \cong T^*_xM \otimes V_eE \cong T^*_xM \otimes E_x \,. \tag{28.2.3}$$

Doing this for each $x \in M$ we obtain a section $\nabla^\omega\epsilon \in \Gamma(T^*M \otimes E)$. The smoothness of this section can be proven using the smoothness of ω and ϵ. Further, given a function $f \in C^\infty(M, \mathbb{R})$ we find that

$$\begin{aligned} \nabla^\omega(f\epsilon)|_x &= (f\epsilon)_*|_x - \sigma_{fe*}|_x \\ &= f\epsilon_*|_x + (\mathrm{d}f \otimes \epsilon)|_x - f\sigma_{e*}|_x \\ &= (f\nabla^\omega\epsilon)|_x + (\mathrm{d}f \otimes \epsilon)|_x \,. \end{aligned} \tag{28.2.4}$$

This shows that ∇^ω satisfies the Leibniz rule and hence is a Koszul connection.

We will not prove the converse direction, but simply provide the construction how to obtain a linear Ehresmann connection from a Koszul connection $\nabla : \Gamma(E) \to \Gamma(T^*M \otimes E)$. For $e \in E$ with $\pi(e) = x \in M$ choose a section $\epsilon \in \Gamma(E)$ such that $\epsilon(x) = e$. Then $\epsilon^\nabla_x = \epsilon_*|_x - \nabla\epsilon|_x : T_xM \to T_eE$ is a linear map, which we can take as an ingredient to construct a section $\omega^\nabla : E \to J^1(E)$ as described in the previous section. Of course, to complete the proof one still needs to show that this is independent of the choice of the section ϵ. ∎

From the construction above one can derive how the coordinate expressions $y^a{}_{b\mu}$ and $\omega^a{}_{b\mu}$ we introduced earlier are related. A quick calculation shows that similarly to the case of general Ehresmann connections we have $y^a{}_{b\mu} = -\omega^a{}_{b\mu}$.

28.3 Affine bundle of connections

Recall from section 26.3 that the bundle $\pi_{1,0} : J^1(E) \to E$ is an affine bundle, and so we could consider connections as sections of this bundle over E. In the case of linear Ehresmann connections on vector bundles $\pi : E \to M$, it is instructive to derive yet another interpretation as sections of an affine bundle over the base manifold M instead of the total space E. To see this, we start with the following observation formulated in the language of Koszul connections.

Theorem 28.3.1. *Let $\pi : E \to M$ be a vector bundle and $\nabla : \Gamma(E) \to \Gamma(T^*M \otimes E)$ a Koszul connection. Then the following holds:*

1. *For every endomorphism-valued covector field $K \in \Gamma(T^*M \otimes \mathrm{End}(E))$ there exists a unique Koszul connection $\nabla' : \Gamma(E) \to \Gamma(T^*M \otimes E)$, and*
2. *for every Koszul connection $\nabla' : \Gamma(E) \to \Gamma(T^*M \otimes E)$ there exists a unique endomorphism-valued covector field $K \in \Gamma(T^*M \otimes \mathrm{End}(E))$,*

such that

$$\nabla'\epsilon = \nabla\epsilon + K\epsilon \,. \tag{28.3.1}$$

Proof. 1. It follows immediately from the definition of K that $K\epsilon \in \Gamma(T^*M \otimes E)$, so that the expression above is well-defined. Further, since ∇ is a Koszul connection and K is a tensor field, both terms on the right hand side, and thus also the right hand side altogether, are $\mathbb{R}$-linear. Finally, given a function $f \in C^\infty(M, \mathbb{R})$ we find

$$\nabla'(f\epsilon) = \nabla(f\epsilon) + K(f\epsilon) = f\nabla\epsilon + \mathrm{d}f \otimes \epsilon + fK\epsilon = f\nabla'\epsilon + \mathrm{d}f \otimes \epsilon \,, \tag{28.3.2}$$

and so ∇' is a Koszul connection. Finally, ∇' is unique, since it is uniquely defined by its action on any section ϵ given in the statement of the theorem.

2. Given two Koszul connections $\nabla, \nabla' : \Gamma(E) \to \Gamma(T^*M \otimes E)$, we define

$$K\epsilon = \nabla'\epsilon - \nabla\epsilon \,, \tag{28.3.3}$$

which is clearly $\mathbb{R}$-linear. To show that it is indeed tensorial (i.e., does not involve derivatives of ϵ), we calculus

$$K(f\epsilon) = \nabla'(f\epsilon) - \nabla(f\epsilon) = f\nabla'\epsilon_i + \mathrm{d}f \otimes \epsilon - f\nabla\epsilon - \mathrm{d}f \otimes \epsilon = fK\epsilon \,. \tag{28.3.4}$$

Finally, K is uniquely defined by its action on any section ϵ. ■

Since K in the theorem above is a section of the bundle $T^*M \otimes \mathrm{End}(E)$, this result suggests that also Koszul connections form sections of an affine bundle modeled over $T^*M \otimes \mathrm{End}(E)$. This is supported by the index structure of the connection coefficients $\omega^a{}_{b\mu}$, since we can write the components of K with the same index positions $K^a{}_{b\mu}$. Indeed, if we write

$$\nabla\epsilon = \left[\partial_\mu y^a + \omega^a{}_{b\mu} y^b\right] \mathrm{d}x^\mu \otimes e_a \,, \quad \nabla'\epsilon = \left[\partial_\mu y^a + \omega'^a{}_{b\mu} y^b\right] \mathrm{d}x^\mu \otimes e_a \,, \tag{28.3.5}$$

then we have

$$K\epsilon = K^a{}_{b\mu} y^b \mathrm{d}x^\mu \otimes e_a \tag{28.3.6}$$

with

$$K^a{}_{b\mu} = \omega'^a{}_{b\mu} - \omega^a{}_{b\mu} \,, \tag{28.3.7}$$

and so one may expect that the connection coefficients $\omega^a{}_{b\mu}$ are related to a local trivialization of this affine bundle.

►...◄

28.4 Parallel transport

Theorem 28.4.1. *Every linear Ehresmann connection is complete.*

Proof. ►...◄ ■

Theorem 28.4.2. *Let $\pi : E \to M$ be a vector bundle equipped with a linear Ehresmann connection and $\gamma : [a, b] \to M$ a curve. Then the parallel transport $\mathcal{P}_\gamma : E_{\gamma(a)} \to E_{\gamma(b)}$ is a linear bijection.*

Proof. Let $e_1, e_2 \in E_{\gamma(a)}$ and denote the corresponding horizontal lifts of γ by $\hat{\gamma}_1, \hat{\gamma}_2 : [a, b] \to E$, where $\hat{\gamma}_1(a) = e_1$ and $\hat{\gamma}_2(a) = e_2$, and hence $\hat{\gamma}_1(b) = \mathcal{P}_\gamma(e_1)$ and $\hat{\gamma}_2(b) = \mathcal{P}_\gamma(e_2)$. For $\mu, \nu \in \mathbb{R}$, define a curve

$$\begin{array}{rcccc} \hat{\gamma} & : & [a,b] & \to & E \\ & & t & \mapsto & \mu\hat{\gamma}_1(t) + \nu\hat{\gamma}_2(t) \end{array} . \tag{28.4.1}$$

This curve satisfies $\mu\hat{\gamma}_1(a) + \nu\hat{\gamma}_2(a) = \mu e_1 + \nu e_2$. ►Show horizontality.◄ Finally, it follows that

$$\mathcal{P}_\gamma(\mu e_1 + \nu e_2) = \hat{\gamma}(b) = \mu\hat{\gamma}_1(b) + \nu\hat{\gamma}_2(b) = \mu\mathcal{P}_\gamma(e_1) + \nu\mathcal{P}_\gamma(e_2)\,, \tag{28.4.2}$$

and so $\mathcal{P}_\gamma$ is indeed linear. Finally, recall from theorem 26.8.2 that $\mathcal{P}_\gamma$ defines a diffeomorphism, and is thus bijective. ■

28.5 Covariant derivative

A Koszul connection allows us to perform another operation on vector bundles. Given a section of a vector bundle and a vector field on the base manifold, it allows us to take the derivative of this section along the vector field. This is defined as follows.

Definition 28.5.1 (Covariant derivative)**.** Let $\pi : E \to M$ be a vector bundle with a Koszul connection ∇. The *covariant derivative* of a section $\epsilon \in \Gamma(E)$ with respect to a vector field $X \in \mathrm{Vect}(M)$ is the section $\nabla_X\epsilon = \iota_X(\nabla\epsilon)$.

From its definition it is clear that the covariant derivative satisfies a number of relations, which immediately follow from the properties of the interior product and the Koszul connection. Here we show the most important relations.

Theorem 28.5.1. *Let ∇ be a Koszul connection on a vector bundle $\pi : E \to M$, $X, Y \in \mathrm{Vect}(M)$ vector fields, $\epsilon, \zeta \in \Gamma(E)$ sections and $f \in C^\infty(M, \mathbb{R})$. The covariant derivative satisfies the following relations:*

$$\nabla_{X+Y}\epsilon = \nabla_X\epsilon + \nabla_Y\epsilon\,, \tag{28.5.1a}$$
$$\nabla_X(\epsilon + \zeta) = \nabla_X\epsilon + \nabla_X\zeta\,, \tag{28.5.1b}$$
$$\nabla_{fX}\epsilon = f\nabla_X\epsilon\,, \tag{28.5.1c}$$
$$\nabla_X(f\epsilon) = f\nabla_X\epsilon + (Xf)\epsilon\,. \tag{28.5.1d}$$

Proof. The first and the third property are obvious from the linearity of the interior product. From the linearity of the Koszul connection immediately follows the second property, while the last one becomes clear by calculating

$$\nabla_X(f\epsilon) = \iota_X(f\nabla\epsilon + \mathrm{d}f \otimes \epsilon) = f\nabla_X\epsilon + (Xf)\epsilon\,. \tag{28.5.2}$$

■

Also the coordinate expression of the covariant derivative should be clear from its definition. If we write $\nabla\epsilon = y^a{}_{;\mu}\mathrm{d}x^\mu \otimes e_a$ and $X = X^\mu\partial_\mu$, then

$$\nabla_X\epsilon = X^\mu y^a{}_{;\mu}e_a = X^\mu(\partial_\mu y^a + \omega^a{}_{b\mu}y^b)e_a\,. \tag{28.5.3}$$

This form is also in agreement with the properties shown in theorem 28.5.1.

In the definition of the covariant derivative given above we have focused on the algebraic perspective in terms of a Koszul connection, which immediately yields a prescription to calculate the covariant derivative in coordinates. Another approach towards the same notion is to take a more geometric perspective and makes use of the properties of the parallel transport which we studied in the previous section. We first show how the same notion arises as follows.

Theorem 28.5.2. *Let $\pi : E \to M$ be a vector bundle equipped with a linear Ehresmann connection and $\epsilon \in \Gamma(E)$ a section. For a vector field $X \in \operatorname{Vect}(M)$, let $\phi : \mathbb{R} \times M \supseteq U \to M$ denote the flow of X. Further, let $\mathcal{P}_{x,t} : E_{\phi_t(x)} \to E_x$ be the parallel transport from $\phi_t(x)$ to x along the flow line of X. Then the covariant derivative is given by*

$$(\nabla_X \epsilon)(x) = \lim_{t \to 0} \frac{\mathcal{P}_{x,t}(\epsilon(\phi_t(x))) - \epsilon(x)}{t} \,. \tag{28.5.4}$$

Proof. ▶...◀ ∎

28.6 Connections on frame bundles

In section 22.1 we have shown how to construct the frame bundle of a vector bundle. In the following, we will study how this construction allows us to obtain a connection on the frame bundle, provided we already have a connection on the underlying vector bundle. The starting point of this discussion is the following definition.

Definition 28.6.1 (Frame bundle connection)**.** Let $\pi : E \to M$ be a vector bundle of rank k, $\varpi : F(E) \to M$ its frame bundle and $\omega : E \to J^1(E)$ a linear Ehresmann connection. The *frame bundle connection* on $F(E)$ is the connection $\hat{\omega} : F(E) \to J^1(F(E))$ such that for each $p \in F(E)$ holds $\hat{\omega}(p) = j^1_{\varpi}\sigma$, where $\sigma \in \Gamma_{\varpi(p)}(F(E))$ is the local section defined on $U \subset M$ such that $\sigma(x)(v) = \sigma_v(x)$ for all $x \in U$, where σ_v is defined by the jet $\omega(p(v)) = j^1_{\varpi(p)}\sigma_v$ for all $v \in \mathbb{R}^k$.

To see that this definition makes sense, we need to check that σ indeed constitutes a local section of the frame bundle, i.e., that $\sigma(x)$ is a bijective linear function which maps $\mathbb{R}^k$ into $E_x = \pi^{-1}(x)$ and that $\sigma(\varpi(p)) = p$. To check this, note that $\sigma_v \in \Gamma_{\varpi(p)}(E)$ is a local section of $\pi : E \to M$, by the definition of a linear Ehresmann connection, and so $\sigma(x)(v) = \sigma_v(x) \in E_x$. Further, $\sigma(x)$ is linear, since both ω and p are linear, and so the assignment $v \mapsto \omega_v$ is linear. In particular, for $x = \varpi(p)$ holds $\sigma(\varpi(p))(v) = \sigma_v(\varpi(p)) = p(v)$, since σ_v is a representative of a jet $\omega(p(v))$. At this point, $\omega(\varpi(p)) = p$ is thus invertible. Hence, there also exists a neighborhood around $\varpi(p)$ where $\sigma(x)$ is invertible, and thus σ constitutes a local section of the frame bundle. Finally, its first jet depends only on the first jet of σ_v by construction.

It is helpful to illustrate the construction given above in coordinates. Let (x^μ) be local coordinates on M and (x^μ, y^a) adapted local coordinates on E which correspond to a local basis. On the frame bundle, we can then introduce local coordinates $(x^\mu, p^a{}_i)$. Coordinates on the first jet bundles are given by (x^μ, y^a, y^a_μ) and $(x^\mu, p^a{}_i, p^a{}_{i\mu})$, respectively. In these coordinates, a frame $p \in \varpi^{-1}(x)$ defines a linear function that maps $v = v^i \mathbb{e}_i \in \mathbb{R}^k$ to $y = y^a e_a = p^a{}_i v^i e_a \in E_x$. The linear Ehresmann connection ω assigns to this element y a jet, which we denote $(x^\mu, y^a, -\omega^a{}_{b\mu}(x) y^b)$ as before. In other words, the values and first partial derivatives of the local sections σ_v at $x = \varpi(p)$ are given by

$$(\sigma_v)^a(x) = p^a{}_i v^i \,, \quad \partial_\mu (\sigma_v)^a(x) = -\omega^a{}_{b\mu}(x) p^b{}_i v^i \,. \tag{28.6.1}$$

Writing the local section σ in the frame bundle coordinates $(x^\mu, p^a{}_i)$ as $\sigma^a{}_i$, we thus have

$$\sigma^a{}_i(x) = p^a{}_i \,, \quad \partial_\mu \sigma^a{}_i(x) = -\omega^a{}_{b\mu}(x) p^b{}_i \,. \tag{28.6.2}$$

Its first jet therefore has the coordinates $(x^\mu, p^a{}_i, -\omega^a{}_{b\mu}(x)p^b{}_i)$. This is the coordinate expression of the frame bundle connection $\hat{\omega}$.

Since we have now constructed a connection on a frame bundle, and frame bundles are in particular principal bundles, the first question which arises is whether this construction yields a principal connection. This will be shown next.

Theorem 28.6.1. *The frame bundle connection $\hat{\omega}$ is a principal Ehresmann connection.*

Proof. In the following, we will show only the equivariance of the connection. For this purpose, consider $p, p' \in F(E)$ with $\varpi(p) = \varpi(p') = x \in M$ and $p' = p \cdot g$ with $g \in G = \mathrm{GL}(k, \mathbb{R})$. Denoting the local section defining the jet $\hat{\omega}(p) = j^1_x\sigma^p$ by σ^p, we then have

$$\hat{\omega}(R_g(p)) = \hat{\omega}(p \cdot g) = j^1_x\sigma^{p\cdot g}\,, \tag{28.6.3}$$

where

$$\sigma^{p\cdot g}(x)(v) = \sigma^{p\cdot g}_v(x) \tag{28.6.4}$$

and

$$j^1_x\sigma^{p\cdot g}_v = \omega((p\cdot g)(v)) = \omega(p(gv)) = j^1_x\sigma^p_{gv}\,. \tag{28.6.5}$$

Hence, we have

$$\sigma^{p\cdot g}(x) = \sigma^p(x) \circ g = \sigma^p(x) \cdot g = R_g(\sigma^p(x))\,, \tag{28.6.6}$$

and thus

$$j^1_x\sigma^{p\cdot g} = j^1_x(R_g \circ \sigma^p) = R_g(j^1_x(\sigma^p)) = R_g(\hat{\omega}(p))\,, \tag{28.6.7}$$

so that $\hat{\omega}$ is indeed equivariant, and hence a principal Ehresmann connection. ■

The coordinate form we have found above thus does not come as a surprise, since it agrees with the one we have found in example 27.1.1 for principal connections on bundles whose structure group is a matrix group, and $\mathrm{GL}(k, \mathbb{R})$ is simply a special case.

28.7 Connections on associated vector bundles

In the previous section we have seen how a linear Ehresmann connection gives rise to a principal connection on its frame bundle. We now come to a complementary construction, that allows us to obtain a linear connection from a principal one. For this purpose, we recall from theorem 20.4.1 that the associated bundle obtained from a principal bundle and a linear representation of its structure group is a vector bundle. We have already given a construction for a connection on arbitrary associated bundles in section 27.5. It remains to show that this connection is a linear Ehresmann connection if the bundle is obtained from a linear representation. This can be shown as follows.

Theorem 28.7.1. *Let $\varpi : P \to M$ be a principal G-bundle, $\omega : P \to J^1(P)$ a principal Ehresmann connection and $\rho : G \to \mathrm{GL}(V)$ a linear representation of G on a k-dimensional vector space V. Then the induced connection $\omega_\rho : E \to J^1(E)$ is a linear Ehresmann connection on the vector bundle $\pi = \varpi_\rho : E = P \times_\rho V \to M$.*

Proof. Recall from theorem 20.4.1 that E is a vector bundle and that the vector space structure on each fiber E_x for $x \in M$ is induced from the relation

$$\mu e + \mu' e' = \mu[p, v] + \mu'[p, v'] = [p, \mu v + \mu' v'] \tag{28.7.1}$$

for all $\mu, \mu' \in \mathbb{R}$, $p \in P_x$, $v, v' \in V$ and $e = [p, v], e' = [p, v'] \in E_x$. This allows us to write the induced connection as

$$\begin{aligned} \omega_\rho(\mu e + \mu' e') &= \omega_\rho([p, \mu v + \mu' v']) \\ &= j^1_x[\sigma_p, \mu v + \mu' v'] \\ &= j^1_x(\mu[\sigma_p, v] + \mu'[\sigma_p, v']) \\ &= \mu j^1_x[\sigma_p, v] + \mu' j^1_x[\sigma_p, v'] \\ &= \mu\omega_\rho([p, v]) + \mu'\omega_\rho([p, v']) \\ &= \mu\omega_\rho(e) + \mu'\omega_\rho(e'), \end{aligned} \tag{28.7.2}$$

where we further used the fact that the fiber $J^1_x(E)$ also inherits its vector space structure from the vector space structure of local sections, so that j^1_x acts as a linear operator. Hence, ω_ρ is a linear Ehresmann connection. ■

While using jets is convenient to define the connection on any associated bundle, it is more practical for calculations to express a linear connection as a Koszul connection, and hence a covariant derivative, which acts on sections of the underlying vector bundle $\pi : E \to M$. In the case that $E = P \times_\rho M$ is an associated bundle obtained from a principal G-bundle $\varpi : P \to M$ and a linear representation $\rho : G \to \mathrm{GL}(V)$ of its structure group on a vector space V, we can make use of theorem 20.3.3, which states that sections $\epsilon \in \Gamma(P \times_\rho M)$ are in one-to-one correspondence with equivariant maps $\tilde{\epsilon} \in C^\infty_G(P, V)$. The latter are basic zero-forms on P with values in V, so that we can apply the exterior covariant derivative to obtain a basic one-form with values in V. This allows us to obtain a Koszul connection as follows.

Theorem 28.7.2. *Let $\varpi : P \to M$ be a principal G-bundle, $\vartheta \in \Omega^1(P, \mathfrak{g})$ a principal G-connection and $\rho : G \to \mathrm{GL}(V)$ a linear representation of G on a k-dimensional vector space V. Then the induced Koszul connection $\nabla : \Gamma(E) \to \Gamma(T^*M \otimes E)$ satisfies* ▶. . . ◀

Proof. ▶. . . ◀ ■

Example 28.7.1. We consider the same principal bundle $\pi : P \to M$ as in example 27.1.1, equipped with the same connection and coordinates. In addition we consider $F = \mathbb{R}^n$ with Cartesian coordinates (f^a). Recall that we assumed the structure group G to be a matrix group with $n \times n$ matrices. This group possesses a natural left action ρ on F by matrix multiplication, $\rho(g, f) = gf$. We denote by $\pi_\rho : E = P \times_\rho F \to M$ the corresponding associated bundle. To define a local trivialization $\phi_\rho : \pi_\rho^{-1}(U) \to U \times F$ of E over a set $U \subset M$, we make use of the trivialization $\phi : \pi^{-1}(U) \to U \times G$ of P and define

$$\phi_\rho([\phi^{-1}(x, e), f]) = (x, f), \tag{28.7.3}$$

where $e \in G$ is the unit element. We write the induced coordinates on $\pi^{-1}(U)$ as (x^μ, y^a). For the corresponding coordinates on the first jet bundle $J^1(E)$ we use the notation (x^μ, y^a, y^a_μ).

To construct the induced connection $\omega_\rho : E \to J^1(E)$ we pick an element $[p, f] \in E_x = \pi_\rho^{-1}(x)$ of the fiber over $x \in U$, and we keep $x \in U$, $p \in P$ and $f \in F$ fixed. Note that the same element $[p, f]$ can equally well be described by any representative $(p \cdot g, \rho(g^{-1}, f))$, and so without loss of generality we may assume that $\phi(p) = (x, e) \in U \times G$. It follows that the coordinate representations of p and f are given by (x^μ, δ^a_b) and (x^μ, f^a), respectively, where (f^a) is the coordinate representation of f. We then choose a section $\sigma : U \to P$ such that $\sigma(x) = p$ and $j^1_x\sigma = \omega(p)$. In coordinates this section can be expressed as

$$\sigma : (x^\mu) \mapsto (x^\mu, \sigma^a{}_b(x)). \tag{28.7.4}$$

while its first jet prolongation reads

$$j^1\sigma : (x^\mu) \mapsto (x^\mu, \sigma^a{}_b(x), \partial_\mu\sigma^a{}_b(x)). \tag{28.7.5}$$

Together with the fixed element $f \in F$ it further defines a section $\tilde{\sigma} = [\sigma(\bullet), f] : U \to E$, which we write in coordinates as

$$\tilde{\sigma} : (x^\mu) \mapsto (x^\mu, \tilde{\sigma}^a(x)) = (x^\mu, \sigma^a{}_b(x) f^b) \,. \tag{28.7.6}$$

Its first jet prolongation is therefore given by

$$j^1\tilde{\sigma} : (x^\mu) \mapsto (x^\mu, \tilde{\sigma}^a(x), \partial_\mu \tilde{\sigma}^a(x)) = (x^\mu, \sigma^a{}_b(x) f^b, \partial_\mu \sigma^a{}_b(x) f^b) \,. \tag{28.7.7}$$

We now evaluate σ, $\tilde{\sigma}$ and their jet prolongations at the fixed point x we have chosen. Due to our choice $\phi(p) = (x, e)$ we have $\sigma^a{}_b(x) = \delta^a_b$. We then use $j^1_x\sigma = \omega(p)$, where evaluating the latter yields

$$\omega : (x^\mu, \delta^a_b) \mapsto (x^\mu, \delta^a{}_b, -\Gamma^a{}_{b\mu}(x)) \,. \tag{28.7.8}$$

We thus have $\partial_\mu \sigma^a{}_b(x) = -\Gamma^a{}_b(x)$. This finally leads to $\tilde{\sigma}^a(x) = f^a$ and $\partial_\mu \tilde{\sigma}^a(x) = -\Gamma^a{}_{b\mu}(x) f^b$. We can thus write the connection $\omega_\rho : E \to J^1(E)$ as

$$\omega_\rho : (x^\mu, y^a) \mapsto (x^\mu, y^a, -\Gamma^a{}_{b\mu}(x) y^b) \,. \tag{28.7.9}$$

Note in particular that the last component is linear in the fiber coordinates y^a, so this is a linear Ehresmann connection. This is due to the fact that we have chosen ρ to be a linear representation of G on the vector space F; it does not hold for general actions ρ which do not have this property.

We finally remark that is we would have chosen the dual representation $\rho^*(g, f) = f g^{-1}$ we would have obtained the dual connection

$$\omega_{\rho^*} : (x^\mu, \bar{y}_a) \mapsto (x^\mu, \bar{y}_a, \Gamma^b{}_{a\mu}(x) \bar{y}_b) \tag{28.7.10}$$

on the dual bundle E^*. Note in particular the change of the sign in the last component.

28.8 Connections on the dual bundle

Given a linear connection on a vector bundle $\pi : E \to M$, one finds that also on other, related bundles a number of connections are defined, which are derived from the original connection. This holds for all operations on vector bundles as discussed in chapter 4. The most straightforward construction is found for the dual vector bundle $\bar{\pi} : E^* \to M$ introduced in section 4.1. Recall that for every $x \in M$, the fiber E^*_x is the dual vector space of the fiber E_x. Given sections $\epsilon \in \Gamma(E)$ and $\zeta \in \Gamma(E^*)$, one therefore can construct a function

$$\begin{array}{cccc} \langle \epsilon, \zeta \rangle & : & M & \to & \mathbb{R} \\ & & x & \mapsto & \langle \epsilon(x), \zeta(x) \rangle \end{array} , \tag{28.8.1}$$

by pointwise application of the canonical pairing of the dual vector space. Since this is a real function, it allows the following definition.

Definition 28.8.1 (Dual bundle connection)**.** Let $\pi : E \to M$ be a vector bundle, equipped with a Koszul connection ∇. The *dual bundle connection* on the dual vector bundle $\bar{\pi} : E^* \to M$ is the unique Koszul connection ∇ on E^* such that for any sections $\epsilon \in \Gamma(E)$ and $\zeta \in \Gamma(E^*)$ and vector fields $X \in \mathrm{Vect}(M)$ the derivatives satisfy

$$\langle \nabla_X \epsilon, \zeta \rangle + \langle \epsilon, \nabla_X \zeta \rangle = X \langle \epsilon, \zeta \rangle \,, \tag{28.8.2}$$

where ∇_X denotes the covariant derivatives on the respective bundles.

This means that we demand a version of the Leibniz rule also for the pairing of sections of dual bundles. We illustrate this definition using the same coordinates (x^μ, y^a) we used also in sections 28.2 and 28.5, as well as coordinates (x^μ, z_a) on the dual bundle E^* as shown in section 4.1. Expressing the sections ϵ and ζ in these coordinates, we find the expressions

$$\begin{aligned}\langle\epsilon, \nabla_X\zeta\rangle &= X\langle\epsilon,\zeta\rangle - \langle\nabla_X\epsilon,\zeta\rangle\\ &= X^\mu\partial_\mu(y^a z_a) - X^\mu y^a{}_{;\mu}z_a\\ &= X^\mu\partial_\mu y^a z_a + X^\mu y^a\partial_\mu z_a - X^\mu(\partial_\mu y^a + \omega^a{}_{b\mu}y^b)z_a\\ &= X^\mu y^a(\partial_\mu z_a - \omega^b{}_{a\mu}z_b)\\ &= X^\mu y^a z_{a;\mu}\,,\end{aligned} \tag{28.8.3}$$

where the last line and the condition that this holds independent of X and ϵ yield the relation

$$z_{a;\mu} = \nabla_\mu z_a = \partial_\mu z_a - \omega^b{}_{a\mu}z_b\,. \tag{28.8.4}$$

Comparing this result with the coordinate formula (28.2.2) for the Koszul connection on E, we see that the sign in front of the connection coefficient $\omega^a{}_{b\mu}$ changes.

28.9 Connections on tensor bundles

For the tensor product of two vector bundles $\pi_1 : E_1 \to M$ and $\pi_2 : E_2 \to M$ over a common base M one can proceed similarly, provided that each of them is equipped with a connection. Here one demands that the Leibniz rule holds for the tensor product of sections. Hence, one defines as follows.

Definition 28.9.1 (Tensor product connection)**.** Let $\pi_1 : E_1 \to M$ and $\pi_2 : E_2 \to M$ be vector bundles over a common base M equipped with Koszul connections, which we denote by ∇ for both bundles. The *tensor product connection* on the tensor product bundle $\pi_1 \otimes \pi_2 : E_1 \otimes E_2 \to M$ is the unique connection on $E_1 \otimes E_2$ such that for all sections $\epsilon_1 \in \Gamma(E_1)$ and $\epsilon_2 \in \Gamma(E_2)$ and vector fields $X \in \mathrm{Vect}(M)$ holds

$$\nabla_X(\epsilon_1 \otimes \epsilon_2) = \nabla_X\epsilon_1 \otimes \epsilon_2 + \epsilon_1 \otimes \nabla_X\epsilon_2\,, \tag{28.9.1}$$

where ∇_X denotes the covariant derivatives on the respective bundles.

Also in this case it is helpful to illustrate the definition by using coordinates. Here the coordinates (x^μ, y^a) on E_1 and $(x^\mu, z^{\bar{a}})$ on E_2, as well as $(x^\mu, w^{a\bar{a}})$ introduced in section 4.3 turn out to be useful. Denoting the sections with these coordinates, and writing $w^{a\bar{a}} = y^a z^{\bar{a}}$, we find that the covariant derivative in coordinates is given by

$$\begin{aligned}\nabla_X(\epsilon_1\otimes\epsilon_2) &= \nabla_X\epsilon_1\otimes\epsilon_2 + \epsilon_1\otimes\nabla_X\epsilon_2\\ &= (X^\mu y^a{}_{;\mu}z^{\bar{a}} + X^\mu y^a z^{\bar{a}}{}_{;\mu})e_a\otimes\bar{e}_{\bar{a}}\\ &= \left[X^\mu(\partial_\mu y^a + \omega^a{}_{b\mu}y^b)z^{\bar{a}} + X^\mu y^a(\partial_\mu z^{\bar{a}} + \bar{\omega}^{\bar{a}}{}_{\bar{b}\mu}z^{\bar{b}})\right]e_a\otimes\bar{e}_{\bar{a}}\\ &= X^\mu\left(\partial_\mu y^a z^{\bar{a}} + y^a\partial_\mu z^{\bar{a}} + \omega^a{}_{b\mu}y^b z^{\bar{a}} + \bar{\omega}^{\bar{a}}{}_{\bar{b}\mu}y^a z^{\bar{b}}\right)e_a\otimes\bar{e}_{\bar{a}}\\ &= X^\mu\left(\partial_\mu w^{a\bar{a}} + \omega^a{}_{b\mu}w^{b\bar{a}} + \bar{\omega}^{\bar{a}}{}_{\bar{b}\mu}w^{a\bar{b}}\right)e_a\otimes\bar{e}_{\bar{a}}\\ &= X^\mu w^{a\bar{a}}{}_{;\mu}e_a\otimes\bar{e}_{\bar{a}}\,,\end{aligned} \tag{28.9.2}$$

where we denoted the connection coefficients on E_2 by $\bar{\omega}^{\bar{a}}{}_{\bar{b}\mu}$. From the last line we read off the relation

$$w^{a\bar{a}}{}_{;\mu} = \nabla_\mu w^{a\bar{a}} = \partial_\mu w^{a\bar{a}} + \omega^a{}_{b\mu}w^{b\bar{a}} + \bar{\omega}^{\bar{a}}{}_{\bar{b}\mu}w^{a\bar{b}}\,. \tag{28.9.3}$$

A few special cases are worth mentioning. Recall from section 4.6 that the homomorphism bundle $\mathrm{Hom}(E_1, E_2)$ is canonically isomorphic to $E_1^* \otimes E_2$. If one introduces coordinates $(x^\mu, w_a{}^{\bar{a}})$ on $\mathrm{Hom}(E_1, E_2)$, ond finds that the formula for the covariant derivative in terms of the connection coefficients $\omega^a{}_{b\mu}$ on E_1 and $\bar{\omega}^{\bar{a}}{}_{\bar{b}\mu}$ on E_2 is given by

$$\nabla_\mu w_a{}^{\bar{a}} = \partial_\mu w_a{}^{\bar{a}} - \omega^b{}_{a\mu} w_b{}^{\bar{a}} + \bar{\omega}^{\bar{a}}{}_{\bar{b}\mu} w_a{}^{\bar{b}} . \tag{28.9.4}$$

In particular, if $E_1 = E_2$ and we consider the same connection on both copies of E, we find the formula

$$\nabla_\mu w_a{}^b = \partial_\mu w_a{}^b - \omega^c{}_{a\mu} w_c{}^b + \omega^b{}_{c\mu} w_a{}^c . \tag{28.9.5}$$

The latter is a special case of the tensor product $E_s^r = \bigotimes^r E \otimes \bigotimes^s E^*$. Denoting the fiber coordinates by $T^{a_1 \cdots a_r}{}_{b_1 \cdots b_s}$, the covariant derivative is given by

$$\begin{aligned} \nabla_\mu T^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} = {} & \partial_\mu T^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} \\ & + \omega^{a_1}{}_{c\mu} T^{c a_2 \cdots a_r}{}_{b_1 \cdots b_s} + \ldots + \omega^{a_r}{}_{c\mu} T^{a_1 \cdots a_{r-1} c}{}_{b_1 \cdots b_s} \\ & - \omega^c{}_{b_1 \mu} T^{a_1 \cdots a_r}{}_{c b_2 \cdots b_s} - \ldots - \omega^c{}_{b_s \mu} T^{a_1 \cdots a_r}{}_{b_1 \cdots b_{s-1} c} . \end{aligned} \tag{28.9.6}$$

This formula may be compared to the Lie derivative (16.2.12) expressed in coordinates. While in the Lie derivative formula for each upper index a term with a negative sign appears, as well as for each lower index a term with a positive sign, the opposite is the case for the covariant derivative.

Other special cases can be obtained by restricting to particular subbundles of the tensor product bundle, in particular the exterior power bundle introduced in section 4.4 and the symmetric power bundle introduced in section 4.5. Here we only discuss the former as a particular example. Recall that we can write a section σ of the bundle $\Lambda^r E$ in coordinates in the form

$$\sigma = \frac{1}{k!} \sigma^{a_1 \cdots a_r} e_{a_1} \wedge \ldots \wedge e_{a_r} = \sigma^{a_1 \cdots a_r} e_{a_1} \otimes \ldots \otimes e_{a_r} , \tag{28.9.7}$$

where the coefficients are totally antisymmetric, $\sigma^{a_1 \cdots a_r} = \sigma^{[a_1 \cdots a_r]}$.

28.10 Connections on density bundles

Finally, we also discuss the case of densities. ►...◄

$$\begin{aligned} \nabla_\mu \mathfrak{T}^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} = {} & \partial_\mu \mathfrak{T}^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} - w \omega^c{}_{c\mu} \mathfrak{T}^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} \\ & + \omega^{a_1}{}_{c\mu} \mathfrak{T}^{c a_2 \cdots a_r}{}_{b_1 \cdots b_s} + \ldots + \omega^{a_r}{}_{c\mu} \mathfrak{T}^{a_1 \cdots a_{r-1} c}{}_{b_1 \cdots b_s} \\ & - \omega^c{}_{b_1 \mu} \mathfrak{T}^{a_1 \cdots a_r}{}_{c b_2 \cdots b_s} - \ldots - \omega^c{}_{b_s \mu} \mathfrak{T}^{a_1 \cdots a_r}{}_{b_1 \cdots b_{s-1} c} . \end{aligned} \tag{28.10.1}$$

28.11 Pullback connections

Definition 28.11.1 (Pullback connection)**.** Let $\pi : E \to N$ be a vector bundle equipped with a Koszul connection $\nabla : \Gamma(E) \to \Gamma(E \otimes T^*N)$ and $\psi : M \to N$ a smooth map. The *pullback* $\psi^*\nabla$ of ∇ along ψ is the unique Koszul connection $\psi^*\nabla : \Gamma(\psi^*E) \to \Gamma(\psi^*E \otimes T^*M)$ on the pullback bundle $\psi^*\pi : \psi^*E \to M$ which satisfies

$$(\psi^*\nabla)(\psi^*\sigma) = \blacktriangleright \ldots \blacktriangleleft \tag{28.11.1}$$

for all sections $\sigma \in \Gamma(E)$.

28.12 Curvature

Recall that we have introduced the curvature of a connection in section 26.10 as the vertical part of the commutator of horizontal vector fields, expressed through the Nijenhuis tensor of the connection form, and for principal connections in section 27.3 as the exterior covariant derivative of the algebra-valued connection form. For linear connections on vector bundles, which can be expressed as Koszul connections and hence give rise to the notion of a covariant derivative, we can introduce curvature in yet another way as the commutator between covariant derivatives. We start with the following definition.

Definition 28.12.1 (Curvature of a linear connection)**.** Let $\pi : E \to M$ be a vector bundle and ∇ a Koszul connection on E. The *curvature* of ∇ is the unique endomorphism valued two-form $F \in \Gamma(\Lambda^2 T^*M \otimes \mathrm{End}(E))$ such that for any vector fields $X, Y \in \mathrm{Vect}(M)$ and section $\epsilon \in \Gamma(E)$ holds

$$F(X,Y)\epsilon = \nabla_X \nabla_Y \epsilon - \nabla_Y \nabla_X \epsilon - \nabla_{[X,Y]}\epsilon\,. \tag{28.12.1}$$

In this definition we have already claimed that the curvature is a two-form which takes values in the endomorphism bundle. For this to be true, it needs to satisfy a number of properties: in particular it must be antisymmetric in the first two arguments and linear in all arguments. In the following, we show that it indeed possesses the required properties.

Theorem 28.12.1. *For all $X, Y, Z \in \mathrm{Vect}(M)$, $\epsilon, \zeta) \in \Gamma(E)$ and $f \in C^\infty(M, \mathbb{R})$, the curvature F of a linear connection on a vector bundle $\pi : E \to M$ satisfies:*

$$F(Y,X)\epsilon = -F(X,Y)\epsilon\,, \tag{28.12.2a}$$
$$F(X+Y,Z)\epsilon = F(X,Z)\epsilon + F(Y,Z)\epsilon\,, \tag{28.12.2b}$$
$$F(X,Y)(\epsilon+\zeta) = F(X,Y)\epsilon + F(X,Y)\zeta\,, \tag{28.12.2c}$$
$$F(fX,Y)\epsilon = fF(X,Y)\epsilon\,, \tag{28.12.2d}$$
$$F(X,Y)(f\epsilon) = fF(X,Y)\epsilon\,. \tag{28.12.2e}$$

Proof. The first three properties follows immediately from the fact that the covariant derivative is linear and that the Lie bracket is linear and antisymmetric in its arguments. For the remaining two propositions, we use the properties 28.5.1 of the covariant derivative and calculate

$$\begin{aligned}
F(fX,Y)\epsilon &= \nabla_{fX}\nabla_Y\epsilon - \nabla_Y\nabla_{fX}\epsilon - \nabla_{[fX,Y]}\epsilon \\
&= f\nabla_X\nabla_Y\epsilon - \nabla_Y(f\nabla_X\epsilon) - \nabla_{f[X,Y]-(Yf)X}\epsilon \\
&= f\nabla_X\nabla_Y\epsilon - f\nabla_Y\nabla_X\epsilon - (Yf)\nabla_X\epsilon - f\nabla_{[X,Y]}\epsilon + (Yf)\nabla_X\epsilon \\
&= f\nabla_X\nabla_Y\epsilon - f\nabla_Y\nabla_X\epsilon - f\nabla_{[X,Y]}\epsilon \\
&= fF(X,Y)\epsilon
\end{aligned} \tag{28.12.3}$$

and

$$\begin{aligned}
F(X,Y)(f\epsilon) &= \nabla_X\nabla_Y(f\epsilon) - \nabla_Y\nabla_X(f\epsilon) - \nabla_{[X,Y]}(f\epsilon) \\
&= \nabla_X(f\nabla_Y\epsilon + (Yf)\epsilon) - \nabla_Y(f\nabla_X\epsilon + (Xf)\epsilon) - f\nabla_{[X,Y]}\epsilon - ([X,Y]f)\epsilon \\
&= f\nabla_X\nabla_Y\epsilon + (Xf)\nabla_Y\epsilon + (Yf)\nabla_X\epsilon + (XYf)\epsilon - f\nabla_X\nabla_Y\epsilon - (Yf)\nabla_X\epsilon \\
&\quad - (Xf)\nabla_Y\epsilon - (YXf)\epsilon - f\nabla_{[X,Y]}\epsilon - (XYf)\epsilon + (YXf)\epsilon \\
&= f\nabla_X\nabla_Y\epsilon - f\nabla_Y\nabla_X\epsilon - f\nabla_{[X,Y]}\epsilon \\
&= fF(X,Y)\epsilon\,. \quad \blacksquare
\end{aligned} \tag{28.12.4}$$

It follows that the curvature is indeed a section of the bundle $\Lambda^2 T^*M \otimes \mathrm{End}(E)$. In coordinates, it can therefore be expressed in the form

$$F = \frac{1}{2} F^a{}_{b\mu\nu} \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \otimes e_a \otimes e^b \,. \tag{28.12.5}$$

To obtain an explicit expression for the components $F^a{}_{b\mu\nu}$, we calculate

$$\begin{aligned}
F(X,Y)\epsilon &= \nabla_X \nabla_Y \epsilon - \nabla_Y \nabla_X \epsilon - \nabla_{[X,Y]}\epsilon \\
&= \nabla_{X^\mu\partial_\mu} \nabla_{Y^\nu\partial_\nu}(y^a e_a) - \nabla_{Y^\nu\partial_\nu} \nabla_{X^\mu\partial_\mu}(y^a e_a) - \nabla_{[X^\mu\partial_\mu, Y^\nu\partial_\nu]}(y^a e_a) \\
&= \big(X^\mu\{\partial_\mu[Y^\nu(\partial_\nu y^a + \omega^a{}_{b\nu} y^b)] + \omega^a{}_{c\mu}[Y^\nu(\partial_\nu y^c + \omega^c{}_{b\nu} y^b)]\} \\
&\quad - Y^\nu\{\partial_\nu[X^\mu(\partial_\mu y^a + \omega^a{}_{b\mu} y^b)] + \omega^a{}_{c\nu}[X^\mu(\partial_\mu y^c + \omega^c{}_{b\mu} y^b)]\} \\
&\quad - (X^\mu\partial_\mu Y^\nu - Y^\mu\partial_\mu X^\nu)(\partial_\nu y^a + \omega^a{}_{b\nu} y^b)\big) e_a \\
&= X^\mu Y^\nu(\partial_\mu\omega^a{}_{b\nu} - \partial_\nu\omega^a{}_{b\mu} + \omega^a{}_{c\mu}\omega^c{}_{b\nu} - \omega^a{}_{c\nu}\omega^c{}_{b\mu}) y^b e_a \,.
\end{aligned} \tag{28.12.6}$$

Hence, it follows that the components of the curvature are given by

$$F^a{}_{b\mu\nu} = \partial_\mu\omega^a{}_{b\nu} - \partial_\nu\omega^a{}_{b\mu} + \omega^a{}_{c\mu}\omega^c{}_{b\nu} - \omega^a{}_{c\nu}\omega^c{}_{b\mu} \,. \tag{28.12.7}$$

One may now pose the question how the notion of curvature we introduced here is related to the ones we introduced before. Since we have related the linear connection to a principal connection in sections 28.6 and 28.7, we start with the notion of curvature for principal connections discussed in section 27.3. Recall that the curvature form $\Omega \in \Omega^2(P, \mathfrak{g})$ of a principal G-connection ϑ is a basic form of type Ad, where the adjoint representation Ad of the structure group G is a representation on its Lie algebra $\mathfrak{g}$. Given a representation $\rho : G \to \mathrm{GL}(V)$ of G on a vector space V, which gives rise to the associated bundle $E = P \times_\rho V$, its differential $\rho_* : \mathfrak{g} \to \mathrm{End}(V)$ is a Lie algebra representation. Hence, $\rho_* \circ \Omega \in \Omega^2(P, \mathrm{End}(V))$ is an endomorphism-valued two-form on P. Note that $\mathrm{End}(V) \cong V \otimes V^*$ naturally carries a representation of G, which is the tensor product representation $\mathrm{End}(\rho) = \rho \otimes \bar{\rho}$, which gives rise to the bundle $\mathrm{End}(E) = P \times_{\mathrm{End}(\rho)} \mathrm{End}(V)$. Finally, recalling that the curvature form $F \in \Gamma(\Lambda^2 T^*M \otimes \mathrm{End}(E))$ is an endomorphism-valued two-form on M, and that basic forms on P and vector-valued forms on M are related by theorem 27.2.1 (for any representation, and thus also for $\mathrm{End}(\rho)$), the following statement suggests itself.

Theorem 28.12.2. *The curvature form $\Omega \in \Omega^2(P, \mathfrak{g})$ of a principal G-connection $\vartheta \in \Omega^1(P, \mathfrak{g})$ on a principal G-bundle $\varpi : P \to M$ and the curvature form $F \in \Gamma(\Lambda^2 T^*M \otimes \mathrm{End}(E))$ of the induced linear connection on the associated vector bundle $\pi : E = P \times_\rho V$ are related such that for each $p \in P$ holds*

$$\rho_*(\Omega_p) = [p]^{-1} \circ (\varpi^* F)_p \circ [p] \,, \tag{28.12.8}$$

where $[p] : V \to P_{\varpi(p)} \times_\rho V$ is the fiber diffeomorphism, and conversely

$$F_{\varpi(p)}(\varpi_*(v), \varpi_*(w)) = [p, \rho_*(\Omega_p(v, w))] \tag{28.12.9}$$

for all $v, w \in T_pP$.

Proof. ►...◄ ■

Another relation concerns the curvature $R \in \Gamma(\Lambda^2 T^*E \otimes VE)$ of a general Ehresmann connection discussed in section 26.10, and can be derived from the fact that a Koszul connection can also be seen as a linear Ehresmann connection, as discussed in section 28.2. The curvature R of this linear Ehresmann connection can be related to the curvature form $F \in \Gamma(\Lambda^2 T^*M \otimes \mathrm{End}(E))$ as follows. One may first take the pullback $\pi^*F \in \Gamma(\Lambda^2 T^*E \otimes \pi^* \mathrm{End}(E))$. The pullback bundle $\pi^*E \cong E \times_\pi E$ has a canonical diagonal section $\delta : e \mapsto (e, e)$, and so one has a vector-valued two-form $(\pi_*F) \circ \delta \in \Gamma(\Lambda^2 T^*E \otimes \pi^*E)$. Since π^*E is canonically isomorphic to VE as shown in theorem 19.8.1, this can also be interpreted as a section of $\Lambda^2 T^*E \otimes VE$, and so one may expect that this yields the curvature R. We now show that this is indeed the case.

Theorem 28.12.3. *The curvature form* $F \in \Gamma(\Lambda^2 T^*M \otimes \mathrm{End}(E))$ *of a linear Ehresmann connection is related to the curvature* $R \in \Gamma(\Lambda^2 T^*M \otimes VE)$ *by*

$$R = (\pi^* F) \circ \delta\,, \tag{28.12.10}$$

where $\delta : E \to \pi^* E, e \mapsto (e, e)$ *is the diagonal section.*

Proof. ▶. . . ◀ ■

28.13 Exterior covariant derivative

By definition 28.2.1, a Koszul connection is a linear assignment $\nabla : \Gamma(E) \to \Gamma(T^*M \otimes E)$. Note that we can also regard the source of this assignment as the space $\Omega^0(M, E)$ of vector-valued zero-forms on M with values on E, while the target space is the space $\Omega^1(M, E)$ of vector-valued one-forms on M with values on E. Hence, a Koszul connection can be regarded as a linear assignment $\nabla : \Omega^0(M, E) \to \Omega^1(M, E)$. This is reminiscent of the total differential $\mathrm{d} : \Omega^0(M) \to \Omega^1(M)$ defined in section 8.4, which assigns to a function $f \in C^\infty(M, \mathbb{R}) \cong \Omega^0(M)$ the one-form $\mathrm{d}f \in \Omega^1(M)$, and which we generalized to the exterior derivative $\mathrm{d} : \Omega^k(M) \to \Omega^{k+1}(M)$ acting on differential forms of arbitrary rank k in section 9.3. One may wonder whether such a generalization is also possible in the case of Koszul connections and E-valued differential forms. It turns out that this is indeed the case, and we define it as follows.

Definition 28.13.1 (Exterior covariant derivative)**.** Let $\pi : E \to M$ be a vector bundle with a Koszul connection ∇. The *exterior covariant derivative* associated to ∇ is the unique linear function $\mathrm{d}^\nabla : \Omega^k(M, E) \to \Omega^{k+1}(M, E)$ such that

$$\mathrm{d}^\nabla(\sigma \otimes \epsilon) = \mathrm{d}\sigma \otimes \epsilon + (-1)^k \sigma \wedge \nabla\epsilon \tag{28.13.1}$$

for all $\sigma \in \Omega^k(M)$ and $\epsilon \in \Gamma(E)$.

Like the ordinary exterior derivative, also the exterior covariant derivative on vector bundles satisfies a few relations, which turn out to be helpful in practical calculations. We start by showing the following.

Theorem 28.13.1. *The exterior covariant derivative satisfies*

$$\mathrm{d}^\nabla(\omega \wedge \alpha) = \mathrm{d}\omega \wedge \alpha + (-1)^k \omega \wedge \mathrm{d}^\nabla \alpha \tag{28.13.2}$$

for all $\omega \in \Omega^k(M)$ *and* $\alpha \in \Omega^l(M, E)$.

Proof. Due to the linearity, it is sufficient to show this relation for $\alpha = \sigma \otimes \epsilon$ with $\sigma \in \Omega^l(M)$. Then $\omega \wedge \sigma \in \Omega^{k+l}(M)$, and by direct calculation one has

$$\begin{aligned}
\mathrm{d}^\nabla(\omega \wedge (\sigma \otimes \epsilon)) &= \mathrm{d}^\nabla((\omega \wedge \sigma) \otimes \epsilon) \\
&= \mathrm{d}(\omega \wedge \sigma) \otimes \epsilon + (-1)^{k+l} \omega \wedge \sigma \wedge \nabla\epsilon \\
&= (\mathrm{d}\omega \wedge \sigma) \otimes \epsilon + (-1)^k \omega \wedge \mathrm{d}\sigma \otimes \epsilon + (-1)^{k+l} \omega \wedge \sigma \wedge \nabla\epsilon \\
&= \mathrm{d}\omega \wedge \sigma \otimes \epsilon + (-1)^k [\omega \wedge \mathrm{d}\sigma \otimes \epsilon + (-1)^l \omega \wedge \sigma \wedge \nabla\epsilon] \\
&= \mathrm{d}\omega \wedge \sigma \otimes \epsilon + (-1)^k \omega \wedge \mathrm{d}^\nabla(\sigma \otimes \epsilon)\,.
\end{aligned} \tag{28.13.3}$$

■

For the ordinary exterior derivative, it is well known that it squares to zero. This is not the case for the exterior covariant derivative, whose square is closely related to the curvature. This is what we show next.

Theorem 28.13.2. *The exterior covariant derivative satisfies*

$$\mathrm{d}^\nabla \mathrm{d}^\nabla \alpha = F \wedge \alpha \tag{28.13.4}$$

for all $\alpha \in \Omega^k(M, E)$.

Proof. Once again, we consider $\alpha = \sigma \otimes \epsilon$ and then conclude on the general case by linearity. By direct calculation, we have

$$\begin{aligned} \mathrm{d}^\nabla \mathrm{d}^\nabla(\sigma \otimes \epsilon) &= \mathrm{d}^\nabla[\mathrm{d}\sigma \otimes \epsilon + (-1)^k \sigma \wedge \nabla\epsilon] \\ &= \mathrm{dd}\sigma \otimes \epsilon + (-1)^{k+1}\mathrm{d}\sigma \wedge \nabla\epsilon + (-1)^k \mathrm{d}\sigma \wedge \nabla\epsilon + (-1)^{2k}\sigma \wedge \mathrm{d}^\nabla\nabla\epsilon \\ &= \sigma \wedge \mathrm{d}^\nabla \nabla\epsilon\,. \end{aligned} \tag{28.13.5}$$

Hence, it is sufficient to show the formula for a zero-form $\epsilon \in \Omega^0(M, E) \cong \Gamma(E)$. Here $\nabla\epsilon \in \Omega^1(M, E)$. Again we use the linearity and assume a form $\nabla\epsilon = \tau \otimes \zeta$ with a one-form $\tau \in \Omega^1(M)$. For vector fields $X, Y \in \mathrm{Vect}(M)$ we then have

$$\begin{aligned} \mathrm{d}^\nabla(\tau \otimes \zeta)(X, Y) &= [\mathrm{d}\tau \otimes \zeta - \tau \wedge \nabla\zeta](X, Y) \\ &= \mathrm{d}\tau(X, Y) \otimes \zeta - \tau(X) \otimes \nabla_Y\zeta + \tau(Y) \otimes \nabla_X\zeta \\ &= [X(\tau(Y)) - Y(\tau(X)) - \tau([X, Y])] \otimes \zeta - \tau(X) \otimes \nabla_Y\zeta + \tau(Y) \otimes \nabla_X\zeta \\ &= \nabla_X[\tau(Y) \otimes \zeta] - \nabla_Y[\tau(X) \otimes \zeta] - \tau([X, Y]) \otimes \zeta\,, \end{aligned} \tag{28.13.6}$$

using the relation (9.4.5). Substituting back, we find

$$(\mathrm{d}^\nabla\nabla\epsilon)(X, Y) = \nabla_X\nabla_Y\epsilon - \nabla_Y\nabla_X\epsilon - \nabla_{[X,Y]}\epsilon = F(X, Y)(\epsilon). \tag{28.13.7}$$

Here F is a two-form, and so its exterior product with σ commutes, so that we finally find

$$\mathrm{d}^\nabla \mathrm{d}^\nabla(\sigma \otimes \epsilon) = F \wedge \sigma \otimes \epsilon\,. \tag{28.13.8}$$

■

From the general theory of connections we know that the curvature of any connection satisfies the Bianchi identity. Using the exterior covariant derivative, we can also formulate the Bianchi identity for the curvature form on vector bundles as follows.

Theorem 28.13.3 (Bianchi identity)**.** *The curvature form satisfies* $\mathrm{d}^\nabla F = 0$.

Proof. ►...◄ ■

In the following, we will discuss another possibility to obtain the exterior covariant derivative we have defined above. This construction can be applied in the case of an associated vector bundle $E = P \times_\rho F$, on which a Koszul connection ∇ is induced by a principal G-connection ϑ on the principal G-bundle P as shown in section 28.7. Recall from section 27.2 that a principal connection ϑ on P gives rise to an exterior covariant derivative d_ϑ acting on basic forms on the total space on the principal bundle, as given in definition 27.2.2. In this section, however, we study vector-valued differential forms on the *base manifold* of a bundle instead. As we have shown in theorem 27.2.1, these are, in fact, in one-to-one correspondence with basic forms. One may thus wonder whether the exterior covariant derivative d_ϑ on the principal bundle and the exterior covariant derivative d^∇ acting on vector-valued forms given in definition 28.13.1 are related by this one-to-one correspondence. We confirm that this is the case as follows:

Theorem 28.13.4. *The exterior covariant derivative of a basic form of type* ρ *corresponds to the induced exterior covariant derivative on the associated vector bundle* $P \times_\rho F$.

Proof. ►...◄ ■

28.14 Holonomy

In section 27.8 we have seen that the parallel transport from one point back to itself along all closed curves yields a group, which is a subgroup of the structure group. Using the fact that the parallel transport on vector bundles is linear and bijective, one finds a similar concept, which comes from the fact that linear and bijective maps from a vector space to itself also form a group. We may thus define the following.

Definition 28.14.1 (Holonomy group)**.** Let $\pi : E \to M$ be a vector bundle equipped with a linear connection ω. For $x \in M$, the *holonomy group* is given by

$$\mathrm{Hol}_x(\omega) = \{\mathcal{P}_x(\gamma), \gamma \in \mathcal{C}_{\pi(p)}(M)\} \subset \mathrm{GL}(E_x)\,, \tag{28.14.1}$$

where $\mathcal{C}_x(M)$ denotes the set of curves $\gamma : [0,1] \to M$ with $\gamma(0) = \gamma(1) = x$, while its *reduced holonomy group* is

$$\mathrm{Hol}^0_x(\omega) = \{\mathcal{P}_x(\gamma), \gamma \in \mathcal{C}^0_{\pi(p)}(M)\} \subset \mathrm{GL}(E_x)\,, \tag{28.14.2}$$

where $\mathcal{C}^0_x(M)$ denotes the set of zero-homotopic curves $\gamma : [0,1] \to M$ with $\gamma(0) = \gamma(1) = x$.

Also here we have already claimed that the sets we have defined above are groups. As in the case of principal bundles, this can be proven as follows.

Theorem 28.14.1. *For each point $x \in M$ of a vector bundle $\pi : E \to M$ with connection ω, the holonomy group $\mathrm{Hol}_x(\omega)$ is a normal Lie subgroup of $\mathrm{GL}(E_x)$ and $\mathrm{Hol}^0_p(\omega)$ is its connected component containing the unit element.*

Proof. ▶. . . ◀ ■

In the case of principal bundles, we have further seen that there exists a close relation between the Lie algebra of the holonomy group and the curvature of a principal connection. Given that the curvature of a linear connection is a two-form with values in the endomorphism bundle, and that at each point $x \in M$ the fiber $\mathrm{End}_x(E)$ is just the Lie algebra of $\mathrm{GL}(E_x)$, one may expect a similar relation also for linear connections. We first show the following.

Theorem 28.14.2. *For each point $x \in M$ of a vector bundle $\pi : E \to M$ with linear connection ω and $v, w \in T_xM$, the curvature satisfies*

$$F_x(v, w) \in \mathrm{Lie}(\mathrm{Hol}_x(\omega))\,. \tag{28.14.3}$$

Proof. ▶. . . ◀ ■

Part II

Particular geometries

Chapter 29

Canonical tangent bundle geometry

29.1 Coordinates on the tangent bundle

In the following we will consider a particular vector bundle, namely the tangent bundle $\tau : TM \to M$ over a manifold M. Its total space TM is again a manifold, and we will be interested in tensor fields defined on this manifold. For this purpose, we will be working with its tangent and cotangent bundles, which we write as $\varpi : TTM \to TM$ and $\bar{\varpi} : T^*TM \to TM$, respectively. In order to work with objects on TM, it is often convenient to introduce a particular set of coordinates on TM and use the corresponding coordinate bases on TTM and T^*TM. We define these coordinates as follows.

Definition 29.1.1 (Induced coordinates on the tangent bundle)**.** Let M be a manifold of dimension $\dim M = n$ and $(x^a, a = 1, \ldots, n)$ a set of coordinates on $U \subset M$. The corresponding *induced coordinates* are the coordinates $(x^a, \bar{x}^a)$ on $\tau^{-1}(U) \subset TM$, where for each element in $\tau^{-1}(U)$, (x^a) are the coordinates of the base point $x \in U$ and the fiber coordinates are used to denote $\bar{x}^a \partial_a \in T_x M$, where ∂_a is the coordinate basis of TM induced by the coordinates (x^a).

There exist different conventions in the literature how to denote the fiber coordinates. It is denoted (y^a) in [BCS91, MA94, BM07]. Two different notations are used in [Run59], namely $(\dot{x}^a)$, which is reminiscent of the coordinate convention for jet bundles over the one-dimensional base manifold $\mathbb{R}$, and $(\mathrm{d}x^a)$. Both conventions are also used elsewhere in the literature. However, we avoid them here, since $\dot{x}^a$ will be used to denote the (components of the) tangent vector of curves, while $\mathrm{d}x^a$ is used for the coordinate basis of the cotangent bundle T^*M.

For later use, it is also instructive to note the change of the induced coordinates which arises from a change of the coordinates on the base manifold. For this purpose, assume we are given another set (x'^a) of coordinates on the base manifold, as functions of the original coordinates (x^a). Then one finds that the fiber part of the induced coordinates $(x'^a, \bar{x}'^a)$ is given by

$$\bar{x}'^a = \frac{\partial x'^a}{\partial x^b} \bar{x}^b \,. \tag{29.1.1}$$

Hence, the induced fiber coordinates transform as if they were components of vector fields.

Given coordinates on TM, it is straightforward to construct the coordinate bases on the bundles TTM and T^*TM. We will denote these bases by

$$\left(\partial_a = \frac{\partial}{\partial x^a}, \bar{\partial}_a = \frac{\partial}{\partial \bar{x}^a} \right) \tag{29.1.2}$$

and

$$(\mathrm{d}x^a, \mathrm{d}\bar{x}^a)\,. \tag{29.1.3}$$

Note that the symbol ∂_a now denotes both a coordinate vector field on M and on TM. However, it should be clear from the context which of these objects is meant. The same holds for the basis element $\mathrm{d}x^a$.

Also for later use, it is helpful to note how these coordinate bases change under a change of coordinates on the base manifold. The coordinate bases obtained from a new set (x'^a) of coordinates on M can be derived from the standard formulas

$$\partial'_a = \frac{\partial x^b}{\partial x'^a}\partial_b + \frac{\partial \bar{x}^b}{\partial x'^a}\bar{\partial}_b\,, \quad \bar{\partial}'_a = \frac{\partial x^b}{\partial \bar{x}'^a}\partial_b + \frac{\partial \bar{x}^b}{\partial \bar{x}'^a}\bar{\partial}_b \tag{29.1.4}$$

and

$$\mathrm{d}x'^a = \frac{\partial x'^a}{\partial x^b}\mathrm{d}x^b + \frac{\partial x'^a}{\partial \bar{x}^b}\mathrm{d}\bar{x}^b\,, \quad \mathrm{d}\bar{x}'^a = \frac{\partial \bar{x}'^a}{\partial x^b}\mathrm{d}x^b + \frac{\partial \bar{x}'^a}{\partial \bar{x}^b}\mathrm{d}\bar{x}^b\,. \tag{29.1.5}$$

These formulas simplify, since the coordinates x'^a on the base manifold depend only on the coordinates x^a, but not on the fiber coordinates $\bar{x}^a$; the same holds for the converse direction. Hence,

$$\frac{\partial x'^a}{\partial \bar{x}^b} = 0\,, \quad \frac{\partial x^a}{\partial \bar{x}'^b} = 0\,. \tag{29.1.6}$$

Further, making use of the dependence (29.1.1) of the transformed fiber coordinates on the original coordinates, one can write

$$\frac{\partial \bar{x}'^a}{\partial \bar{x}^b} = \frac{\partial x'^a}{\partial x^b}\,, \quad \frac{\partial \bar{x}'^a}{\partial x^b} = \bar{x}^c\frac{\partial x'^a}{\partial x^b \partial x^c}\,, \tag{29.1.7}$$

and analogously by exchanging the old and new (primed and unprimed) coordinates. Using these relations, the basis transformations can further be expanded as

$$\partial'_a = \frac{\partial x^b}{\partial x'^a}\partial_b + \bar{x}'^c\frac{\partial x^b}{\partial x'^a \partial x'^c}\bar{\partial}_b\,, \quad \bar{\partial}'_a = \frac{\partial x^b}{\partial x'^a}\bar{\partial}_b \tag{29.1.8}$$

and

$$\mathrm{d}x'^a = \frac{\partial x'^a}{\partial x^b}\mathrm{d}x^b\,, \quad \mathrm{d}\bar{x}'^a = \bar{x}^c\frac{\partial x'^a}{\partial x^b \partial x^c}\mathrm{d}x^b + \frac{\partial x'^a}{\partial x^b}\mathrm{d}\bar{x}^b\,. \tag{29.1.9}$$

We will make use of these formulas in later sections.

29.2 Tangent structure

The tangent bundle TM of a manifold M is canonically equipped with a number of geometric objects. We have already encountered the Liouville vector field in definition 19.9.2, which is defined on every vector bundle, and hence also on TM. Another object, which is specific to the tangent bundle TM, is defined as follows.

Definition 29.2.1 (Tangent structure)**.** Let M be a manifold, $\tau : TM \to M$ its tangent bundle and $\varpi : TTM \to TM$ the double tangent bundle. For each $\psi \in TTM$ we then have projections $v = \varpi(\psi)$ and $p = \tau(v)$, so that $\psi \in T_vTM$ and $v \in T_pM$. For the differential $\tau_* : TTM \to TM$ of τ we also have $\tau_*(\psi) \in T_pM$. We can thus consider a curve

$$\begin{array}{rcrcl} \gamma_\psi & : & \mathbb{R} & \to & TM \\ & & \lambda & \mapsto & \gamma_\psi(\lambda) = \varpi(\psi) + \lambda\tau_*(\psi) \end{array}\,. \tag{29.2.1}$$

The map $J : TTM \to TTM$ which assigns to each $\psi \in TTM$ the tangent vector $\dot{\gamma}_\psi(0) \in T_vTM$ of the corresponding curve γ_ψ is called the *tangent structure*.

There are different alternative definitions of the tangent structure, which lead to canonically equivalent objects. To see this, and derive a coordinate expression for J, we first note the following.

Theorem 29.2.1. *The tangent structure $J : TTM \to TTM$ is a vector bundle homomorphism from $\varpi : TTM \to TM$ to itself, covering the identity on TM.*

Proof. We will not show smoothness here for brevity. Using the notations from definition 29.2.1, we have

$$J(\psi) = \dot{\gamma}_\psi(0) \in T_vTM \tag{29.2.2}$$

and hence $\varpi(J(\psi)) = v = \varpi(\psi)$, so that J indeed preserves the fibers of $\varpi : TTM \to TM$. To see that it is linear on each fiber, consider $\psi, \psi' \in T_vTM$ and $\mu, \mu' \in \mathbb{R}$. Then we have $\varpi(\psi) = \varpi(\psi') = v$ and, using the fact that τ_* is linear, we find

$$\gamma_{\mu\psi+\mu'\psi'}(\lambda) = v + \lambda[\mu\tau_*(\psi) + \mu'\tau_*(\psi')]\,. \tag{29.2.3}$$

Taking the tangent vector

$$J(\mu\psi + \mu'\psi') = \dot{\gamma}_{\mu\psi+\mu'\psi'}(0) = \mu\dot{\gamma}_\psi(0) + \mu'\dot{\gamma}_{\psi'}(0) = \mu J(\psi) + \mu' J(\psi') \tag{29.2.4}$$

shows that J is indeed linear on each fiber. Hence, it is a vector bundle homomorphism. ■

Recall from section 4.6 that one can identify a vector bundle homomorphism with a section of the homomorphism bundle, or equivalently, a section of a particular tensor product bundle. Here we consider the tangent structure J as a section of the homomorphism bundle $\mathrm{Hom}(TTM, TTM) \cong TTM \otimes T^*TM$, and use it to derive a coordinate expression for J. For this purpose, we use the induced coordinates $(x^a, \bar{x}^a)$ on TM defined in section 29.1 and write ψ in the coordinate basis as $\psi = \psi^a\partial_a + \bar{\psi}^a\bar{\partial}_a \in T_vTM$, using again the notation from definition 29.2.1. Similar we can write $v = v^a\partial_a \in T_pM$, as well as $\tau_*(\psi) = \psi^a\partial_a \in T_pM$. Finally denoting the coordinates of the base point $p = \tau(v) \in M$ by (p^a), we have the curve γ_ψ in coordinates given by

$$\gamma_\psi : \lambda \mapsto (p^a, v^a + \lambda\psi^a) \in TM\,. \tag{29.2.5}$$

Note that for $\lambda = 0$ the curve passes through the point $v \in TM$ with coordinates (p^a, v^a), with the (vertical) tangent vector $J(\psi) = \psi^a\bar{\partial}_a \in T_vTM$. Hence, we can write J in induced coordinates as the rank $(1, 1)$ tensor field

$$J = \bar{\partial}_a \otimes \mathrm{d}x^a \in \Gamma(TTM \otimes T^*TM)\,, \tag{29.2.6}$$

which acts on ψ as

$$J(\psi) = \bar{\partial}_a \otimes \mathrm{d}x^a(\psi^b\partial_b + \bar{\psi}^b\bar{\partial}_b) = \psi^b\delta^a_b\bar{\partial}_a = \psi^a\bar{\partial}_a\,. \tag{29.2.7}$$

From the coordinate expression of the tangent structure one may conclude a few more interesting properties. The first one concerns its image. We find the following:

Theorem 29.2.2. *The image of the tangent structure J is VTM, i.e., for every $\psi \in TTM$, its image is vertical, $J(\psi) \in VTM$.*

Proof. By definition 29.2.1, the curve γ_ψ is entirely contained in the tangent space $T_{\tau(\varpi(\psi))}M$. Hence, $\tau(\gamma_\psi(\lambda)) = \tau(\varpi(\psi))$ is constant and does not depend on λ. The tangent vector of $\tau \circ \gamma_\psi$ at $\lambda = 0$ therefore vanishes, and so $\tau_*(\dot{\gamma}_\psi(0)) = 0$. This means that $\dot{\gamma}_\psi(0) = J(\psi)$ is vertical. ■

This becomes apparent also from the coordinate expression (29.2.7), since $\psi^a\bar{\partial}_a \in VTM$ is vertical.

The previous result implies that J is not an isomorphism on TTM, and since its image is not the whole bundle, it must also possess a non-trivial kernel. Indeed, the kernel is easily found:

Theorem 29.2.3. *The kernel of the tangent structure is VTM, i.e., $J(\psi) = 0 \in T_{\varpi(\psi)}TM$ for all $\psi \in VTM$.*

Proof. It follows immediately from the definition 29.2.1 that $J(\psi) = 0$ if and only if $\tau_*(\psi) = 0$, so that J and τ_* have the same kernel. By definition, the vertical tangent bundle VTM is the kernel of τ_*. ■

Also this becomes apparent from the coordinate expression (29.2.7), since $J(\bar{\psi}^a\bar{\partial}_a) = 0$ if $\psi \in VTM$ vertical.

From the two previous statements now the next one immediately follows.

Theorem 29.2.4. *The tangent structure $J : TTM \to TTM$ satisfies $J \circ J = 0$, where $0 : TTM \to TTM$ denotes the vector bundle homomorphism which assigns to every $\psi \in TTM$ the zero element $0 \in T_{\varpi(\psi)}TM$ in the corresponding fiber over $\varpi(\psi) \in TM$.*

Proof. This follows immediately from theorems 29.2.2 and 29.2.3. ■

In the following, we will consider the tangent structure as a tensor field J of rank $(1, 1)$ on the tangent bundle TM, whose coordinate expression is given by (29.2.6), in order to derive a few more properties. The first is its homogeneity.

Theorem 29.2.5. *The tangent structure, viewed as a tensor field of rank $(1, 1)$ on the tangent bundle TM, is homogeneous of order -1.*

Proof. Given a vector $\psi \in TTM$, the pullback $\chi_\lambda^* J$ of J along the dilatation χ_λ with $\lambda \in \mathbb{R}$ acts as

$$(\chi_\lambda^* J)\psi = \chi_{-\lambda*} J \chi_{\lambda*}\psi\,. \tag{29.2.8}$$

We then follow the construction in the definition 29.2.1. First note that

$$\tau_*(\chi_{\lambda*}\psi) = (\tau \circ \chi_\lambda)_*(\psi) = \tau_*(\psi)\,, \tag{29.2.9}$$

since $\chi_\lambda : TM \to TM$ covers the identity on M, and hence $\tau \circ \chi_\lambda = \tau$. Further, we have

$$\varpi(\chi_{\lambda*}\psi) = \chi_\lambda \varpi(\psi) = e^\lambda \varpi(\psi)\,, \tag{29.2.10}$$

since $\chi_{\lambda*} : TTM \to TTM$ covers χ_λ on TM, and so $\varpi \circ \chi_{\lambda*} = \chi_\lambda \circ \varpi$. Thus, we obtain a curve

$$\begin{array}{rcccl} \gamma_{\chi_{\lambda*}\psi} & : & \mathbb{R} & \to & TM \\ & & t & \mapsto & \gamma_{\chi_{\lambda*}\psi}(t) = e^\lambda \varpi(\psi) + t\tau_*(\psi) \end{array}. \tag{29.2.11}$$

We further need to calculate its tangent vector $\dot{\gamma}_{\chi_{\lambda*}\psi}(0)$, which we then push along $\chi_{-\lambda}$. These steps can be combined by using

$$\chi_{-\lambda*}\dot{\gamma}_{\chi_{\lambda*}\psi}(0) = \dot{\tilde{\gamma}}_{\chi_{\lambda*}\psi}(0)\,, \tag{29.2.12}$$

where

$$\tilde{\gamma}_{\chi_{\lambda*}\psi}(t) = (\chi_{-\lambda} \circ \gamma_{\chi_{\lambda*}\psi})(t) = \varpi(\psi) + te^{-\lambda}\tau_*(\psi)\,. \tag{29.2.13}$$

Comparing with the curve γ_ψ used in the construction of $J(\psi)$, we now see that

$$\tilde{\gamma}_{\chi_{\lambda*}\psi}(t) = \gamma_\psi(te^{-\lambda})\,, \tag{29.2.14}$$

and hence

$$(\chi_\lambda^* J)\psi = \dot{\tilde{\gamma}}_{\chi_{\lambda*}\psi}(0) = e^{-\lambda}\dot{\gamma}_\psi(0) = e^{-\lambda}J\psi. \tag{29.2.15}$$

Since this holds for all $\psi \in TTM$, we conclude

$$\chi_\lambda^* J = e^{-\lambda}J\,, \tag{29.2.16}$$

and so J is homogeneous of order -1. ■

Another set of properties can be derived by realizing that the tangent structure is also a vector valued one-form on TM, hence $J \in \Omega^1(TM, TTM)$. This allows us to apply the theory of graded derivations introduced in chapter 17, from which we derive the following properties.

Theorem 29.2.6. *The tangent structure $J \in \Omega^1(TM, TTM)$ and the Liouville vector field $\mathbf{c} \in \Omega^0(TM, TTM)$, viewed as vector-valued differential forms on TM, satisfy the following graded commutation relations:*

$$[\iota_{\mathbf{c}}, \iota_J] = 0\,, \tag{29.2.17a}$$
$$[\iota_J, \mathcal{L}_{\mathbf{c}}] = \iota_J\,, \tag{29.2.17b}$$
$$[\iota_{\mathbf{c}}, \mathcal{L}_J] = \iota_J\,, \tag{29.2.17c}$$
$$[\mathcal{L}_J, \mathcal{L}_{\mathbf{c}}] = \mathcal{L}_J\,, \tag{29.2.17d}$$
$$[\mathcal{L}_J, \mathcal{L}_J] = 0\,, \tag{29.2.17e}$$
$$[\iota_J, \mathcal{L}_J] = 0\,. \tag{29.2.17f}$$

Proof. We make frequent use of the fact that $\mathbf{c}$ is vertical, and so $\iota_{\mathbf{c}} J = J\mathbf{c} = 0$, as well as the homogeneity of J, from which follows $[\![\mathbf{c}, J]\!] = \mathcal{L}_{\mathbf{c}} J = -J$, to show the stated relations:

1. By the definition 17.4.1 of the Nijenhuis-Richardson bracket, one has
$$[\iota_{\mathbf{c}}, \iota_J] = \iota_{[\mathbf{c},J]^\wedge}\,, \tag{29.2.18}$$
where
$$[\mathbf{c}, J]^\wedge = \iota_{\mathbf{c}} J - \iota_J \mathbf{c} = 0\,, \tag{29.2.19}$$
where the second term vanishes, since $\mathbf{c}$ is a vector-valued zero-form.

2. Following theorem 17.7.2, we have
$$[\iota_J, \mathcal{L}_{\mathbf{c}}] = \iota_{[\![J,\mathbf{c}]\!]} - \mathcal{L}_{\iota_J \mathbf{c}} = \iota_J\,. \tag{29.2.20}$$

3. Again using theorem 17.7.2, we find
$$[\iota_{\mathbf{c}}, \mathcal{L}_J] = -\iota_{[\![\mathbf{c},J]\!]} - \mathcal{L}_{\iota_{\mathbf{c}} J} = \iota_J\,. \tag{29.2.21}$$

4. From the definition 17.6.1 of the Frölicher-Nijenhuis bracket and its relation 17.6.7 follows
$$[\mathcal{L}_J, \mathcal{L}_{\mathbf{c}}] = \mathcal{L}_{[\![J,\mathbf{c}]\!]} = \mathcal{L}_J\,. \tag{29.2.22}$$

5. ▶. . . ◀

6. ▶. . . ◀ ■

Of particular interest is the relation (29.2.17e), which is related to the fact that the vertical tangent bundle VTM is an integrable distribution. It can equivalently be formulated as follows.

Theorem 29.2.7. *The Nijenhuis tensor $N_J = [\![J, J]\!]/2$ of the tangent structure vanishes, $N_J = 0$.*

Proof. This follows immediately from the relation (29.2.17e), together with the definition of the Frölicher-Nijenhuis bracket as
$$\mathcal{L}_{[\![J,J]\!]} = [\mathcal{L}_J, \mathcal{L}_J]\,. \quad ■ \tag{29.2.23}$$

Another helpful relation is the following, which relates it to the Liouville vector field.

Theorem 29.2.8. *For any curve $\gamma : \mathbb{R} \to M$, the tangent structure and Liouville vector field are related by*
$$J \circ \ddot{\gamma} = \mathbf{c} \circ \dot{\gamma}\,. \tag{29.2.24}$$

Proof. Let $t \in \mathbb{R}$. Then by definition of the tangent structure we have

$$\begin{aligned}(J \circ \ddot{\gamma})(t) &= \frac{\mathrm{d}}{\mathrm{d}\lambda}(\dot{\gamma}(t) + \lambda(\tau_* \circ \ddot{\gamma})(t))\Big|_{\lambda=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}\lambda}(\dot{\gamma}(t) + \lambda\dot{\gamma}(t))\Big|_{\lambda=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}\lambda}\dot{\gamma}(t)(1+\lambda)\Big|_{\lambda=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}\lambda}\dot{\gamma}(t)e^{\lambda}\Big|_{\lambda=0} \\ &= (\mathbf{c} \circ \dot{\gamma})(t)\,,\end{aligned} \tag{29.2.25}$$

where we replaced $1 + \lambda$ by e^{λ}, since both have the same value and first derivative at $\lambda = 0$. ■

29.3 Cotangent structure

In analogy to the tangent structure discussed above, there also exists an objects which is in a certain sense dual to the tangent structure. This is defined as follows.

Definition 29.3.1 (Cotangent structure)**.** Let M be a manifold and J its tangent structure. The *cotangent structure* is the unique map $J^* : T^*TM \to T^*TM$ such that for all vector fields $X \in \mathrm{Vect}(TM)$ and one-forms $\sigma \in \Omega^1(TM)$ on TM holds

$$\iota_X(J^* \circ \sigma) = \iota_{J \circ X}\sigma\,. \tag{29.3.1}$$

Note that despite the notation, which is conventional in the literature such as [BM07], J^* is not a pullback. In analogy to theorem 29.2.1, also the cotangent structure has a few useful properties, such as the following.

Theorem 29.3.1. *The cotangent structure $J^* : T^*TM \to T^*TM$ is a vector bundle homomorphism from $\varpi : T^*TM \to TM$ to itself, covering the identity on TM.*

Proof. ▶... ◀ ■

Theorem 29.3.2. *For any $\alpha \in T^*TM$ holds $J^*(\alpha) = 0$ if and only if α is horizontal, $\alpha \in H^*TM$.*

Proof. ▶... ◀ ■

We finally derive a coordinate expression for J^*. Writing $X \in \mathrm{Vect}(TM)$ and $\sigma \in \Omega^1(TM)$ in the induced coordinates as

$$X = X^a\partial_a + \bar{X}^a\bar{\partial}_a\,, \quad \sigma = \sigma_a\mathrm{d}x^a + \bar{\sigma}_a\mathrm{d}\bar{x}^a\,, \tag{29.3.2}$$

we have

$$J \circ X = X^a\bar{\partial}_a\,, \tag{29.3.3}$$

and hence

$$\iota_X(J^* \circ \sigma) = \iota_{J \circ X}\sigma = X^a\bar{\sigma}_a\,. \tag{29.3.4}$$

By comparing this with the left hand side, and demanding that it holds for any vector field $X \in \mathrm{Vect}(TM)$, we find that we must have

$$J^* \circ \sigma = \bar{\sigma}_a \mathrm{d}x^a \,. \tag{29.3.5}$$

We can make use of this result and interpret the cotangent structure as a tensor field of rank $(1,1)$ on TM, for which we now find the coordinate expression

$$J^* = \mathrm{d}x^a \otimes \bar{\partial}_a \,. \tag{29.3.6}$$

Observe that this is very similar to the coordinate expression (29.2.6), the only difference being the order of the factors in the tensor product.

29.4 Lifts of functions

One peculiar property of the tangent bundle is the possibility to lift certain objects, which are defined on the base manifold M, to objects defined on the tangent bundle. We have already encountered such a lift in definition 7.3.2 of the canonical lift of a curve into the tangent bundle. In this section we will see that also functions can be lifted from the base manifold to the tangent bundle. In fact, it turns out that there are even several possibilities to construct such a lift. We start with the following.

Definition 29.4.1 (Vertical lift of a function)**.** Let $f \in C^\infty(M, \mathbb{R})$ be a function. Its *vertical lift* is the function $\overset{\mathrm{V}}{f} = f \circ \tau \in C^\infty(TM, \mathbb{R})$ on the tangent bundle $\tau : TM \to M$.

Comparing this definition with definition 11.1.1, we see that the vertical lift is simply the pullback $\overset{\mathrm{V}}{f} = \tau^* f$ of f along the tangent bundle projection. The following property of the vertical lift becomes immediately clear.

Theorem 29.4.1. *The vertical lift $\overset{V}{f}$ of a function $f \in C^\infty(M, \mathbb{R})$ is a homogeneous function on TM of order* 0.

Proof. Given a dilatation $\chi_\lambda : TM \to TM$ we have

$$\chi_\lambda^* \overset{\mathrm{V}}{f} = \overset{\mathrm{V}}{f} \circ \chi_\lambda = f \circ \tau \circ \chi_\lambda = f \circ \tau = \overset{\mathrm{V}}{f} = e^{0 \cdot \lambda} \overset{\mathrm{V}}{f} \,, \tag{29.4.1}$$

where $\tau \circ \chi_\lambda = \tau$ since χ_λ is vertical. ∎

This can also easily be seen in coordinates. Using induced coordinates $(x^a, \bar{x}^a)$, the functional dependence on these coordinates is simply expressed as $\overset{\mathrm{V}}{f}(x^a, \bar{x}^a) = f(x^a)$.

Another type of lift is the following.

Definition 29.4.2 (Complete lift of a function)**.** Let $f \in C^\infty(M, \mathbb{R})$ be a function. Its *complete lift* is the function $\overset{\mathrm{C}}{f} \in C^\infty(TM, \mathbb{R})$ on the tangent bundle $\tau : TM \to M$ defined by

$$\begin{array}{rccc} \overset{\mathrm{C}}{f} : & TM & \to & \mathbb{R} \\ & v & \mapsto & v(f) \end{array} \,. \tag{29.4.2}$$

To clarify this definition, recall that we defined tangent vectors as derivations, according to definition 7.1.1. Every derivation acts on functions, and the result is a real number. In induced coordinates, the complete lift is thus expressed as

$$\overset{C}{f} = \bar{x}^a \partial_a f\,. \tag{29.4.3}$$

On may ask whether also in this case the resulting function on TM is homogeneous. Intuitively one may expect homogeneity of order 1, since the complete lift depends linearly on the vector argument. We show that this is indeed the case.

Theorem 29.4.2. *The complete lift $\overset{C}{f}$ of a function $f \in C^\infty(M, \mathbb{R})$ is a homogeneous function on TM of order 1.*

Proof. Given a dilatation $\chi_\lambda : TM \to TM$ we have

$$\left(\chi_\lambda^* \overset{C}{f}\right)(v) = \overset{C}{f}(\chi_\lambda(v)) = \overset{C}{f}(e^\lambda v) = e^\lambda v(f) = e^\lambda \overset{C}{f}(v)\,. \tag{29.4.4}$$

■

We further recall that a derivation, by definition, satisfies the Leibniz rule. One may thus expect that this property also propagates to the complete lift of a function. Indeed also this is the case, and we find the following relation.

Theorem 29.4.3. *The vertical and complete lifts satisfy the Leibniz rule*

$$\overset{C}{(fg)} = \overset{C}{f}\overset{V}{g} + \overset{V}{f}\overset{C}{g} \tag{29.4.5}$$

for any functions $f, g \in C^\infty(M, \mathbb{R})$.

Proof. By the Leibniz rule for derivations we have

$$\begin{aligned} \overset{C}{(fg)}(v) &= v(fg) \\ &= v(f)g(\tau(v)) + f(\tau(v))v(g) \\ &= \overset{C}{f}(v)\overset{V}{g}(v) + \overset{V}{f}(v)\overset{C}{g}(v) \\ &= (\overset{C}{f}\overset{V}{g} + \overset{V}{f}\overset{C}{g})(v) \end{aligned} \tag{29.4.6}$$

for all $v \in TM$. ■

We see that the complete and vertical lifts of functions satisfy a number of relations. One may ask, conversely, whether there are any relations which are satisfied *only* by these lifts, and which can thus be used as a criterion to determine whether an arbitrary function on the tangent bundle comes from lifting a function on the base manifold. This is indeed the case. For the vertical lift, it is immediately clear that a function $F \in C^\infty(TM, \mathbb{R})$ is a vertical lift if and only if it is constant on the fibers. We find that there is another possibility to formulate this condition, by using the cotangent structure J^* as follows.

Theorem 29.4.4. *A function $F \in C^\infty(TM, \mathbb{R})$ is a vertical lift, $F = \overset{V}{f}$ for some $f \in C^\infty(M, \mathbb{R})$, if and only if $J^*\mathrm{d}F = 0$.*

Proof. Following theorem 29.3.2, $J^*\mathrm{d}F = 0$ if and only if $\mathrm{d}F$ is horizontal, $\mathrm{d}F \in \Gamma(H^*TM)$. This is equivalent to the statement that $\mathrm{d}F$ vanishes on every vertical vector field $X \in \Gamma(VTM)$,

$$0 = \iota_X \mathrm{d}F = XF\,. \tag{29.4.7}$$

Now this can equivalently be stated as F being constant along the fibers of TM, and hence the vertical lift of a function on M. ■

For the complete lift, we find a similarly simple condition.

Theorem 29.4.5. *A function $F \in C^\infty(TM, \mathbb{R})$ is the sum of a complete and a vertical lift, $F = \overset{C}{f_1} + \overset{V}{f_2}$ for some $f_1, f_2 \in C^\infty(M, \mathbb{R})$, if and only if $\mathrm{d}J^*\mathrm{d}F = 0$.*

Proof. ▶. . . ◀ ■

The latter two statements can also be illustrated using coordinates. For a function $F \in C^\infty(TM, \mathbb{R})$, we have

$$\mathrm{d}F = \partial_a F \mathrm{d}x^a + \bar{\partial}_a F \mathrm{d}\bar{x}^a \,, \tag{29.4.8}$$

$$J^*\mathrm{d}F = \bar{\partial}_a F \mathrm{d}x^a \tag{29.4.9}$$

and

$$\mathrm{d}J^*\mathrm{d}F = \partial_a \bar{\partial}_b F \mathrm{d}x^a \wedge \mathrm{d}x^b + \bar{\partial}_a \bar{\partial}_b F \mathrm{d}\bar{x}^a \wedge \mathrm{d}x^b \,. \tag{29.4.10}$$

From these expressions one easily sees that $J^*\mathrm{d}F = 0$ if and only if $\bar{\partial}_a F = 0$, which is the case if and only if F is a vertical lift. Similarly, $\mathrm{d}J^*\mathrm{d}F = 0$ if and only if

$$\partial_{[a} \bar{\partial}_{b]} F = \bar{\partial}_a \bar{\partial}_b F = 0 \,. \tag{29.4.11}$$

The latter implies that F must be of the form

$$F(x^a, \bar{x}^a) = f(x^a) + \bar{x}^a v_a(x^a) \,. \tag{29.4.12}$$

The former then gives the condition

$$\partial_{[a} v_{b]}(x^a) = 0 \,, \tag{29.4.13}$$

from which follows

$$v_a = \partial_a \tilde{f}(x^a) \,. \tag{29.4.14}$$

We thus see that $F = \overset{V}{f} + \overset{C}{\tilde{f}}$.

29.5 Lifts of vector fields

The next class of objects which we will lift from the manifold M to its tangent bundle are vector fields. As it was also the case for functions, we find that there are two canonical ways how this can be done. We first define the following.

Definition 29.5.1 (Vertical lift of a vector field)**.** Let M be a manifold and $X \in \mathrm{Vect}(M)$ a vector field. Its *vertical lift* is the vector field $\overset{V}{X} \in \mathrm{Vect}(TM)$ defined by

$$\overset{V}{X}(v) = \frac{\mathrm{d}}{\mathrm{d}t}(v + tX(\tau(v)))\bigg|_{t=0} \tag{29.5.1}$$

as the tangent vector to the curve $t \mapsto v + tX(\tau(v))$ at $t = 0$, where $\tau : TM \to M$ is the bundle map of the tangent bundle.

To illustrate this definition, we derive its coordinate expression using the induced tangent bundle coordinates $(x^a, \bar{x}^a)$ introduced in section 29.1. In these coordinates, the vector field X is expressed by component functions $X^a = X^a(x)$ on the base manifold M as $X = X^a \partial_a$. To derive its vertical lift $\overset{V}{X}(x, \bar{x})$ at a point with coordinates $(x^a, \bar{x}^a)$ in the tangent bundle, we construct the curve

$$\lambda \mapsto (x^a, \bar{x}^a + \lambda X^a(x)) \tag{29.5.2}$$

on TM. At $\lambda = 0$, we find its tangent vector $X^a(x)\bar{\partial}_a \in T_{(x,\bar{x})}TM$. Hence, the vertical lift has the coordinate expression

$$\overset{\mathrm{V}}{X} = X^a\bar{\partial}_a \,. \tag{29.5.3}$$

One easily checks the following property of the vertical lift.

Theorem 29.5.1. *The vertical lift $\overset{\mathrm{V}}{X}$ of a vector field $X \in \mathrm{Vect}(M)$ is a homogeneous vector field on TM of order -1.*

Proof. We will check the homogeneity using its definition 19.9.3, together with the pullback 12.1.1 of a vector field and the pushforward 10.2.1 of the tangent vector of a curve. Putting these elements together, we find the formula

$$\begin{aligned}
\left(\chi_\lambda^*\overset{\mathrm{V}}{X}\right)(v) &= \chi_{-\lambda*}\left(\overset{\mathrm{V}}{X}(\chi_\lambda(v))\right)\\
&= \chi_{-\lambda*}\left(\overset{\mathrm{V}}{X}(e^\lambda v)\right)\\
&= \chi_{-\lambda*}\left[\frac{\mathrm{d}}{\mathrm{d}t}(e^\lambda v + tX(\tau(e^\lambda v)))\bigg|_{t=0}\right]\\
&= \frac{\mathrm{d}}{\mathrm{d}t}\chi_{-\lambda}(e^\lambda v + tX(\tau(v)))\bigg|_{t=0}\\
&= \frac{\mathrm{d}}{\mathrm{d}t}(v + te^{-\lambda}X(\tau(v)))\bigg|_{t=0}\\
&= e^{-\lambda}\overset{\mathrm{V}}{X}(v)
\end{aligned} \tag{29.5.4}$$

for all $X \in \mathrm{Vect}(X)$, $\lambda \in \mathbb{R}$ and $v \in TM$. ■

This can also be seen from the coordinate expression (29.5.3) and theorem 19.9.1, from which one finds

$$\mathcal{L}_{\mathbf{c}}\overset{\mathrm{V}}{X} = \left[\mathbf{c}, \overset{\mathrm{V}}{X}\right] = [\bar{x}^a\bar{\partial}_a, X^b\bar{\partial}_b] = -X^a\bar{\partial}_a = -\overset{\mathrm{V}}{X} \,. \tag{29.5.5}$$

Another type of lift arises from the fact that the tangent bundle is a *natural* bundle, in the sense we studied in chapter 10. Such bundles allow a functorial, or canonical, lift of diffeomorphisms $\varphi : M \to M$ from the base manifold M to the total space of the bundle. In the case of the tangent bundle, this lift is given by the differential $\varphi_* : TM \to TM$. Considering not a single diffeomorphism φ, but a family of diffeomorphisms generated by the flow of a vector field, we arrive at the notion of the complete lift of a vector field, which we can define as follows.

Definition 29.5.2 (Complete lift of a vector field). Let M be a manifold and $X \in \mathrm{Vect}(M)$ a vector field with flow $\phi : \mathbb{R} \times M \supseteq U \to M$. Its *complete lift* is the vector field $\overset{\mathrm{C}}{X} \in \mathrm{Vect}(TM)$ whose flow is given by

$$\overset{\mathrm{C}}{\phi} \;:\; \begin{array}{ccc} \mathbb{R} \times TM \supseteq (\mathrm{id}_{\mathbb{R}}, \tau)^{-1}(U) & \to & TM \\ (t, v) & \mapsto & \phi_{t*}(v) \end{array} \,. \tag{29.5.6}$$

To see that $\overset{\mathrm{C}}{\phi}$ given above indeed defines the flow of a vector field, one must check that every curve $\Gamma_v : t \mapsto \overset{\mathrm{C}}{\phi}_t(v)$ defined by fixing v is an integral curve, and thus in particular

$$\dot{\Gamma}_v(t) = \overset{\mathrm{C}}{X}(\Gamma_v(t)) = \dot{\Gamma}_{\overset{\mathrm{C}}{\phi}_t(v)}(0) \,. \tag{29.5.7}$$

This follows from the fact that $\phi_{t+s} = \phi_s \circ \phi_t$, since ϕ is a flow, and hence $\phi_{(t+s)*} = \phi_{s*} \circ \phi_{t*}$, by differentiating. For a function $f \in C^\infty(TM, \mathbb{R})$ one has

$$\begin{aligned}\dot{\Gamma}_v(t)(f) &= \frac{\mathrm{d}}{\mathrm{d}s} f(\Gamma_v(t+s))\bigg|_{s=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}s} f(\phi_{(t+s)*}(v))\bigg|_{s=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}s} f(\phi_{s*}(\phi_{t*}(v)))\bigg|_{s=0} \\ &= \dot{\Gamma}_{\phi_{t*}(v)}(0)(f)\,.\end{aligned} \tag{29.5.8}$$

There exists another possibility to define the canonical lift of a vector field, which is equivalent to the definition we gave above. This alternative definition is given by its relation to the complete lift of a function, which we defined in the previous section. We find the following relation.

Theorem 29.5.2. *The complete lift* $\overset{C}{X} \in \mathrm{Vect}(TM)$ *is the unique vector field on* TM *such that*

$$\overset{C}{(Xf)} = \overset{C}{X}\overset{C}{f} \tag{29.5.9}$$

for all functions $f \in C^\infty(M, \mathbb{R})$.

Proof. Recall from section 16.1 that any vector field is uniquely defined by its flow, through the tangent vectors $X(x) = \dot{\gamma}_x$ to the integral curves $\gamma_x(t) = \phi_t(x)$. Instead of the vector fields, we can thus work with their flows. Further, recall from section 16.3 that the action of a vector field on a function is given by the Lie derivative, and hence

$$Xf = \frac{\mathrm{d}}{\mathrm{d}t}(f \circ \phi_t)\bigg|_{t=0}\,. \tag{29.5.10}$$

To calculate the complete lift of Xf, consider a vector $v \in TM$, seen as a derivation. By commuting the action of v and the derivative with respect to t, and then using the definition of the pushforward, we have

$$\overset{C}{(Xf)}(v) = v(Xf) = \frac{\mathrm{d}}{\mathrm{d}t} v(f \circ \phi_t)\bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\phi_{t*}(v)(f)\bigg|_{t=0}\,. \tag{29.5.11}$$

Now using the definition of the complete lift of f again, the expression under the derivative becomes

$$\phi_{t*}(v)(f) = \overset{C}{f}(\phi_{t*}(v)) = \left(\overset{C}{f} \circ \phi_{t*}\right)(v)\,, \tag{29.5.12}$$

and so we see that we can write $\overset{C}{(Xf)}(v)$ as the action of a vector on $\overset{C}{f}$, which arises as the tangent vector to the curve $t \mapsto \phi_{t*}(v) \in TM$. By the relation (29.5.9), this tangent vector defines $\overset{C}{X}(v)$, i.e.,

$$(\overset{C}{X}\overset{C}{f})(v) = \overset{C}{(Xf)}(v) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\overset{C}{f} \circ \phi_{t*}\right)(v)\bigg|_{t=0}\,. \tag{29.5.13}$$

This shows that the flow $\overset{C}{\phi}$ of $\overset{C}{X}$ and ϕ_* agree at $t = 0$, and so it uniquely defines $\overset{C}{X}$. ∎

It is helpful to derive a coordinate expression for the canonical lift using the induced coordinates on the tangent bundle. Denote the coordinates of the vector $v \in TM$ by $(x^a, \bar{x}^a)$. Its base point $\tau(v)$ therefore has coordinates (x^a). Further, we write the coordinates of $\phi_{t*}(v)$ as $(x'^a, \bar{x}'^a)$, and hence $\tau(\phi_{t*}(v)) = \phi_t(\tau(v))$ has coordinates (x'^a). By definition of the differential and the induced coordinates we have

$$\bar{x}'^a = \frac{\partial x'^a}{\partial x^b}\bar{x}^b\,. \tag{29.5.14}$$

Using the definition of the flow of a vector field $X = X^a\partial_a$, one then finds

$$\begin{aligned}
\overset{C}{X} &= \left.\frac{\mathrm{d}x'^a}{\mathrm{d}t}\right|_{t=0}\partial_a + \left.\frac{\mathrm{d}\bar{x}'^a}{\mathrm{d}t}\right|_{t=0}\bar{\partial}_a \\
&= \left.\frac{\mathrm{d}x'^a}{\mathrm{d}t}\right|_{t=0}\partial_a + \left.\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial x'^a}{\partial x^b}\bar{x}^b\right)\right|_{t=0}\bar{\partial}_a \\
&= \left.\frac{\mathrm{d}x'^a}{\mathrm{d}t}\right|_{t=0}\partial_a + \bar{x}^b\frac{\partial}{\partial x^b}\left.\frac{\mathrm{d}x'^a}{\mathrm{d}t}\right|_{t=0}\bar{\partial}_a \\
&= X^a\partial_a + \bar{x}^b\partial_b X^a\bar{\partial}_a\,.
\end{aligned} \tag{29.5.15}$$

One can compare this derivation with the alternative definition given in theorem 29.5.2. Using $\overset{C}{f} = \bar{x}^a\partial_a f$, the left hand side of (29.5.9) reads

$$(\overset{C}{X f}) = \bar{x}^a\partial_a(X^b\partial_b f) = \bar{x}^a\partial_a X^b\partial_b f + \bar{x}^a X^b\partial_a\partial_b f\,. \tag{29.5.16}$$

For the right hand side, we write $\overset{C}{X} = \overset{C}{X}{}^a\partial_a + \overset{C}{\bar{X}}{}^a\bar{\partial}_a$ with yet to be determined components $\overset{C}{X}{}^a$ and $\overset{C}{\bar{X}}{}^a$. This yields

$$\overset{C}{X}\overset{C}{f} = \overset{C}{X}{}^a\bar{x}^b\partial_a\partial_b f + \overset{C}{\bar{X}}{}^a\partial_a f\,. \tag{29.5.17}$$

By comparison with the left hand side, one thus identifies $\overset{C}{X}{}^a = X^a$ and $\overset{C}{\bar{X}}{}^a = \bar{x}^b\partial_b X^a$, which yields again the formula (29.5.15).

The complete lift has a number of interesting properties. A question which naturally arises is whether it is homogeneous, and of which order. This is answered by the following statement.

Theorem 29.5.3. *The complete lift $\overset{C}{X}$ of a vector field $X \in \mathrm{Vect}(M)$ is a homogeneous vector field on TM of order* 0.

Proof. Again it is most convenient to use the definition 19.9.3 for the proof. Using the linearity of the pushforward one finds

$$\begin{aligned}
\left(\chi_\lambda^*\overset{C}{X}\right)(v) &= \chi_{-\lambda*}\left(\overset{C}{X}(\chi_\lambda(v))\right) \\
&= \chi_{-\lambda*}\left(\overset{C}{X}(e^\lambda v)\right) \\
&= \chi_{-\lambda*}\left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\phi_{t*}(e^\lambda v)\right|_{t=0}\right) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\chi_{-\lambda}(\phi_{t*}(e^\lambda v))\right|_{t=0} \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\chi_{-\lambda}(e^\lambda\phi_{t*}(v))\right|_{t=0} \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\phi_{t*}(v)\right|_{t=0} \\
&= \overset{C}{X}(v)
\end{aligned} \tag{29.5.18}$$

for all $X \in \mathrm{Vect}(X)$, $\lambda \in \mathbb{R}$ and $v \in TM$. ■

Of course, this can also be derived using the coordinate expression (29.5.15). Together with the statement 19.9.1, we can write

$$\begin{aligned}
\mathcal{L}_{\mathbf{c}}\overset{C}{X} &= [\mathbf{c}, \overset{C}{X}] \\
&= [\bar{x}^c\bar{\partial}_c, X^a\partial_a + \bar{x}^b\partial_b X^a\bar{\partial}_a] \\
&= \bar{x}^c\delta_c^b\partial_b X^a\bar{\partial}_a - \bar{x}^b\partial_b X^a\delta_a^c\bar{\partial}_c \\
&= \bar{x}^b\partial_b X^a\bar{\partial}_a - \bar{x}^b\partial_b X^a\bar{\partial}_a \\
&= 0\,.
\end{aligned} \tag{29.5.19}$$

One may wonder whether there is any relation between the complete and vertical lifts of a vector field. This is indeed the case, and it is given by the tangent structure as follows.

Theorem 29.5.4. *The vertical and complete lifts of a vector field $X \in \mathrm{Vect}(M)$ are related by $\overset{V}{X} = J\overset{C}{X}$.*

Proof. We can use the fact that both the vertical and complete lift, as well as the tangent structure, are defined via tangent vectors to certain curves, and explicitly construct the corresponding curves. For $v \in TM$ we then find

$$\begin{aligned}
\left(J\overset{\mathrm{C}}{X}\right)(v) &= \frac{\mathrm{d}}{\mathrm{d}s}\left(\varpi(\overset{\mathrm{C}}{X}(v)) + s\tau_*(\overset{\mathrm{C}}{X}(v))\right)\Big|_{s=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\left(v + s\tau_*\left(\frac{\mathrm{d}}{\mathrm{d}t}\phi_{t*}(v)\Big|_{t=0}\right)\right)\Big|_{s=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\left(v + s\,\frac{\mathrm{d}}{\mathrm{d}t}\tau(\phi_{t*}(v))\Big|_{t=0}\right)\Big|_{s=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\left(v + s\,\frac{\mathrm{d}}{\mathrm{d}t}\phi_t(\tau(v))\Big|_{t=0}\right)\Big|_{s=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\left(v + sX(\tau(v))\right)\Big|_{s=0} \\
&= \overset{\mathrm{V}}{X}(v)\,,
\end{aligned} \tag{29.5.20}$$

where ϕ denotes the flow of X. Here we made use of the pushforward of the tangent vector of a curve given by theorem 10.2.1 to conclude

$$\tau_*\left(\frac{\mathrm{d}}{\mathrm{d}t}\phi_{t*}(v)\Big|_{t=0}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\tau(\phi_{t*}(v))\Big|_{t=0}\,, \tag{29.5.21}$$

and theorem 10.1.1 that the differential is a vector bundle homomorphism covering the original map to conclude

$$\tau(\phi_{t*}(v)) = \phi_t(\tau(v))\,. \quad \blacksquare \tag{29.5.22}$$

This can also be seen easily by using coordinates. From the expressions (29.2.6) and (29.5.15) follows

$$J\overset{\mathrm{C}}{X} = (\bar{\partial}_c \otimes \mathrm{d}x^c)(X^a\partial_a + \bar{x}^b\partial_b X^a\bar{\partial}_a) = X^a\bar{\partial}_a = \overset{\mathrm{V}}{X}\,. \tag{29.5.23}$$

Another helpful set of relations between the complete and vertical lifts of vector fields is given by the following formulas for their commutators.

Theorem 29.5.5. *The Lie bracket and the vertical and complete lifts of vector fields $X, Y \in \mathrm{Vect}(M)$ satisfy the relations*

$$\left[\overset{V}{X}, \overset{V}{Y}\right] = 0\,, \quad \left[\overset{V}{X}, \overset{C}{Y}\right] = \overset{V}{[X, Y]}\,, \quad \left[\overset{C}{X}, \overset{C}{Y}\right] = \overset{C}{[X, Y]}\,. \tag{29.5.24}$$

Proof. Let $f \in C^\infty(TM, \mathbb{R})$ be a function on the tangent bundle. To show the propositions, we will make use of the definition 7.5.1 of the commutator of vector fields. This leads to the following relations:

1. For any $v \in TM$, we have by definition of the vertical lift

$$\left(\overset{\mathrm{V}}{X}f\right)(v) = \frac{\mathrm{d}}{\mathrm{d}t}f(v + tX(\tau(v)))\Big|_{t=0}\,, \tag{29.5.25}$$

and analogously for Y. Hence, we find

$$\begin{aligned}
\left(\overset{\mathrm{v}}{X}\overset{\mathrm{v}}{Y}f\right)(v) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\overset{\mathrm{v}}{Y}f\right)(v + tX(\tau(v)))\bigg|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}}{\mathrm{d}s}f(v + tX(\tau(v)) + sY(\tau(v + tX(\tau(v)))))\bigg|_{s=0}\bigg|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}}{\mathrm{d}s}f(v + tX(\tau(v)) + sY(\tau(v)))\bigg|_{s=0}\bigg|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\frac{\mathrm{d}}{\mathrm{d}t}f(v + tX(\tau(v)) + sY(\tau(v)))\bigg|_{t=0}\bigg|_{s=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\frac{\mathrm{d}}{\mathrm{d}t}f(v + tX(\tau(v + sY(\tau(v)))) + sY(\tau(v)))\bigg|_{t=0}\bigg|_{s=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\left(\overset{\mathrm{v}}{X}f\right)(v + sY(\tau(v)))\bigg|_{s=0} \\
&= \left(\overset{\mathrm{v}}{Y}\overset{\mathrm{v}}{X}f\right)(v)\,,
\end{aligned} \tag{29.5.26}$$

where we used the facts that the parameter derivatives with respect to t and s commute for smooth functions, as well as

$$\tau(v + tX(\tau(v))) = \tau(v + sY(\tau(v))) = \tau(v)\,, \tag{29.5.27}$$

since the corresponding curves are entirely contained in $T_{\tau(v)}M$. Thus, $\overset{\mathrm{v}}{X}$ and $\overset{\mathrm{v}}{Y}$ commute, and so $\left[\overset{\mathrm{v}}{X}, \overset{\mathrm{v}}{Y}\right] = 0$.

2. We can proceed similarly to the previous derivation. First note that by definition of the complete lift and the flow we have for $v \in TM$:

$$\left(\overset{\mathrm{c}}{Y}f\right)(v) = \frac{\mathrm{d}}{\mathrm{d}t}f(\phi_{t*}(v))\bigg|_{t=0}\,, \tag{29.5.28}$$

where ϕ denotes the flow of Y. Now one may calculate

$$\begin{aligned}
\left(\left[\overset{\mathrm{v}}{X}, \overset{\mathrm{c}}{Y}\right]f\right)(v) &= \left(\overset{\mathrm{v}}{X}\overset{\mathrm{c}}{Y}f\right)(v) - \left(\overset{\mathrm{c}}{Y}\overset{\mathrm{v}}{X}f\right)(v) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\overset{\mathrm{c}}{Y}f\right)(v + tX(\tau(v)))\bigg|_{t=0} - \frac{\mathrm{d}}{\mathrm{d}s}\left(\overset{\mathrm{v}}{X}f\right)(\phi_{s*}(v))\bigg|_{s=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}}{\mathrm{d}s}f(\phi_{s*}(v + tX(\tau(v))))\bigg|_{s=0}\bigg|_{t=0} \\
&\quad - \frac{\mathrm{d}}{\mathrm{d}s}\frac{\mathrm{d}}{\mathrm{d}t}f(\phi_{s*}(v) + tX(\tau(\phi_{s*}(v))))\bigg|_{t=0}\bigg|_{s=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}}{\mathrm{d}s}f(\phi_{s*}(v) + t\phi_{s*}(X(\tau(v))))\bigg|_{s=0}\bigg|_{t=0} \\
&\quad - \frac{\mathrm{d}}{\mathrm{d}s}\frac{\mathrm{d}}{\mathrm{d}t}f(\phi_{s*}(v) + tX(\phi_s(\tau(v))))\bigg|_{t=0}\bigg|_{s=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}f(v - t(\mathcal{L}_Y X)(\tau(v)))\bigg|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}f(v + t[X, Y](\tau(v)))\bigg|_{t=0} \\
&= \left(\overset{\mathrm{v}}{[X, Y]}f\right)(v)\,.
\end{aligned} \tag{29.5.29}$$

Here we used the facts that the differential ϕ_{s*} is linear, so that

$$\phi_{s*}(v + tX(\tau(v))) = \phi_{s*}(v) + t\phi_{s*}(X(\tau(v)))\,, \tag{29.5.30}$$

and that it is a vector bundle homomorphism, so that

$$\tau(\phi_{s*}(v)) = \phi_s(\tau(v))\,. \tag{29.5.31}$$

Finally, the Lie derivative follows from

$$(\mathcal{L}_Y X)(\tau(v)) = \frac{\mathrm{d}}{\mathrm{d}s}\left(X(\phi_s(\tau(v))) - \phi_{s*}(X(\tau(v)))\right)\Big|_{s=0}\,. \tag{29.5.32}$$

3. We can make use of theorem 29.5.2 which states that for any $f \in C^\infty(M, \mathbb{R})$ holds

$$\begin{aligned}
\left[\overset{\mathrm{C}}{X}, \overset{\mathrm{C}}{Y}\right]\overset{\mathrm{C}}{f} &= \overset{\mathrm{C}}{X}\overset{\mathrm{C}}{Y}\overset{\mathrm{C}}{f} - \overset{\mathrm{C}}{Y}\overset{\mathrm{C}}{X}\overset{\mathrm{C}}{f} \\
&= \overset{\mathrm{C}}{X}(\overset{\mathrm{C}}{Yf}) - \overset{\mathrm{C}}{Y}(\overset{\mathrm{C}}{Xf}) \\
&= (\overset{\mathrm{C}}{XYf}) - (\overset{\mathrm{C}}{YXf}) \\
&= (\overset{\mathrm{C}}{[X,Y]f}) \\
&= \overset{\mathrm{C}}{[X,Y]}\overset{\mathrm{C}}{f}\,.
\end{aligned} \tag{29.5.33}$$

Since the action on arbitrary complete lifts $\overset{\mathrm{C}}{f}$ uniquely determines a vector field on TM, it follows that $\left[\overset{\mathrm{C}}{X}, \overset{\mathrm{C}}{Y}\right] = \overset{\mathrm{C}}{[X,Y]}$. ■

The last statement for the complete lift in particular shows that the function $\overset{\mathrm{C}}{\bullet} : \mathrm{Vect}(M) \to \mathrm{Vect}(TM)$ is a Lie algebra homomorphism. Again it is convenient to illustrate the result by explicitly calculating the Lie brackets from the corresponding coordinate expressions:

1. For the vertical lifts one finds

$$\left[\overset{\mathrm{V}}{X}, \overset{\mathrm{V}}{Y}\right] = [X^a\bar{\partial}_a, Y^b\bar{\partial}_b] = X^a\bar{\partial}_a Y^b\bar{\partial}_b - Y^b\bar{\partial}_b X^a\bar{\partial}_a = 0\,, \tag{29.5.34}$$

which follows from the fact that the coefficients X^a, Y^a are functions on the base manifold M, and thus do not depend on the vertical coordinate $\bar{x}^a$.

2. The mixed commutator is given by

$$\begin{aligned}
\left[\overset{\mathrm{V}}{X}, \overset{\mathrm{C}}{Y}\right] &= [X^a\bar{\partial}_a, Y^b\partial_b + \bar{x}^c\partial_c Y^b\bar{\partial}_b] \\
&= X^a\delta_a^c\partial_c Y^b\bar{\partial}_b - Y^b\partial_b X^a\bar{\partial}_a \\
&= (X^b\partial_b Y^a - Y^b\partial_b X^a)\bar{\partial}_a \\
&= [X,Y]^a\bar{\partial}_a \\
&= \overset{\mathrm{V}}{[X,Y]}\,.
\end{aligned} \tag{29.5.35}$$

3. Finally, for the complete lifts holds

$$\begin{aligned}
\left[\overset{\mathrm{C}}{X}, \overset{\mathrm{C}}{Y}\right] &= [X^a\partial_a + \bar{x}^c\partial_c X^a\bar{\partial}_a, Y^b\partial_b + \bar{x}^d\partial_d Y^b\bar{\partial}_b] \\
&= X^a\partial_a Y^b\partial_b + X^a\bar{x}^d\partial_a\partial_d Y^b\bar{\partial}_b + \bar{x}^c\partial_c X^a\delta_a^d\partial_d Y^b\bar{\partial}_b \\
&\quad - Y^b\partial_b X^a\partial_a - Y^b\bar{x}^c\partial_b\partial_c X^a\bar{\partial}_a - \bar{x}^d\partial_d Y^b\delta_b^c\partial_c X^a\bar{\partial}_a \\
&= (X^b\partial_b Y^a - Y^b\partial_b X^a)\partial_a + \bar{x}^c\partial_c(X^b\partial_b Y^a - Y^b\partial_b X^a)\bar{\partial}_a \\
&= [X,Y]^a\partial_a + \bar{x}^c\partial_c[X,Y]^a\bar{\partial}_a \\
&= \overset{\mathrm{C}}{[X,Y]}\,.
\end{aligned} \tag{29.5.36}$$

One may wonder whether the vertical and complete lifts of vector fields can be characterized by their relations with the tangent structure, similar to theorems 29.4.4 and 29.4.5 for the lifts of functions. It turns out that this is indeed the case. For the vertical lift, we find the following statement.

Theorem 29.5.6. *A vector field $X \in \text{Vect}(TM)$ is a vertical lift, $X = \overset{V}{Z}$ for some $Z \in \text{Vect}(M)$, if and only if $JX = 0$ and $\mathcal{L}_X J = 0$.*

Proof. ►...◄ ■

Similarly, for the complete lift the following relation holds.

Theorem 29.5.7. *A vector field $X \in \text{Vect}(TM)$ is the sum of a complete and a vertical lift, $X = \overset{C}{Y} + \overset{V}{Z}$ for some $Y, Z \in \text{Vect}(M)$, if and only if $\mathcal{L}_X J = 0$.*

Proof. ►...◄ ■

We illustrate these two statements using coordinates. Given a vector field $X = X^a\partial_a + \bar{X}^a\bar{\partial}_a \in \text{Vect}(TM)$ and an auxiliary vector field $V = V^a\partial_a + \bar{V}^a\bar{\partial}_a \in \text{Vect}(TM)$, we use theorem 16.6.1 to calculate the Lie derivative

$$\begin{aligned}(\mathcal{L}_X J)V &= [X, JV] - J[X, V] \\ &= [X^a\partial_a + \bar{X}^a\bar{\partial}_a, V^b\bar{\partial}_b] - J[X^a\partial_a + \bar{X}^a\bar{\partial}_a, V^b\partial_b + \bar{V}^b\bar{\partial}_b] \\ &= (X^b\partial_b V^a + \bar{X}^b\bar{\partial}_b V^a - V^b\bar{\partial}_b\bar{X}^a)\bar{\partial}_a - V^b\bar{\partial}_b X^a\partial_a \\ &\quad - (X^b\partial_b V^a + \bar{X}^b\bar{\partial}_b V^a - V^b\partial_b X^a - \bar{V}^b\bar{\partial}_b X^a)\bar{\partial}_a \\ &= V^b(\partial_b X^a - \bar{\partial}_b\bar{X}^a)\bar{\partial}_a + \bar{\partial}_b X^a(\bar{V}^b\bar{\partial}_a - V^b\partial_a)\,,\end{aligned} \tag{29.5.37}$$

so that

$$\mathcal{L}_X J = (\partial_b X^a - \bar{\partial}_b\bar{X}^a)\bar{\partial}_a \otimes \mathrm{d}x^b + \bar{\partial}_b X^a(\bar{\partial}_a \otimes \mathrm{d}\bar{x}^b - \partial_a \otimes \mathrm{d}x^b)\,. \tag{29.5.38}$$

Now this vanishes if and only if the coefficients in both terms vanish. For the second term, $\bar{\partial}_b X^a = 0$ means that the components X^a are constant along the fibers of TM, as they depend on the base manifold coordinates x^a only. Hence, $X^a = Y^a$ for some vector field $Y = Y^a\partial_a \in \text{Vect}(M)$. The condition from the first term then determines $\bar{X}^a$ to be of the form

$$\bar{X}^a = \bar{x}^b\partial_b Y^a + Z^a\,, \tag{29.5.39}$$

where also Z^a does not depend on $\bar{x}^a$, and so gives another vector field $Z = Z^a\partial_a \in \text{Vect}(M)$. By comparing with the coordinate expressions of the complete and vertical lifts, one sees that $X = \overset{C}{Y} + \overset{V}{Z}$. Further, $JX = 0$ if and only if X is vertical, hence $X^a = Y^a = 0$, whence $X = \overset{V}{Z}$.

Finally, we come to discuss the relation between the lifts of vector fields and those of functions. These come in two types, depending on whether we apply a vector field to a function, or multiply a vector field with a function. They are summarized in the following statement.

Theorem 29.5.8. *The vertical and complete lifts of functions and vector fields are related by*

$$\overset{C}{(Xf)} = \overset{C}{X}\overset{C}{f}\,, \quad \overset{V}{(Xf)} = \overset{V}{X}\overset{C}{f} = \overset{C}{X}\overset{V}{f}\,, \quad \overset{C}{(fX)} = \overset{C}{f}\overset{V}{X} + \overset{V}{f}\overset{C}{X}\,, \quad \overset{V}{(fX)} = \overset{V}{f}\overset{V}{X} \tag{29.5.40}$$

for any function $f \in C^\infty(M, \mathbb{R})$ and vector field $X \in \text{Vect}(M)$.

Proof. The first relation is simply the statement from theorem 29.5.2, which we repeat here for completeness, but do not need to prove again. Let $v \in TM$ and $g \in C^\infty(TM, \mathbb{R})$, and denote the flow of X by ϕ. Then the remaining propositions are proven as follows:

1. The complete lift of a vector field acts on the vertical lift of a function as

$$\begin{aligned}
\left(\overset{C}{X}\overset{V}{f}\right)(v) &= \frac{\mathrm{d}}{\mathrm{d}t}\overset{V}{f}(\phi_{t*}(v))\Big|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}f(\tau(\phi_{t*}(v)))\Big|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}f(\phi_t(\tau(v)))\Big|_{t=0} \\
&= (Xf)(\tau(v)) \\
&= (\overset{V}{Xf})(v)\,.
\end{aligned} \tag{29.5.41}$$

2. The vertical lift of a vector field acts on the complete lift of a function as

$$\begin{aligned}
\left(\overset{V}{X}\overset{C}{f}\right)(v) &= \frac{\mathrm{d}}{\mathrm{d}t}\overset{C}{f}(v + tX(\tau(v)))\Big|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}(v + tX(\tau(v)))f\Big|_{t=0} \\
&= X(\tau(v))f \\
&= (Xf)(\tau(v)) \\
&= (\overset{V}{Xf})(v)\,.
\end{aligned} \tag{29.5.42}$$

3. For the complete lift $(f\overset{C}{X})$ holds

$$(f\overset{C}{X})\overset{C}{g} = (f\overset{C}{X}g)\,. \tag{29.5.43}$$

For the latter, we can use the Leibniz rule, according to which holds

$$(f\overset{C}{X}g)(v) = v(fXg) = v(f)(Xg)(\tau(v)) + f(\tau(v))v(Xg) = \overset{C}{f}(v)(\overset{V}{X}g)(v) + \overset{V}{f}(v)(\overset{C}{X}g)(v)\,. \tag{29.5.44}$$

Using the previously proven propositions, we have

$$(\overset{V}{X}g) = \overset{V}{X}\overset{C}{g}\,, \quad (\overset{C}{X}g) = \overset{C}{X}\overset{C}{g}\,, \tag{29.5.45}$$

and so

$$(f\overset{C}{X})\overset{C}{g} = \left(\overset{C}{f}\overset{V}{X} + \overset{V}{f}\overset{C}{X}\right)\overset{C}{g}\,. \tag{29.5.46}$$

Since this holds for all functions g, the statement follows.

4. We can start from the previous proposition $(f\overset{C}{X}) = \overset{C}{f}\overset{V}{X} + \overset{V}{f}\overset{C}{X}$. Applying J on both sides yields

$$(f\overset{V}{X}) = J(f\overset{C}{X}) = \overset{C}{f}J\overset{V}{X} + \overset{V}{f}J\overset{C}{X} = \overset{V}{f}\overset{V}{X}\,, \tag{29.5.47}$$

where we used theorem 29.5.4 and the fact that J vanishes on vertical vector fields, whence $J\overset{V}{X} = 0$. ■

Also these formulas are easily illustrated in coordinates. Writing $X = X^a\partial_a \in \mathrm{Vect}(M)$, and hence $Xf = X^a\partial_a f$, we have the following formulas (for $(\overset{C}{X}f)$, see the coordinate calculation below theorem 29.5.2):

1. Acting with $\overset{C}{X}$ on $\overset{V}{f}$, we get

$$(X^a\partial_a + \bar{x}^a\partial_a X^b\bar{\partial}_b)(f\circ\tau) = (X^a\partial_a f)\circ\tau = (\overset{V}{Xf})\,. \tag{29.5.48}$$

2. Conversely, acting with $\overset{\mathrm{V}}{X}$ on $\overset{\mathrm{C}}{f}$, one finds

$$X^a\bar{\partial}_a(\bar{x}^b\partial_b(f\circ\tau)) = (X^a\partial_a f)\circ\tau = (\overset{\mathrm{V}}{Xf})\,. \tag{29.5.49}$$

3. The complete lift of fX reads

$$(\overset{\mathrm{C}}{fX}) = f(X^a\partial_a + \bar{x}^a\partial_a X^b\bar{\partial}_b) + \bar{x}^a\partial_a f X^b\bar{\partial}_b = \overset{\mathrm{V}}{f}\overset{\mathrm{C}}{X} + \overset{\mathrm{C}}{f}\overset{\mathrm{V}}{X}\,. \tag{29.5.50}$$

4. Finally, for the vertical lift of fX one has

$$(\overset{\mathrm{V}}{fX}) = fX^a\bar{\partial}_a = \overset{\mathrm{V}}{f}\overset{\mathrm{V}}{X}\,. \tag{29.5.51}$$

29.6 Lifts of covector fields

Definition 29.6.1 (Vertical lift of a covector field)**.** Let M be a manifold and $\omega\in\Omega^1(M)$ a covector field. Its *vertical lift* is the covector field $\overset{\mathrm{V}}{\omega}\in\Omega^1(TM)$ defined by the pullback

$$\overset{\mathrm{V}}{\omega} = \tau^*\omega\,. \tag{29.6.1}$$

Definition 29.6.2 (Complete lift of a covector field)**.** Let M be a manifold and $\omega\in\Omega^1(M)$ a covector field. Its *complete lift* is the covector field $\overset{\mathrm{C}}{\omega}\in\Omega^1(TM)$ defined by ►...◄

Theorem 29.6.1. *The vertical and complete lifts of a covector field $\omega\in\Omega^1(M)$ are related by $\overset{V}{\omega} = J^*\overset{C}{\omega}$.*

Proof. ►...◄ ■

Theorem 29.6.2. *For every function $f\in C^\infty(M,\mathbb{R})$, the vertical and complete lift of the total differential satisfy*

$$(\overset{V}{\mathrm{d}f}) = \mathrm{d}\overset{V}{f}\,,\quad (\overset{C}{\mathrm{d}f}) = \mathrm{d}\overset{C}{f}\,. \tag{29.6.2}$$

Proof. ►...◄ ■

Theorem 29.6.3. *For every function $f\in C^\infty(M,\mathbb{R})$ and covector field $\omega\in\Omega^1(M)$, the vertical and complete lift satisfy*

$$(\overset{V}{f\omega}) = \overset{V}{f}\overset{V}{\omega}\,,\quad (\overset{C}{f\omega}) = \overset{V}{f}\overset{C}{\omega} + \overset{C}{f}\overset{V}{\omega}\,. \tag{29.6.3}$$

Proof. ►...◄ ■

29.7 The canonical involution

Based on the constructions displayed in the previous sections, we now introduce a map, which we define implicitly by the following conditions.

Definition 29.7.1 (Canonical involution)**.** Let M be a manifold, $\tau : TM \to M$ its tangent bundle and $\varpi : TTM \to TM$ the double tangent bundle. The *canonical involution* of TTM is the unique diffeomorphism $\kappa : TTM \to TTM$ satisfying

1. $\tau_* \circ \kappa = \varpi$,
2. $\varpi \circ \kappa = \tau_*$,
3. $\kappa(\psi)(\overset{C}{f}) = \psi(\overset{C}{f})$ for all $v \in TTM$ and $f \in C^\infty(M, \mathbb{R})$.

►. . . ◄

It is instructive to calculate a coordinate expression for the canonical involution. For this purpose, consider a double tangent vector $\psi = \psi^a\partial_a + \bar{\psi}^a\bar{\partial}_a \in T_vTM$, where we assign coordinates $(x^a, \bar{x}^a)$ to the base point $\varpi(\psi) = v$. Hence, ψ is fully characterized by the coordinates $(x^a, \bar{x}^a, \psi^a, \bar{\psi}^a)$ on TTM. The bundle map ϖ, obviously, assigns to ψ the element $\varpi(\psi) = v \in TM$ with coordinates $(x^a, \bar{x}^a)$. However, we also have a second map which allows us to obtain an element of TM, namely the differential $\tau_* : TTM \to TM$ of the bundle map $\tau : TM \to M$. By construction of the induced coordinates, the latter gives the assignment $\tau : (x^a, \bar{x}^a) \mapsto (x^a)$. Hence, its differential acts on the basis vectors of TTM as

$$\tau_*(\partial_a) = \partial_a\,, \quad \tau_*(\bar{\partial}_a) = 0\,, \qquad \Rightarrow \qquad \tau_*(\psi) = \tau_*(\psi^a\partial_a + \bar{\psi}^a\bar{\partial}_a) = \psi^a\partial_a\,, \tag{29.7.1}$$

and so assigns to ψ the element of TM with coordinates (x^a, ψ^a). Further, we use the coordinate expression $\overset{C}{f}(x, \bar{x}) = \bar{x}^a\partial_a f(x)$ for the canonical lift of a function, from which follows

$$\psi(\overset{C}{f}) = \psi^a\bar{x}^b\partial_a\partial_b f(x) + \bar{\psi}^a\partial_a f(x)\,. \tag{29.7.2}$$

Finally, writing the image $\kappa(\psi)$ in coordinates as $(x'^a, \bar{x}'^a, \psi'^a, \bar{\psi}'^a)$, we have the following relations:

1. From $\tau_* \circ \kappa = \varpi$ follows $(x'^a, \psi'^a) = (x^a, \bar{x}^a)$.
2. From $\varpi \circ \kappa = \tau_*$ follows $(x'^a, \bar{x}'^a) = (x^a, \psi^a)$.
3. From $\kappa(\psi)(\overset{C}{f}) = \psi(\overset{C}{f})$ follows $\psi'^{(a}\bar{x}'^{b)} = \psi^{(a}\bar{x}^{b)}$ (since partial derivatives commute) and $\bar{\psi}'^a = \bar{\psi}^a$.

From the first two conditions thus follows $x'^a = x^a$, $\bar{x}'^a = \psi^a$ and $\psi'^a = \bar{x}^a$. It thus follows that $\psi'^a\bar{x}'^b = \psi^b\bar{x}^a$, and thus also $\psi'^{(a}\bar{x}'^{b)} = \psi^{(a}\bar{x}^{b)}$ is already satisfied. Finally, we have $\bar{\psi}'^a = \bar{\psi}^a$, so that we can write the canonical involution as

$$\kappa : (x^a, \bar{x}^a, \psi^a, \bar{\psi}^a) \mapsto (x^a, \psi^a, \bar{x}^a, \bar{\psi}^a)\,. \tag{29.7.3}$$

With this knowledge, we can study the properties of κ. We start with the following, which is already suggested by the name we introduced.

Theorem 29.7.1. *The canonical involution $\kappa : TTM \to TTM$ is an involution, $\kappa \circ \kappa = \mathrm{id}_{TTM}$.*

Proof. ►. . . ◄ ■

Theorem 29.7.2. *The canonical involution is a vector bundle isomorphism from $\varpi : TTM \to TM$ to $\tau_* : TTM \to TM$ and vice versa, covering the identity on TM.*

Proof. ►. . . ◄ ■

Theorem 29.7.3. *The canonical involution satisfies* $\kappa \circ \ddot{\gamma} = \ddot{\gamma}$ *for the second canonical lift* $\ddot{\gamma} \in C^\infty(\mathbb{R}, TTM)$ *of every curve* $\gamma \in C^\infty(\mathbb{R}, M)$.

Proof. ►. . . ◄ ■

Chapter 30

Affine connections

30.1 Frame bundle connections

In the chapters 26, 27 and 28 we have discussed different descriptions of connections on various types of bundles. The most important class of bundles we have encountered so far, since they are canonically defined for any manifold, are constructed as associated bundles to the tangent frame bundle, and are given by tensor density bundles and their special cases, including tensor bundles and scalar density bundles. There are different, equivalent possibilities to define connections on these bundles. Here we choose to start from the principal bundle, since every other bundle we will discuss in this chapter can be obtained from the frame bundle. We will thus use the following definition as the starting point.

Definition 30.1.1 (Affine connection)**.** Let M be a manifold of dimension n. An *affine connection* is a principal $\mathrm{GL}(n,\mathbb{R})$ connection on the general linear frame bundle FM.

Reminding definition 27.1.2 of a principal G-connection, an affine connection is therefore a $\mathfrak{gl}(n,\mathbb{R})$-valued one-form $\vartheta \in \Omega^1(FM, \mathfrak{gl}(n,\mathbb{R}))$ on the frame bundle FM, which is equivariant and reverses the operation $X \mapsto \tilde{X}$ which assigns to $X \in \mathfrak{gl}(n,\mathbb{R})$ the fundamental vector field $\tilde{X} \in \mathrm{Vect}(FM)$.

We illustrate the definition using the coordinates $(x^\mu, p^\mu{}_i)$ on FM which we introduced in section 22.6, where (x^μ) are coordinates on the base manifold M. A general $\mathfrak{gl}(n,\mathbb{R})$-valued one-form $\vartheta \in \Omega^1(FM, \mathfrak{gl}(n,\mathbb{R}))$ takes the form

$$\vartheta = (\vartheta^i{}_{j\mu}\mathrm{d}x^\mu + \bar{\vartheta}^i{}_{j\mu}{}^k\mathrm{d}p^\mu{}_k) \otimes \mathcal{H}_i{}^j\,, \tag{30.1.1}$$

where $\mathcal{H}_i{}^j$ denotes the basis of the Lie algebra $\mathfrak{gl}(n,\mathbb{R})$. In order for this to be a principal connection, we have two conditions. First, we must demand that on the fundamental vector fields, which generate the right translations, it must recover the generating Lie algebra element. Using the expression (22.6.6) for the fundamental vector fields, we have

$$\vartheta(\tilde{a}) = \bar{\vartheta}^i{}_{j\mu}{}^k p^\mu{}_l a^l{}_k \mathcal{H}_i{}^j\,. \tag{30.1.2}$$

Demanding that this is equal to $a^i{}_j\mathcal{H}_i{}^j$, we thus have the condition

$$\bar{\vartheta}^i{}_{j\mu}{}^k p^\mu{}_l = \delta^i_l \delta^k_j\,, \tag{30.1.3}$$

and thus

$$\bar{\vartheta}^i{}_{j\mu}{}^k = p^{-1}_\mu{}^i \delta^k_j\,, \tag{30.1.4}$$

so that ϑ becomes

$$\vartheta = (\vartheta^i{}_{j\mu}\mathrm{d}x^\mu + p^{-1\,i}{}_\mu \mathrm{d}p^\mu{}_j) \otimes \mathcal{H}_i{}^j\,. \tag{30.1.5}$$

The second condition is that of equivariance. Denoting a group element $g \in \mathrm{GL}(n,\mathbb{R})$ by the matrix components $g^i{}_j$, we can use the expression (22.6.4) for the right translation to calculate the pullback

$$R_g^*\vartheta = (\vartheta'^i{}_{j\mu}\mathrm{d}x^\mu + g^{-1\,i}{}_k p^{-1\,k}{}_\mu g^l{}_j \mathrm{d}p^\mu{}_l) \otimes \mathcal{H}_i{}^j\,, \tag{30.1.6}$$

where $\vartheta'^i{}_{j\mu}$ indicates that we must evaluate $\vartheta^i{}_{j\mu}$ at $p \cdot g$ instead of p. Further, we have the adjoint representation

$$\mathrm{Ad}_g(\mathcal{H}_i{}^j) = g^k{}_i g^{-1\,j}{}_l \mathcal{H}_k{}^l\,, \tag{30.1.7}$$

and thus

$$\begin{aligned}\mathrm{Ad}_g \circ R_g^*\vartheta &= (\vartheta'^i{}_{j\mu}\mathrm{d}x^\mu + g^{-1\,i}{}_k p^{-1\,k}{}_\mu g^l{}_j \mathrm{d}p^\mu{}_l) \otimes g^m{}_i g^{-1\,j}{}_n \mathcal{H}_m{}^n\\ &= (g^i{}_k g^{-1\,l}{}_j \vartheta'^k{}_{l\mu}\mathrm{d}x^\mu + p^{-1\,i}{}_\mu \mathrm{d}p^\mu{}_j) \otimes \mathcal{H}_i{}^j\,.\end{aligned} \tag{30.1.8}$$

Note that this agrees with ϑ if and only if the components $\vartheta^i{}_{j\mu}$ and $\vartheta'^i{}_{j\mu}$ at p and $p \cdot g$ are related by

$$\vartheta^i{}_{j\mu} = g^i{}_k g^{-1\,l}{}_j \vartheta'^k{}_{l\mu}\,. \tag{30.1.9}$$

To solve this condition, it is useful to introduce the coefficients $\Gamma^\nu{}_{\rho\mu}$ as

$$\vartheta^i{}_{j\mu} = p^{-1\,i}{}_\nu p^\rho{}_j \Gamma^\nu{}_{\rho\mu} \quad \Leftrightarrow \quad \Gamma^\nu{}_{\rho\mu} = p^\nu{}_i p^{-1\,j}{}_\rho \vartheta^i{}_{j\mu}\,, \tag{30.1.10}$$

so that we have

$$\vartheta'^i{}_{j\mu} = g^{-1\,i}{}_k p^{-1\,k}{}_\nu p^\rho{}_l g^l{}_j \Gamma'^\nu{}_{\rho\mu} \quad \Leftrightarrow \quad \Gamma'^\nu{}_{\rho\mu} = p^\nu{}_k g^k{}_i g^{-1\,j}{}_l p^{-1\,l}{}_\rho \vartheta'^i{}_{j\mu}\,. \tag{30.1.11}$$

With this definition, the equivariance condition simply becomes

$$\Gamma'^\nu{}_{\rho\mu} = \Gamma^\nu{}_{\rho\mu}\,, \tag{30.1.12}$$

which means that these coefficients must be independent of $p \in F_xM$, and can depend on $x \in M$ only. This allows us to finally write the connection as

$$\vartheta = p^{-1\,i}{}_\mu (p^\nu{}_j \Gamma^\mu{}_{\nu\rho}\mathrm{d}x^\rho + \mathrm{d}p^\mu{}_j) \otimes \mathcal{H}_i{}^j\,, \tag{30.1.13}$$

where we relabeled indices to pull out a common factor in the front. We call $\Gamma^\mu{}_{\nu\rho}$ the *connection coefficients*.

30.2 Linear connection in the tangent bundle

Recall from section 28.7 that a connection on a principal bundle induces a Koszul connection on any associated vector bundle, which we constructed explicitly in theorem 28.7.2. As discussed in chapter 22, we can obtain the tangent bundle as such an associated vector bundle by using the canonical representation ρ of the structure group $\mathrm{GL}(n,\mathbb{R})$ on $\mathbb{R}^n$. Hence, an affine connection as defined in section 30.1 as a connection on the frame bundle can also be represented as a Koszul connection on the tangent bundle.

To illustrate the construction of the Koszul connection, we derive its coordinate expression from the coordinate expression (30.1.13) of the connection on the frame bundle. Let $X = X^\mu\partial_\mu \in \mathrm{Vect}(M)$ be a vector field on M. Following theorem 27.2.1 (which reduces to theorem 20.3.3 in this case) we can identify X with the basic zero-form $\hat{X} \in \Omega^0(FM, \mathbb{R}^n)$ given by

$$\hat{X} = X^\mu p^{-1\,i}{}_\mu \mathbb{e}_i\,. \tag{30.2.1}$$

We then calculate its exterior covariant derivative, which takes the form

$$\begin{aligned}
\mathrm{d}_\vartheta\hat{X} &= \mathrm{d}\hat{X} + \rho_*(\vartheta)\wedge\hat{X} \\
&= (\partial_\nu X^\mu p^{-1\,i}{}_\mu \mathrm{d}x^\nu - X^\mu p^{-1\,i}{}_\nu p^{-1\,j}{}_\mu \mathrm{d}p^\nu{}_j)\otimes \mathbb{e}_i + p^{-1\,i}{}_\mu (p^\nu{}_j \Gamma^\mu{}_{\nu\rho}\mathrm{d}x^\rho + \mathrm{d}p^\mu{}_j)\otimes X^\sigma p^{-1\,j}{}_\sigma \mathbb{e}_i \\
&= p^{-1\,i}{}_\mu(\partial_\rho X^\mu + \Gamma^\mu{}_{\nu\rho}X^\nu)\mathrm{d}x^\rho \otimes \mathbb{e}_i\,,
\end{aligned} \tag{30.2.2}$$

where we used the fact that the canonical representation ρ_* of the Lie algebra $\mathfrak{gl}(n,\mathbb{R}^n)$ is simply its matrix representation which we used also in the coordinate expression (30.1.13), and which takes the form

$$\rho_*(\mathcal{H}_i{}^j)\mathbb{e}_k = \delta^j_k \mathbb{e}_i\,, \tag{30.2.3}$$

so that

$$\rho_*(a^i{}_j \mathcal{H}_i{}^j)(v^k \mathbb{e}_k) = a^i{}_j v^j \mathbb{e}_i\,. \tag{30.2.4}$$

Following theorem 27.2.2, the result $\mathrm{d}_\vartheta\hat{X}$ is a basic one-form, and so we can apply theorem 27.2.1 again to obtain a TM-valued one-form on M, which is given by

$$\nabla X = (\partial_\rho X^\mu + \Gamma^\mu{}_{\nu\rho}X^\nu)\mathrm{d}x^\rho \otimes \partial_\mu\,, \tag{30.2.5}$$

which completes the construction of the Koszul connection. Hence, we see that the coefficients of the latter are simply the coefficients $\Gamma^\mu{}_{\nu\rho}$ of the affine connection which we introduced for the connection on the frame bundle.

Given a Koszul connection, we can make use of all constructions shown for linear connections in chapter 28 apply. In particular, it allows us to define a covariant derivative, which we can now write in coordinates as

$$\nabla_X Y = X^\mu(\partial_\mu Y^\nu + \Gamma^\nu{}_{\rho\mu}Y^\rho)\partial_\nu = X^\mu(\nabla_\mu Y^\nu)\partial_\nu\,, \tag{30.2.6}$$

where we used the abbreviation

$$\nabla_\mu Y^\nu = \partial_\mu Y^\nu + \Gamma^\nu{}_{\rho\mu}Y^\rho\,. \tag{30.2.7}$$

We can also apply the same construction to other vector bundles which are associated to the frame bundle, derive the corresponding Koszul connections and covariant derivatives, again following the procedure outlined for arbitrary vector bundles in chapter 28. In particular, we then have for a covector field $\alpha = \alpha_\mu \mathrm{d}x^\mu \in \Omega^1(M)$ the relation

$$\nabla_\mu \alpha_\nu = \partial_\mu \alpha_\nu - \Gamma^\rho{}_{\nu\mu}\alpha_\rho\,, \tag{30.2.8}$$

for an arbitrary tensor field $A \in \Gamma(T^r_s M)$ of rank (r,s) the relation

$$\begin{aligned}
\nabla_\mu A^{\nu_1\cdots\nu_r}{}_{\rho_1\cdots\rho_s} = \partial_\mu A^{\nu_1\cdots\nu_r}{}_{\rho_1\cdots\rho_s}& \\
+ \Gamma^{\nu_1}{}_{\sigma\mu}A^{\sigma\nu_2\cdots\nu_r}{}_{\rho_1\cdots\rho_s} + \ldots + \Gamma^{\nu_r}{}_{\sigma\mu}A^{\nu_1\cdots\nu_{r-1}\sigma}{}_{\rho_1\cdots\rho_s}& \\
- \Gamma^\sigma{}_{\rho_1\mu}A^{\nu_1\cdots\nu_r}{}_{\sigma\rho_2\cdots\rho_s} - \ldots - \Gamma^\sigma{}_{\rho_s\mu}A^{\nu_1\cdots\nu_r}{}_{\rho_1\cdots\rho_{s-1}\sigma}&\,,
\end{aligned} \tag{30.2.9}$$

and for a tensor density $\mathfrak{A}$ of weight w the relation

$$\begin{aligned}
\nabla_\mu \mathfrak{A}^{\nu_1\cdots\nu_r}{}_{\rho_1\cdots\rho_s} = \partial_\mu \mathfrak{A}^{\nu_1\cdots\nu_r}{}_{\rho_1\cdots\rho_s} - w\Gamma^\sigma{}_{\sigma\mu}\mathfrak{A}^{\nu_1\cdots\nu_r}{}_{\rho_1\cdots\rho_s}& \\
+ \Gamma^{\nu_1}{}_{\sigma\mu}\mathfrak{A}^{\sigma\nu_2\cdots\nu_r}{}_{\rho_1\cdots\rho_s} + \ldots + \Gamma^{\nu_r}{}_{\sigma\mu}\mathfrak{A}^{\nu_1\cdots\nu_{r-1}\sigma}{}_{\rho_1\cdots\rho_s}& \\
- \Gamma^\sigma{}_{\rho_1\mu}\mathfrak{A}^{\nu_1\cdots\nu_r}{}_{\sigma\rho_2\cdots\rho_s} - \ldots - \Gamma^\sigma{}_{\rho_s\mu}\mathfrak{A}^{\nu_1\cdots\nu_r}{}_{\rho_1\cdots\rho_{s-1}\sigma}&\,,
\end{aligned} \tag{30.2.10}$$

It should be noted that an affine connection allows for a few additional constructions which are not defined for linear connections on arbitrary vector bundles, and which we discuss in detail in the following sections.

30.3 Curvature

The first property which we discuss and extend for the case of affine connections is the curvature, which we discussed for linear connections in general in section 28.12. Since it is conventional to use the symbol R instead of F for the curvature in this case, we provide the following definition.

Definition 30.3.1 (Curvature of an affine connection)**.** Let M be a manifold equipped with a connection ∇. Its *curvature* is the endomorphism-valued two-form $R \in \Omega^2(M, \mathrm{End}(TM))$ defined by

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z \tag{30.3.1}$$

for all vector fields $X, Y, Z \in \mathrm{Vect}(M)$.

It is helpful to derive the curvature also in coordinates. First, we follow the definition and calculate

$$\begin{aligned} R(X,Y)Z &= \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z \\ &= \nabla_{X^\mu \partial_\mu} \nabla_{Y^\nu \partial_\nu} (Z^\rho \partial_\rho) - \nabla_{Y^\nu \partial_\nu} \nabla_{X^\mu \partial_\mu} (Z^\rho \partial_\rho) - \nabla_{[X^\mu \partial_\mu, Y^\nu \partial_\nu]} (Z^\rho \partial_\rho) \\ &= \big(X^\mu \{\partial_\mu [Y^\nu (\partial_\nu Z^\rho + \Gamma^\rho{}_{\sigma\nu} Z^\sigma)] + \Gamma^\rho{}_{\lambda\mu} [Y^\nu (\partial_\nu Z^\lambda + \Gamma^\lambda{}_{\sigma\nu} Z^\sigma)]\} \\ &\quad - Y^\nu \{\partial_\nu [X^\mu (\partial_\mu Z^\rho + \Gamma^\rho{}_{\sigma\mu} Z^\sigma)] + \Gamma^\rho{}_{\lambda\nu} [X^\mu (\partial_\mu Z^\lambda + \Gamma^\lambda{}_{\sigma\mu} Z^\sigma)]\} \\ &\quad - (X^\mu \partial_\mu Y^\nu - Y^\mu \partial_\mu X^\nu)(\partial_\nu Z^\rho + \Gamma^\rho{}_{\sigma\nu} Z^\sigma)\big) \partial_\rho \\ &= X^\mu Y^\nu (\partial_\mu \Gamma^\rho{}_{\sigma\nu} - \partial_\nu \Gamma^\rho{}_{\sigma\mu} + \Gamma^\rho{}_{\lambda\mu} \Gamma^\lambda{}_{\sigma\nu} - \Gamma^\rho{}_{\lambda\nu} \Gamma^\lambda{}_{\sigma\mu}) Z^\sigma \partial_\rho \,. \end{aligned} \tag{30.3.2}$$

Hence, it follows that the components of the curvature are given by

$$R^\rho{}_{\sigma\mu\nu} = \partial_\mu \Gamma^\rho{}_{\sigma\nu} - \partial_\nu \Gamma^\rho{}_{\sigma\mu} + \Gamma^\rho{}_{\lambda\mu} \Gamma^\lambda{}_{\sigma\nu} - \Gamma^\rho{}_{\lambda\nu} \Gamma^\lambda{}_{\sigma\mu} \,. \tag{30.3.3}$$

These components allow to express the curvature in terms of different interpretations. Following our definition 30.3.1, we can write it as an endomorphism-valued two-form

$$R = \frac{1}{2} R^\rho{}_{\sigma\mu\nu} \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \otimes \partial_\rho \otimes \mathrm{d}x^\sigma \,. \tag{30.3.4}$$

It is also conventional, and historically abundant in the literature, to define write it as the $(1,3)$-tensor field

$$R = R^\rho{}_{\sigma\mu\nu} \partial_\rho \otimes \mathrm{d}x^\sigma \otimes \mathrm{d}x^\mu \otimes \mathrm{d}x^\nu \,, \tag{30.3.5}$$

where the different numerical factor is related to the fact that here we use a tensor product $\mathrm{d}x^\mu \otimes \mathrm{d}x^\nu$ instead of the exterior product $\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu$. Here we make use of either interpretation, depending on which is more convenient.

Clearly, R possesses all the properties of the curvature F of a linear connection detailed in theorem 28.12.1. Further, following theorem 28.12.2 it is closely related to the curvature $\Omega \in \Omega^2(FM, \mathfrak{gl}(n, \mathbb{R}))$ of the principal connection. In the case of the tangent bundle, the latter can also be expressed as follows.

Theorem 30.3.1. *The curvature form and the curvature tensor of an affine connection are related by*

$$p \circ \Omega_p \circ p^{-1} = (\varpi^* R)_p \tag{30.3.6}$$

for all $p \in FM$.

Proof. ▶...◀ ■

Also this relation can easily be illustrated using coordinates. For this purpose, we first use the coordinate expression (30.1.13) of the principal connection to calculate the curvature form

$$\begin{aligned}
\Omega &= \mathrm{d}\vartheta + \frac{1}{2}[\vartheta \wedge \vartheta] \\
&= -p^{-1\,i}{}_{\sigma}p^{-1\,k}{}_{\mu}\mathrm{d}p^{\sigma}{}_{k} \wedge (p^{\nu}{}_{j}\Gamma^{\mu}{}_{\nu\rho}\mathrm{d}x^{\rho} + \mathrm{d}p^{\mu}{}_{j}) \otimes \mathcal{H}_i{}^j \\
&\quad + p^{-1\,i}{}_{\mu}(\Gamma^{\mu}{}_{\nu\rho}\mathrm{d}p^{\nu}{}_{j} + p^{\nu}{}_{j}\partial_{\sigma}\Gamma^{\mu}{}_{\nu\rho}\mathrm{d}x^{\sigma}) \wedge \mathrm{d}x^{\rho} \otimes \mathcal{H}_i{}^j \\
&\quad + \frac{1}{2}p^{-1\,i}{}_{\mu}p^{-1\,k}{}_{\sigma}(p^{\nu}{}_{j}\Gamma^{\mu}{}_{\nu\rho}\mathrm{d}x^{\rho} + \mathrm{d}p^{\mu}{}_{j}) \wedge (p^{\tau}{}_{l}\Gamma^{\sigma}{}_{\tau\omega}\mathrm{d}x^{\omega} + \mathrm{d}p^{\sigma}{}_{l}) \otimes [\mathcal{H}_i{}^j, \mathcal{H}_k{}^l] \\
&= \big[p^{-1\,i}{}_{\mu}(\Gamma^{\mu}{}_{\nu\rho}\mathrm{d}p^{\nu}{}_{j} + p^{\nu}{}_{j}\partial_{\sigma}\Gamma^{\mu}{}_{\nu\rho}\mathrm{d}x^{\sigma}) \wedge \mathrm{d}x^{\rho} - p^{-1\,i}{}_{\sigma}p^{-1\,k}{}_{\mu}\mathrm{d}p^{\sigma}{}_{k} \wedge (p^{\nu}{}_{j}\Gamma^{\mu}{}_{\nu\rho}\mathrm{d}x^{\rho} + \mathrm{d}p^{\mu}{}_{j}) \\
&\quad + p^{-1\,i}{}_{\mu}p^{-1\,k}{}_{\sigma}(p^{\nu}{}_{k}\Gamma^{\mu}{}_{\nu\rho}\mathrm{d}x^{\rho} + \mathrm{d}p^{\mu}{}_{k}) \wedge (p^{\tau}{}_{j}\Gamma^{\sigma}{}_{\tau\omega}\mathrm{d}x^{\omega} + \mathrm{d}p^{\sigma}{}_{j})\big] \otimes \mathcal{H}_i{}^j \\
&= \big[p^{-1\,i}{}_{\mu}(\Gamma^{\mu}{}_{\nu\rho}\mathrm{d}p^{\nu}{}_{j} + p^{\nu}{}_{j}\partial_{\sigma}\Gamma^{\mu}{}_{\nu\rho}\mathrm{d}x^{\sigma}) \wedge \mathrm{d}x^{\rho} - p^{-1\,i}{}_{\sigma}p^{-1\,k}{}_{\mu}\mathrm{d}p^{\sigma}{}_{k} \wedge (p^{\nu}{}_{j}\Gamma^{\mu}{}_{\nu\rho}\mathrm{d}x^{\rho} + \mathrm{d}p^{\mu}{}_{j}) \\
&\quad + p^{-1\,i}{}_{\mu}(\Gamma^{\mu}{}_{\sigma\rho}\mathrm{d}x^{\rho} + p^{-1\,k}{}_{\sigma}\mathrm{d}p^{\mu}{}_{k}) \wedge (p^{\tau}{}_{j}\Gamma^{\sigma}{}_{\tau\omega}\mathrm{d}x^{\omega} + \mathrm{d}p^{\sigma}{}_{j})\big] \otimes \mathcal{H}_i{}^j \\
&= p^{-1\,i}{}_{\mu}p^{\nu}{}_{j}(\partial_{\rho}\Gamma^{\mu}{}_{\nu\sigma} + \Gamma^{\mu}{}_{\lambda\rho} \wedge \Gamma^{\lambda}{}_{\nu\sigma})\mathrm{d}x^{\rho} \wedge \mathrm{d}x^{\sigma} \otimes \mathcal{H}_i{}^j \\
&= \frac{1}{2}p^{-1\,i}{}_{\mu}p^{\nu}{}_{j}R^{\mu}{}_{\nu\rho\sigma}\mathrm{d}x^{\rho} \wedge \mathrm{d}x^{\sigma} \otimes \mathcal{H}_i{}^j\,,
\end{aligned} \tag{30.3.7}$$

where we used the Lie algebra relation

$$[\mathcal{H}_i{}^j, \mathcal{H}_k{}^l] = \delta^j_k\mathcal{H}_i{}^l - \delta^l_i\mathcal{H}_k{}^j\,. \tag{30.3.8}$$

Hence, we find again the components of the curvature.

We have seen in section 28.12 that the curvature $F \in \Omega^2(M, \mathrm{End}(E))$ of a linear connection is an endomorphism-valued two-form. Hence, together with two vector fields $X, Y \in \mathrm{Vect}(M)$, it defines an endomorphism-valued function $\iota_Y\iota_X F = F(X,Y) \in \Omega^0(M, \mathrm{End}(E))$. This obviously also holds for the curvature $R \in \Omega^2(M, \mathrm{End}(TM))$ of an affine connection. However, in this case there exists also another possibility to define an endomorphism-valued function $Z \mapsto R(Z,Y)X$, by letting $R(\bullet, Y)$ act on X instead. The trace of this function has a particular role.

Definition 30.3.2 (Ricci curvature)**.** Let M be a manifold equipped with an affine connection ∇. The *Ricci tensor* $R\!\!\!/ \in \Gamma(T^0_2M)$ is defined as the trace

$$R\!\!\!/(X,Y) = \mathrm{tr}(R(\bullet, Y)X) \tag{30.3.9}$$

of the endomorphism $\bullet \mapsto R(\bullet, Y)X$.

In coordinates, we have

$$R(\bullet, Y)X = R^{\rho}{}_{\mu\sigma\nu}X^{\mu}Y^{\nu}\partial_{\rho} \otimes \mathrm{d}x^{\sigma}\,, \tag{30.3.10}$$

and thus

$$R\!\!\!/(X,Y) = R^{\rho}{}_{\mu\rho\nu}X^{\mu}Y^{\nu}\,, \tag{30.3.11}$$

so that we write the Ricci tensor as the $(0,2)$-tensor

$$R\!\!\!/ = R^{\rho}{}_{\mu\rho\nu}\mathrm{d}x^{\mu} \otimes \mathrm{d}x^{\nu} = R_{\mu\nu}\mathrm{d}x^{\mu} \otimes \mathrm{d}x^{\nu}\,. \tag{30.3.12}$$

Naturally the question arises whether this tensor has any particular symmetry. The answer to this question is given by the following statement.

Theorem 30.3.2. *The Ricci tensor is symmetric if and only if there exists locally a covariantly constant volume form.*

Proof. ▶. . . ◀ ■

30.4 Torsion

Due to the particular properties of the tangent bundle, an affine connection can be characterized by another tensor field, which is not defined for a general linear connection on a vector bundle. Recall from definition 22.6.7 that the frame bundle of a manifold M of dimension n is equipped with a canonical $\mathbb{R}^n$-valued one-form $\theta \in \Omega^1(FM, \mathbb{R}^n)$, which is basic according to theorem 22.6.5. Given a principal connection on the frame bundle, we may thus calculate the exterior covariant derivative of the canonical one-form, as given in definition 27.2.2. This yields the following object.

Definition 30.4.1 (Torsion form)**.** Let $\vartheta \in \Omega^1(FM, \mathfrak{gl}(n, \mathbb{R}))$ be an affine connection on a manifold M. The *torsion form* of ϑ is given by

$$\Theta = \mathrm{d}_\vartheta \theta = \mathrm{d}\theta + \vartheta \wedge \theta \in \Omega^2(FM, \mathbb{R}^n)\,. \tag{30.4.1}$$

We have already seen for the curvature that in addition to the description in terms of an equivariant, vector-valued form on the frame bundle, also a dual description as a tensor field on the base manifold exists. This is also the case for the torsion, which can equivalently be described in terms of a tensor field, or a vector-valued two-form on the base manifold. Its definition in terms of the covariant derivative is given as follows.

Definition 30.4.2 (Torsion of an affine connection)**.** Let M be a manifold equipped with a connection ∇. Its *torsion* is the vector-valued two-form $T \in \Omega^2(M, TM)$ defined by

$$T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y] \tag{30.4.2}$$

for all vector fields $X, Y \in \mathrm{Vect}(M)$.

Note that this definition relies on the fact that an affine connection is not simply a Koszul connection on an arbitrary vector bundle, but on the tangent bundle in particular, so that one can take the covariant derivative of vector fields. Note that in order for T to be a vector-valued two-form as claimed in the definition, it needs to satisfy a number of properties, which we show next.

Theorem 30.4.1. *For all $X, Y, Z \in \mathrm{Vect}(M)$ and $f \in C^\infty(M, \mathbb{R})$, the torsion F of an affine connection:*

$$T(Y, X) = -T(X, Y)\,, \tag{30.4.3a}$$

$$T(X + Y, Z) = T(X, Z) + T(Y, Z)\,, \tag{30.4.3b}$$

$$T(fX, Y) = fT(X, Y)\,. \tag{30.4.3c}$$

Proof. The first two properties are an immediate consequence of the linearity of the covariant derivative, as well as the linearity and antisymmetry of the Lie bracket. For the last condition, one obtains by direct calculation

$$\begin{aligned} T(fX, Y) &= \nabla_{fX} Y - \nabla_Y (fX) - [fX, Y] \\ &= f\nabla_X Y - f\nabla_Y X - (Yf)X - f[X, Y] + (Yf)X \\ &= f(\nabla_X Y - \nabla_Y X - [X, Y]) \\ &= fT(X, Y)\,. \quad \blacksquare \end{aligned} \tag{30.4.4}$$

From its definition, one can now easily derive a coordinate expression for the torsion. For this purpose, we calculate

$$\begin{aligned} T(X,Y) &= \nabla_X Y - \nabla_Y X - [X,Y] \\ &= \nabla_{X^\mu\partial_\mu}(Y^\nu\partial_\nu) - \nabla_{Y^\nu\partial_\nu}(X^\mu\partial_\mu) - [X^\mu\partial_\mu, Y^\nu\partial_\nu] \\ &= X^\mu(\partial_\mu Y^\nu + \Gamma^\nu{}_{\rho\mu}Y^\rho)\partial_\nu - X^\mu(\partial_\mu Y^\nu + \Gamma^\mu{}_{\rho\nu}Y^\rho)\partial_\mu - X^\mu\partial_\mu Y^\nu\partial_\nu + Y^\nu\partial_\nu X^\mu\partial_\mu \\ &= X^\mu Y^\nu(\Gamma^\rho{}_{\nu\mu} - \Gamma^\rho{}_{\mu\nu})\partial_\rho\,. \end{aligned} \tag{30.4.5}$$

Hence, in the coordinate basis, the torsion is written in the form

$$T = \frac{1}{2}T^\rho{}_{\mu\nu}\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \otimes \partial_\rho \tag{30.4.6}$$

with

$$T^\rho{}_{\mu\nu} = \Gamma^\rho{}_{\nu\mu} - \Gamma^\rho{}_{\mu\nu}\,. \tag{30.4.7}$$

We see that the torsion tensor is obtained from the part of the connection coefficients which is antisymmetrized in its lower indices. In particular, the torsion vanishes if and only if the connection coefficients are symmetric in these two indices. This justifies the following naming for this case.

Definition 30.4.3 (Symmetric affine connection)**.** An affine connection is called *symmetric* if and only if $T(X,Y) = 0$ for all $X, Y \in \mathrm{Vect}(M)$.

As with the curvature, also for the torsion the question arises how the two descriptions given on the frame bundle and as a tensor field are related to each other. This relation takes the following form.

Theorem 30.4.2. *The torsion form and the torsion tensor of an affine connection are related by*

$$p \circ \Theta_p = (\varpi^* T)_p \tag{30.4.8}$$

for all $p \in FM$.

Proof. ▶. . . ◀ ■

We also derive the torsion form in coordinates, and find that

$$\begin{aligned} \Theta &= \mathrm{d}\theta + \vartheta \wedge \theta \\ &= -p^{-1\,i}{}_\nu p^{-1\,j}{}_\mu \mathrm{d}p^\nu{}_j \wedge \mathrm{d}x^\mu \otimes \mathbb{e}_i + p^{-1\,i}{}_\mu(p^\nu{}_j\Gamma^\mu{}_{\nu\rho}\mathrm{d}x^\rho + \mathrm{d}p^\mu{}_j) \wedge (p^{-1\,j}{}_\sigma \mathrm{d}x^\sigma) \otimes \mathbb{e}_i \\ &= p^{-1\,i}{}_\mu \Gamma^\mu{}_{\nu\rho}\mathrm{d}x^\rho \wedge \mathrm{d}x^\nu \otimes \mathbb{e}_i \\ &= \frac{1}{2}p^{-1\,i}{}_\mu T^\mu{}_{\nu\rho}\mathrm{d}x^\nu \wedge \mathrm{d}x^\rho \otimes \mathbb{e}_i\,, \end{aligned} \tag{30.4.9}$$

which reproduces the components of the torsion tensor.

30.5 Bianchi identities

We have seen for different types of connections that the curvature satisfies a particular relation, known as the Bianchi identity - for a general connection in theorem 26.10.3, for a principal connection in theorem 27.3.4 and for a linear connection in theorem 28.13.3. In the case of an affine connection, one finds another, similar relation, which concerns the exterior covariant derivative of the torsion form. It is conventional to denote this set of relations, given below, as first and second Bianchi identity.

Theorem 30.5.1 (Bianchi identities)**.** *The curvature form Ω and torsion form Θ of an affine connection satisfy the Bianchi identities*

$$\mathrm{d}_\vartheta\Theta = \Omega \wedge \theta\,, \quad \mathrm{d}_\vartheta\Omega = 0\,. \tag{30.5.1}$$

Proof. From the definition 30.4.1 of the torsion form, together with the relation 27.3.5 and the fact that the canonical one-form is basic of type $\rho = \mathrm{id}_{\mathrm{GL}(n,\mathbb{R})}$ follows the first Bianchi identity

$$\mathrm{d}_\vartheta\Theta = \mathrm{d}_\vartheta\mathrm{d}_\vartheta\theta = \Omega \wedge \theta\,. \tag{30.5.2}$$

The second Bianchi identity is simply the Bianchi identity 27.3.4 for a principal connection. ∎

We could also show this in coordinates, but we will defer this to the end of this section, and continue here with providing another form of the Bianchi identities first. From the fact that one can equivalently express the curvature and torsion forms by the corresponding objects on the base manifold M follows that also the curvature R and torsion T should satisfy a similar set of identities on the base manifold. This is indeed the case, and these identities are given as follows.

Theorem 30.5.2 (Bianchi identities)**.** *The curvature R and torsion T of an affine connection satisfy the Bianchi identities*

$$\mathrm{Cycl}_{X,Y,Z}[R(X,Y)Z] = \mathrm{Cycl}_{X,Y,Z}[T(T(X,Y),Z) + (\nabla_X T)(Y,Z)] \tag{30.5.3}$$

and

$$\mathrm{Cycl}_{X,Y,Z}[(\nabla_X R)(Y,Z) + R(T(X,Y),Z)] = 0\,, \tag{30.5.4}$$

where

$$\mathrm{Cycl}_{X,Y,Z}[F(X,Y,Z)] = F(X,Y,Z) + F(Y,Z,X) + F(Z,X,Y)\,. \tag{30.5.5}$$

Proof. We make use of the fact that we can write the covariant derivative of the torsion tensor as

$$(\nabla_X T)(Y,Z) = \nabla_X(T(Y,Z)) - T(\nabla_X Y,Z) - T(Y,\nabla_X Z)\,. \tag{30.5.6}$$

Further, we can expand the inner torsion tensor using its definition, to obtain

$$T(T(X,Y),Z) = T(\nabla_X Y,Z) - T(\nabla_Y X,Z) - T([X,Y],Z)\,. \tag{30.5.7}$$

Using the fact that we can permute X,Y,Z cyclically if they appear under $\mathrm{Cycl}_{X,Y,Z}$, as well as the fact that the torsion is antisymmetric, we can combine these relations to

$$\mathrm{Cycl}_{X,Y,Z}[T(T(X,Y),Z) + (\nabla_X T)(Y,Z)] = \mathrm{Cycl}_{X,Y,Z}[\nabla_X(T(Y,Z)) - T([X,Y],Z)]\,. \tag{30.5.8}$$

Once again expanding the torsion tensor we have

$$\nabla_X(T(Y,Z)) - T([X,Y],Z) = \nabla_X\nabla_Y Z - \nabla_X\nabla_Z Y - \nabla_X[Y,Z] - \nabla_{[X,Y]}Z + \nabla_Z[X,Y] - [[X,Y],Z]\,. \tag{30.5.9}$$

Under the cyclic sum, the two derivatives acting on Lie brackets cancel, and the last term vanishes due to the Jacobi identity, so that we are left with

$$\begin{aligned}\mathrm{Cycl}_{X,Y,Z}[T(T(X,Y),Z) + (\nabla_X T)(Y,Z)] &= \mathrm{Cycl}_{X,Y,Z}[\nabla_X\nabla_Y Z - \nabla_X\nabla_Z Y - \nabla_{[X,Y]}Z]\\ &= \mathrm{Cycl}_{X,Y,Z}[\nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z]\\ &= \mathrm{Cycl}_{X,Y,Z}[R(X,Y)Z]\,.\end{aligned} \tag{30.5.10}$$

For the second Bianchi identity, we proceed similarly. The derivative of the curvature tensor, acting on a vector field U, is given by

$$(\nabla_X R)(Y,Z)U = \nabla_X(R(Y,Z)U) - R(\nabla_X Y,Z)U - R(Y,\nabla_X Z)U - R(Y,Z)\nabla_X U\,. \tag{30.5.11}$$

Expanding the torsion tensor yields

$$R(T(X,Y),Z) = R(\nabla_X Y, Z) - R(\nabla_Y X, Z) - R([X,Y],Z)\,. \tag{30.5.12}$$

Under the cyclic sum, these combine to

$$\begin{aligned}&\mathrm{Cycl}_{X,Y,Z}[(\nabla_X R)(Y,Z)U + R(T(X,Y),Z)U]\\ &\qquad = \mathrm{Cycl}_{X,Y,Z}[\nabla_X(R(Y,Z)U) - R(Y,Z)\nabla_X U - R([X,Y],Z)U]\,.\end{aligned} \tag{30.5.13}$$

Now we expand the curvature, which yields

$$\begin{aligned}&\nabla_X(R(Y,Z)U) - R(Y,Z)\nabla_X U - R([X,Y],Z)U = \nabla_X\nabla_Y\nabla_Z U - \nabla_X\nabla_Z\nabla_Y U - \nabla_X\nabla_{[Y,Z]}U\\ &\quad - \nabla_Y\nabla_Z\nabla_X U + \nabla_Z\nabla_Y\nabla_X U + \nabla_{[Y,Z]}\nabla_X U - \nabla_{[X,Y]}\nabla_Z U + \nabla_Z\nabla_{[X,Y]}U + \nabla_{[[X,Y],Z]}U\,.\end{aligned} \tag{30.5.14}$$

Under the cyclic sum, all terms cancel, and we find

$$\mathrm{Cycl}_{X,Y,Z}[(\nabla_X R)(Y,Z) + R(T(X,Y),Z)] = 0\,. \tag{30.5.15}$$

■

We also illustrate this calculation using coordinates. First, we write the appearing terms in coordinates, which yields

$$\mathrm{Cycl}_{X,Y,Z}[R(X,Y)Z] = 3R^{\mu}{}_{\nu\rho\sigma}X^{[\rho}Y^{\sigma}Z^{\nu]}\,, \tag{30.5.16a}$$
$$\mathrm{Cycl}_{X,Y,Z}[T(T(X,Y),Z)] = 3T^{\mu}{}_{\omega\nu}T^{\omega}{}_{\rho\sigma}X^{[\rho}Y^{\sigma}Z^{\nu]}\,, \tag{30.5.16b}$$
$$\mathrm{Cycl}_{X,Y,Z}[(\nabla_X T)(Y,Z)] = 3\nabla_{\nu}T^{\mu}{}_{\rho\sigma}X^{[\nu}Y^{\rho}Z^{\sigma]}\,, \tag{30.5.16c}$$
$$\mathrm{Cycl}_{X,Y,Z}[(\nabla_X R)(Y,Z)] = 3\nabla_{\omega}R^{\mu}{}_{\nu\rho\sigma}X^{[\omega}Y^{\rho}Z^{\sigma]}\,, \tag{30.5.16d}$$
$$\mathrm{Cycl}_{X,Y,Z}[R(T(X,Y),Z)] = 3R^{\mu}{}_{\nu\tau\omega}T^{\tau}{}_{\rho\sigma}X^{[\rho}Y^{\sigma}Z^{\omega]}\,, \tag{30.5.16e}$$

where we used the fact that each in each term two of the vector fields X,Y,Z are contracted with the antisymmetric indices of a tensor, so that we can replace the cyclic permutation by a complete antisymmetrization, together with an appropriate normalization factor. Using these expressions, we can write the Bianchi identities as

$$R^{\mu}{}_{[\nu\rho\sigma]} = \nabla_{[\nu}T^{\mu}{}_{\rho\sigma]} + T^{\mu}{}_{\omega[\nu}T^{\omega}{}_{\rho\sigma]} \tag{30.5.17}$$

and

$$\nabla_{[\omega}R^{\mu}{}_{|\nu|\rho\sigma]} + R^{\mu}{}_{\nu\tau[\omega}T^{\tau}{}_{\rho\sigma]} = 0\,. \tag{30.5.18}$$

This can easily be seen by direct calculation. For the first identity (30.5.17), one has

$$\nabla_{[\nu}T^{\mu}{}_{\rho\sigma]} = -2\partial_{[\nu}\Gamma^{\mu}{}_{\rho\sigma]} - 2\Gamma^{\mu}{}_{\omega[\nu}\Gamma^{\omega}{}_{\rho\sigma]} + 2\Gamma^{\omega}{}_{[\rho\nu}\Gamma^{\mu}{}_{|\omega|\sigma]} + 2\Gamma^{\omega}{}_{[\sigma\nu}\Gamma^{\mu}{}_{\rho]\omega}\,, \tag{30.5.19}$$

and

$$T^{\mu}{}_{\omega[\nu}T^{\omega}{}_{\rho\sigma]} = 2\Gamma^{\mu}{}_{\omega[\nu}\Gamma^{\omega}{}_{\rho\sigma]} - 2\Gamma^{\mu}{}_{[\nu|\omega|}\Gamma^{\omega}{}_{\rho\sigma]}\,. \tag{30.5.20}$$

Combining these two terms, and sorting the indices, one obtains the expression

$$\nabla_{[\omega}R^{\mu}{}_{|\nu|\rho\sigma]} + T^{\mu}{}_{\omega[\nu}T^{\omega}{}_{\rho\sigma]} = 2\partial_{[\rho}\Gamma^{\mu}{}_{\nu\sigma]} + 2\Gamma^{\mu}{}_{\omega[\rho}\Gamma^{\omega}{}_{\nu\sigma]} = R^{\mu}{}_{[\nu\rho\sigma]}\,, \tag{30.5.21}$$

which is the left hand side of the Bianchi identity (30.5.17). Similarly, the second identity (30.5.18) can be proven. Expanding the covariant derivative in the first term yields

$$\nabla_{[\omega}R^{\mu}{}_{|\nu|\rho\sigma]} = \partial_{[\omega}R^{\mu}{}_{|\nu|\rho\sigma]} + \Gamma^{\mu}{}_{\tau[\omega}R^{\tau}{}_{|\nu|\rho\sigma]} - \Gamma^{\tau}{}_{\nu[\omega}R^{\mu}{}_{|\tau|\rho\sigma]} - \Gamma^{\tau}{}_{[\rho\omega}R^{\mu}{}_{|\nu\tau|\sigma]} - \Gamma^{\tau}{}_{[\sigma\omega}R^{\mu}{}_{|\nu|\rho]\tau}\,. \tag{30.5.22}$$

Now it is easy to see that the last two terms are identical,

$$\Gamma^{\tau}{}_{[\sigma\omega}R^{\mu}{}_{|\nu|\rho]\tau} = -\Gamma^{\tau}{}_{[\sigma\omega}R^{\mu}{}_{|\nu\tau|\rho]} = \Gamma^{\tau}{}_{[\rho\omega}R^{\mu}{}_{|\nu\tau|\sigma]}\,. \tag{30.5.23}$$

Further using the expression (30.4.7) for the torsion, as well as cyclically rearranging the indices within square brackets, the last two terms of the expansion (30.5.22) can be combined to

$$-2\Gamma^{\tau}{}_{[\rho\omega}R^{\mu}{}_{|\nu\tau|\sigma]} = -R^{\mu}{}_{\nu\tau[\omega}T^{\mu}{}_{\rho\sigma]}\,, \tag{30.5.24}$$

which cancels the second term of the Bianchi identity (30.5.18). It remains to show that the first three terms on the right hand side of the expansion (30.5.22) vanish. Expanding the curvature in the derivative term using the expression (30.3.3) yields

$$\partial_{[\omega}R^{\mu}{}_{|\nu|\rho\sigma]} = 2\partial_{[\omega}\partial_{\rho}\Gamma^{\mu}{}_{|\nu|\sigma]} + 2\partial_{[\omega}\Gamma^{\mu}{}_{|\tau|\rho}\Gamma^{\tau}{}_{|\nu|\sigma]} + 2\Gamma^{\mu}{}_{\tau[\rho}\partial_{\omega}\Gamma^{\tau}{}_{|\nu|\sigma]}\,. \tag{30.5.25}$$

The first term on the right hand side vanishes, since partial derivatives commute. In the remaining two terms one can substitute the derivatives on the connection coefficients by curvature tensors, which yields

$$2\partial_{[\omega}\Gamma^{\mu}{}_{|\tau|\rho}\Gamma^{\tau}{}_{|\nu|\sigma]} = \Gamma^{\tau}{}_{\nu[\sigma}R^{\mu}{}_{|\tau|\omega\rho]} - 2\Gamma^{\tau}{}_{\nu[\sigma}\Gamma^{\mu}{}_{|\phi|\omega}\Gamma^{\phi}{}_{|\tau|\rho]}\,, \tag{30.5.26a}$$

$$2\Gamma^{\mu}{}_{\tau[\rho}\partial_{\omega}\Gamma^{\tau}{}_{|\nu|\sigma]} = \Gamma^{\mu}{}_{\tau[\rho}R^{\tau}{}_{|\nu|\omega\sigma]} - 2\Gamma^{\mu}{}_{\tau[\rho}\Gamma^{\tau}{}_{|\phi|\omega}\Gamma^{\phi}{}_{|\nu|\sigma]}\,. \tag{30.5.26b}$$

After permuting the indices in square brackets, one finds that the two curvature terms found here cancel the second and third term on the right hand side of the expansion (30.5.22). Finally, for the two terms cubic in the connection coefficients, one can exchange the dummy indices $\tau \leftrightarrow \phi$ in the first term to find

$$\Gamma^{\mu}{}_{\tau[\omega}\Gamma^{\tau}{}_{|\phi|\rho}\Gamma^{\phi}{}_{|\nu|\sigma]} + \Gamma^{\mu}{}_{\tau[\rho}\Gamma^{\tau}{}_{|\phi|\omega}\Gamma^{\phi}{}_{|\nu|\sigma]} = 0\,, \tag{30.5.27}$$

due to the antisymmetry in the indices ρ and ω. This completes the proof of the second Bianchi identity (30.5.18).

Note that there is a fundamental difference between the expressions for the Bianchi identities in theorems 30.5.1 and 30.5.2: in contrast to the former, the latter contains additional terms where the torsion and curvature tensors are multiplied with another torsion tensor. These additional terms are related to the fact that in theorem 30.5.2 R and T are tensors in the tangent bundle, and since ∇ is a connection in the tangent bundle, it acts on *all* tensor indices, including those which turn R and T into vector-valued two-forms. In contrast, in theorem 30.5.1 the connection acts *only* on the representation space indices, and the two-form indices are inert.

►...◄

We can also show this in coordinates. First calculating

$$\begin{aligned}
\mathrm{d}_{\vartheta}\Theta &= \mathrm{d}\Theta + \vartheta \wedge \Theta \\
&= -\frac{1}{2}p^{-1\,i}{}_{\mu}p^{-1\,j}{}_{\sigma}T^{\sigma}{}_{\nu\rho}\mathrm{d}p^{\mu}{}_{j} \wedge \mathrm{d}x^{\nu} \wedge \mathrm{d}x^{\rho} \otimes \mathbb{e}_i + \frac{1}{2}p^{-1\,i}{}_{\sigma}\partial_{\mu}T^{\sigma}{}_{\nu\rho}\mathrm{d}x^{\mu} \wedge \mathrm{d}x^{\nu} \wedge \mathrm{d}x^{\rho} \otimes \mathbb{e}_i \\
&\quad + \frac{1}{2}p^{-1\,i}{}_{\mu}(p^{\nu}{}_{j}\Gamma^{\mu}{}_{\nu\rho}\mathrm{d}x^{\rho} + \mathrm{d}p^{\mu}{}_{j}) \wedge (p^{-1\,j}{}_{\sigma}T^{\sigma}{}_{\tau\omega}\mathrm{d}x^{\tau} \wedge \mathrm{d}x^{\omega}) \otimes \mathbb{e}_i \\
&= \frac{1}{2}p^{-1\,i}{}_{\sigma}(\partial_{[\mu}T^{\sigma}{}_{\nu\rho]} + \Gamma^{\sigma}{}_{\omega[\mu}T^{\omega}{}_{\nu\rho]})\mathrm{d}x^{\mu} \wedge \mathrm{d}x^{\nu} \wedge \mathrm{d}x^{\rho} \otimes \mathbb{e}_i \\
&= \frac{1}{2}p^{-1\,i}{}_{\sigma}(\nabla_{[\mu}T^{\sigma}{}_{\nu\rho]} + \Gamma^{\omega}{}_{[\nu\mu}T^{\sigma}{}_{|\omega|\rho]} + \Gamma^{\omega}{}_{[\rho\mu}T^{\sigma}{}_{\nu]\omega})\mathrm{d}x^{\mu} \wedge \mathrm{d}x^{\nu} \wedge \mathrm{d}x^{\rho} \otimes \mathbb{e}_i \\
&= \frac{1}{2}p^{-1\,i}{}_{\sigma}(\nabla_{[\mu}T^{\sigma}{}_{\nu\rho]} + T^{\sigma}{}_{\omega[\rho}T^{\omega}{}_{\mu\nu]})\mathrm{d}x^{\mu} \wedge \mathrm{d}x^{\nu} \wedge \mathrm{d}x^{\rho} \otimes \mathbb{e}_i\,,
\end{aligned} \tag{30.5.28}$$

using the fact that the exterior product is totally antisymmetric, we find the right hand side of the first Bianchi identity (30.5.17). Similarly, we have

$$\Omega \wedge \theta = \frac{1}{2}p^{-1\,i}{}_{\sigma}R^{\sigma}{}_{[\rho\mu\nu]}\mathrm{d}x^{\mu} \wedge \mathrm{d}x^{\nu} \wedge \mathrm{d}x^{\rho} \otimes \mathbb{e}_i\,, \tag{30.5.29}$$

which is the right hand side of the first Bianchi identity (30.5.17). For the second Bianchi identity, we calculate

$$
\begin{aligned}
\mathrm{d}_\vartheta\Omega &= \mathrm{d}\Omega + \mathrm{ad}(\vartheta)\wedge\Omega \\
&= \mathrm{d}\Omega + [\vartheta\wedge\Omega] \\
&= \frac{1}{2}{p^{-1}}^i{}_\mu R^\mu{}_{\nu\rho\sigma}\mathrm{d}p^\nu{}_j\wedge\mathrm{d}x^\rho\wedge\mathrm{d}x^\sigma\otimes\mathcal{H}_i{}^j - \frac{1}{2}{p^{-1}}^i{}_\mu {p^{-1}}^k{}_\sigma p^\nu{}_j R^\sigma{}_{\nu\rho\sigma}\mathrm{d}p^\mu{}_k\wedge\mathrm{d}x^\rho\wedge\mathrm{d}x^\sigma\otimes\mathcal{H}_i{}^j \\
&\quad + \frac{1}{2}{p^{-1}}^i{}_\mu p^\nu{}_j\partial_\omega R^\mu{}_{\nu\rho\sigma}\mathrm{d}x^\omega\wedge\mathrm{d}x^\rho\wedge\mathrm{d}x^\sigma\otimes\mathcal{H}_i{}^j \\
&\quad + \frac{1}{2}{p^{-1}}^i{}_\omega(p^\tau{}_j\Gamma^\omega{}_{\tau\lambda}\mathrm{d}x^\lambda + \mathrm{d}p^\omega{}_j)\wedge({p^{-1}}^k{}_\mu p^\nu{}_l R^\mu{}_{\nu\rho\sigma}\mathrm{d}x^\rho\wedge\mathrm{d}x^\sigma)\otimes[\mathcal{H}_i{}^j,\mathcal{H}_k{}^l] \\
&= \frac{1}{2}{p^{-1}}^i{}_\mu R^\mu{}_{\nu\rho\sigma}\mathrm{d}p^\nu{}_j\wedge\mathrm{d}x^\rho\wedge\mathrm{d}x^\sigma\otimes\mathcal{H}_i{}^j - \frac{1}{2}{p^{-1}}^i{}_\mu {p^{-1}}^k{}_\sigma p^\nu{}_j R^\sigma{}_{\nu\rho\sigma}\mathrm{d}p^\mu{}_k\wedge\mathrm{d}x^\rho\wedge\mathrm{d}x^\sigma\otimes\mathcal{H}_i{}^j \\
&\quad + \frac{1}{2}{p^{-1}}^i{}_\mu p^\nu{}_j\partial_\omega R^\mu{}_{\nu\rho\sigma}\mathrm{d}x^\omega\wedge\mathrm{d}x^\rho\wedge\mathrm{d}x^\sigma\otimes\mathcal{H}_i{}^j \\
&\quad + \frac{1}{2}{p^{-1}}^i{}_\omega(p^\tau{}_k\Gamma^\omega{}_{\tau\lambda}\mathrm{d}x^\lambda + \mathrm{d}p^\omega{}_k)\wedge({p^{-1}}^k{}_\mu p^\nu{}_j R^\mu{}_{\nu\rho\sigma}\mathrm{d}x^\rho\wedge\mathrm{d}x^\sigma)\otimes\mathcal{H}_i{}^j \\
&\quad - \frac{1}{2}(p^\tau{}_j\Gamma^\nu{}_{\tau\lambda}\mathrm{d}x^\lambda + \mathrm{d}p^\nu{}_j)\wedge({p^{-1}}^i{}_\mu R^\mu{}_{\nu\rho\sigma}\mathrm{d}x^\rho\wedge\mathrm{d}x^\sigma)\otimes\mathcal{H}_i{}^j \\
&= \frac{1}{2}{p^{-1}}^i{}_\mu p^\nu{}_j(\partial_{[\omega}R^\mu{}_{|\nu|\rho\sigma]} + \Gamma^\mu{}_{\tau[\omega}R^\tau{}_{|\nu|\rho\sigma]} - \Gamma^\tau{}_{\nu[\omega}R^\mu{}_{|\tau|\rho\sigma]})\mathrm{d}x^\omega\wedge\mathrm{d}x^\rho\wedge\mathrm{d}x^\sigma\otimes\mathcal{H}_i{}^j \\
&= \frac{1}{2}{p^{-1}}^i{}_\mu p^\nu{}_j(\nabla_{[\omega}R^\mu{}_{|\nu|\rho\sigma]} + \Gamma^\tau{}_{[\rho\omega}R^\mu{}_{|\nu\tau|\sigma]} + \Gamma^\tau{}_{[\sigma\omega}R^\mu{}_{|\nu|\rho]\tau})\mathrm{d}x^\omega\wedge\mathrm{d}x^\rho\wedge\mathrm{d}x^\sigma\otimes\mathcal{H}_i{}^j \\
&= \frac{1}{2}{p^{-1}}^i{}_\mu p^\nu{}_j(\nabla_{[\omega}R^\mu{}_{|\nu|\rho\sigma]} + R^\mu{}_{\nu\tau[\omega}T^\tau{}_{\rho\sigma]})\mathrm{d}x^\omega\wedge\mathrm{d}x^\rho\wedge\mathrm{d}x^\sigma\otimes\mathcal{H}_i{}^j\,,
\end{aligned}
\tag{30.5.30}
$$

and so we indeed find the second Bianchi identity (30.5.18).

30.6 Higher order covariant derivatives

The fact that affine connections constitute Koszul connections in any vector bundle associated to the tangent frame bundle, and thus in particular the tangent bundle, its dual and any tensor bundles constructed from these, allows for another construction, which is not possible for arbitrary vector bundles. Recall that given a (r,s)-tensor field $A\in\Gamma(T^r_sM)$ on a manifold M equipped with an affine connection ∇, applying the Koszul connection yields a section $\nabla A\in\Gamma(T^M\otimes T^r_sM)$. However, keeping in mind that $T^M\otimes T^r_sM\cong T^r_{s+1}M$, we see that ∇A is again a tensor field to which we can apply the Koszul connection ∇ in $T^r_{s+1}M$, and thus construct $\nabla\nabla A$. Note that this is not the case for an arbitrary vector bundle $\pi: E\to M$, since a Koszul connection on E yields a section of $T^*M\otimes E$, which is not equipped with a Koszul connection by the Koszul connection on E. Also note that the construction $\nabla\nabla A$ is distinct from other iterated covariant derivatives which exist for arbitrary vector bundles, namely $\nabla_X\nabla_YA$ with two vector fields $X,Y\in\mathrm{Vect}(M)$ and the exterior covariant derivative $\mathrm{d}^\nabla\mathrm{d}^\nabla A$, as we will show below. The reason for this is closely related to the fact that even though we use the same symbol ∇ for the two Koszul connections in the expression $\nabla\nabla A$, these act on different bundles T^r_sM and $T^r_{s+1}M$. We start our discussion with the following definition.

Definition 30.6.1 (Higher order covariant derivative)**.** Let M be a manifold equipped with an affine connection ∇ and $A\in\Gamma(T^r_sM)$ a tensor field. For $n\in\mathbb{N}$, we denote by

$$\nabla^nA = \underbrace{\nabla\cdots\nabla}_{n\text{ times}}A \tag{30.6.1}$$

the n times iterated Koszul connection, giving rise to the *nth order covariant derivative*

$$\nabla^n_{X_n,\ldots,X_1} A = \mathrm{tr}_1^1 \cdots \mathrm{tr}_n^n (X^n \otimes \ldots \otimes X^1 \otimes \nabla^n A) \tag{30.6.2}$$

for n vector fields $X_1, \ldots, X_n \in \mathrm{Vect}(M)$.

Note in particular that there are no derivatives acting on the vector fields $X_1, \ldots X_n$, since they are simply contracted with the tensor field $\nabla^n A$. Note also the order of the vector fields in this notation, which indicates that X^1 is contracted with the slot resulting from the *first* application of the Koszul connection, while X^n is paired with the *last* application of the Koszul connection, i.e., derivatives are applied in the order from right (closed to A) to left (furthest from A). Given coordinates (x^μ) on M, this is also illustrated by the notation

$$\begin{aligned}\nabla^n A &= \nabla_{\rho_n} \cdots \nabla_{\rho_1} A^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_s} \mathrm{d}x^{\rho_n} \otimes \ldots \otimes \mathrm{d}x^{\rho_1} \otimes \partial_{\mu_1} \otimes \ldots \otimes \partial_{\mu_r} \otimes \mathrm{d}x^{\nu_1} \otimes \ldots \otimes \mathrm{d}x^{\nu_s} \\ &= A^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_s;\rho_1\cdots\rho_n} \mathrm{d}x^{\rho_n} \otimes \ldots \otimes \mathrm{d}x^{\rho_1} \otimes \partial_{\mu_1} \otimes \ldots \otimes \partial_{\mu_r} \otimes \mathrm{d}x^{\nu_1} \otimes \ldots \otimes \mathrm{d}x^{\nu_s}\,,\end{aligned} \tag{30.6.3}$$

where covariant derivatives ∇_ρ and basis elements $\mathrm{d}x^\rho$ are prepended on the left, where we follow the same convention as usual for the Koszul connection discussed in section 28.2. Another convention, which is also used in the literature, is to append basis elements $\mathrm{d}x^\rho$ on the right instead, as suggested by the order of indices in the semicolon notation, where derivative indices are appended right of the semicolon. Irrespective of the choice of convention, the higher order covariant derivative with respect to n vector fields is given by

$$\begin{aligned}\nabla^n_{X_n,\ldots,X_1} A &= X_n^{\rho_n} \cdots X_1^{\rho_1} \nabla_{\rho_n} \cdots \nabla_{\rho_1} A^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_s} \partial_{\mu_1} \otimes \ldots \otimes \partial_{\mu_r} \otimes \mathrm{d}x^{\nu_1} \otimes \ldots \otimes \mathrm{d}x^{\nu_s} \\ &= X_n^{\rho_n} \cdots X_1^{\rho_1} A^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_s;\rho_1\cdots\rho_n} \partial_{\mu_1} \otimes \ldots \otimes \partial_{\mu_r} \otimes \mathrm{d}x^{\nu_1} \otimes \ldots \otimes \mathrm{d}x^{\nu_s}\,.\end{aligned} \tag{30.6.4}$$

In the following, we will study the properties of this higher order covariant derivative, and its relation to constructions which we have encountered before. We start with the following helpful relation.

Theorem 30.6.1. *For two vector fields $X, Y \in \mathrm{Vect}(M)$, the second order covariant derivative of a tensor field A satisfies*

$$\nabla^2_{X,Y} A = \nabla_X \nabla_Y A - \nabla_{\nabla_X Y} A\,. \tag{30.6.5}$$

Proof. By definition, we have

$$\begin{aligned}\nabla_X \nabla_Y A &= \nabla_X \mathrm{tr}_1^1 (Y \otimes \nabla A) \\ &= \mathrm{tr}_1^1 (\nabla_X Y \otimes \nabla A + Y \otimes \nabla_X \nabla A) \\ &= \mathrm{tr}_1^1 (\nabla_X Y \otimes \nabla A) + \mathrm{tr}_1^1 (X \otimes Y \otimes \nabla\nabla A) \\ &= \nabla_{\nabla_X Y} A + \nabla^2_{X,Y} A\,. \quad \blacksquare\end{aligned} \tag{30.6.6}$$

We also show this in coordinates, which is instructive to understanding the coordinate notation. The calculation in the proof above then takes the form

$$\begin{aligned}(\nabla_X \nabla_Y A) A^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_s} &= X^\rho \nabla_\rho (Y^\sigma \nabla_\sigma A^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_s}) \\ &= X^\rho \nabla_\rho Y^\sigma \nabla_\sigma A^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_s} + X^\rho Y^\sigma \nabla_\rho \nabla_\sigma A^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_s} \\ &= (\nabla_{\nabla_X Y} A + \nabla^2_{X,Y} A)^{\mu_1\cdots\mu_r}{}_{\nu_1\cdots\nu_s}\,.\end{aligned} \tag{30.6.7}$$

This relation turns out to be helpful in the following calculations, where we discuss in particular the relevance of the order of the appearing vector fields in the higher order covariant derivative. We first introduce the following notion.

Definition 30.6.2 (Covariant Hessian)**.** Let M be a manifold equipped with an affine connection ∇ and $f \in C^\infty(M, \mathbb{R})$ a function. The *covariant Hessian* of f is the $(0,2)$-tensor field $\nabla^2 f$.

Hence, we can write the coordinate expression for the covariant Hessian as

$$\nabla_\mu \nabla_\nu f = \partial_\mu \partial_\nu f - \Gamma^\rho{}_{\nu\mu} \partial_\rho f \,. \tag{30.6.8}$$

In vector analysis, the Hessian is defined with partial derivatives with respect to the Cartesian coordinates, and from the fact that partial derivatives commute follows that the Hessian is symmetric. However, this is in general not the case for covariant derivatives, and so the covariant Hessian turns out to be not symmetric in general, as indicated by the presence of the connection coefficients in the coordinate expression above. We will show this as follows.

Theorem 30.6.2. *For a scalar function $f \in C^\infty(M, \mathbb{R})$, the commutator of second order covariant derivatives satisfies*

$$\nabla^2_{X,Y} f - \nabla^2_{Y,X} f = -\nabla_{T(X,Y)} f \tag{30.6.9}$$

for all vector fields $X, Y \in \mathrm{Vect}(M)$.

Proof. Recall that on a scalar function the covariant derivative acts as $\nabla_X f = Xf$, independently of the connection. We can thus use theorem 30.6.1 to calculate

$$\begin{aligned} \nabla^2_{X,Y} f - \nabla^2_{Y,X} f &= \nabla_X \nabla_Y f - \nabla_{\nabla_X Y} f - \nabla_Y \nabla_X f + \nabla_{\nabla_Y X} f \\ &= X(Yf) - (\nabla_X Y) f - Y(Xf) + (\nabla_Y X) f \\ &= ([X,Y] - \nabla_X Y + \nabla_Y X) f \\ &= -T(X,Y) f \\ &= -\nabla_{T(X,Y)} f \,. \quad \blacksquare \end{aligned} \tag{30.6.10}$$

We also see this relation in coordinates, where we can now omit the vector fields X and Y, which we used above to indicate the order of differentiation, which we can indicate by indices in the coordinate notation, so that we find

$$\begin{aligned} 2\nabla_{[\mu} \nabla_{\nu]} f &= \nabla_\mu \partial_\nu f - \nabla_\nu \partial_\mu f \\ &= \partial_\mu \partial_\nu f - \Gamma^\rho{}_{\nu\mu} \partial_\rho f - \partial_\nu \partial_\mu f + \Gamma^\rho{}_{\mu\nu} \partial_\rho f \\ &= -T^\rho{}_{\mu\nu} \partial_\rho f \\ &= -T^\rho{}_{\mu\nu} \nabla_\rho f \,. \end{aligned} \tag{30.6.11}$$

Hence, we find that the antisymmetric part of the Hessian is closely related to the torsion of the connection. Note that the proof relied on the fact that we can substitute covariant derivatives on a function by the application of the vector field, and so it is valid for scalar functions only. In order to generalize this result to tensor fields of higher rank, we continue with the case of a vector field.

Theorem 30.6.3. *For a vector field $V \in \mathrm{Vect}(M)$, the commutator of second order covariant derivatives satisfies*

$$\nabla^2_{X,Y} V - \nabla^2_{Y,X} V = R(X,Y)V - \nabla_{T(X,Y)} V \tag{30.6.12}$$

for all vector fields $X, Y \in \mathrm{Vect}(M)$.

Proof. Using again theorem 30.6.1, we calculate

$$\begin{aligned} \nabla^2_{X,Y} V - \nabla^2_{Y,X} V &= \nabla_X \nabla_Y V - \nabla_{\nabla_X Y} V - \nabla_Y \nabla_X V + \nabla_{\nabla_Y X} V \\ &= R(X,Y)V + \nabla_{[X,Y]} V - \nabla_{\nabla_X Y} V + \nabla_{\nabla_Y X} V \\ &= R(X,Y)V - \nabla_{T(X,Y)} V \,. \quad \blacksquare \end{aligned} \tag{30.6.13}$$

We also show this in coordinates, where now we have

$$
\begin{aligned}
2\nabla_{[\mu}\nabla_{\nu]}V^\rho &= \nabla_\mu(\partial_\nu V^\rho + \Gamma^\rho{}_{\sigma\nu}V^\sigma) - \nabla_\nu(\partial_\mu V^\rho + \Gamma^\rho{}_{\sigma\mu}V^\sigma)\\
&= \partial_\mu(\partial_\nu V^\rho + \Gamma^\rho{}_{\sigma\nu}V^\sigma) - \Gamma^\omega{}_{\nu\mu}(\partial_\omega V^\rho + \Gamma^\rho{}_{\sigma\omega}V^\sigma) + \Gamma^\rho{}_{\omega\mu}(\partial_\nu V^\omega + \Gamma^\omega{}_{\sigma\nu}V^\sigma)\\
&\quad - \partial_\nu(\partial_\mu V^\rho + \Gamma^\rho{}_{\sigma\mu}V^\sigma) + \Gamma^\omega{}_{\mu\nu}(\partial_\omega V^\rho + \Gamma^\rho{}_{\sigma\omega}V^\sigma) - \Gamma^\rho{}_{\omega\nu}(\partial_\mu V^\omega + \Gamma^\omega{}_{\sigma\mu}V^\sigma)\\
&= \partial_\mu\Gamma^\rho{}_{\sigma\nu}V^\sigma + \Gamma^\rho{}_{\sigma\nu}\partial_\mu V^\sigma - \Gamma^\omega{}_{\nu\mu}(\partial_\omega V^\rho + \Gamma^\rho{}_{\sigma\omega}V^\sigma) + \Gamma^\rho{}_{\omega\mu}(\partial_\nu V^\omega + \Gamma^\omega{}_{\sigma\nu}V^\sigma)\\
&\quad - \partial_\nu\Gamma^\rho{}_{\sigma\mu}V^\sigma - \Gamma^\rho{}_{\sigma\mu}\partial_\nu V^\sigma + \Gamma^\omega{}_{\mu\nu}(\partial_\omega V^\rho + \Gamma^\rho{}_{\sigma\omega}V^\sigma) - \Gamma^\rho{}_{\omega\nu}(\partial_\mu V^\omega + \Gamma^\omega{}_{\sigma\mu}V^\sigma)\\
&= (\partial_\mu\Gamma^\rho{}_{\sigma\nu} - \partial_\nu\Gamma^\rho{}_{\sigma\mu} + \Gamma^\rho{}_{\omega\mu}\Gamma^\omega{}_{\sigma\nu} - \Gamma^\rho{}_{\omega\nu}\Gamma^\omega{}_{\sigma\mu})V^\sigma\\
&\quad + (\Gamma^\omega{}_{\mu\nu} - \Gamma^\omega{}_{\nu\mu})(\partial_\omega V^\rho + \Gamma^\rho{}_{\sigma\omega}V^\sigma)\\
&= R^\rho{}_{\sigma\mu\nu}V^\sigma - T^\omega{}_{\mu\nu}\nabla_\omega V^\rho\,.
\end{aligned}
\tag{30.6.14}
$$

We then continue with the second order covariant derivative of a covector field.

Theorem 30.6.4. *For a covector field $\alpha \in \Omega^1(M)$, the commutator of second order covariant derivatives satisfies*

$$
Z \lrcorner (\nabla^2_{X,Y}\alpha - \nabla^2_{Y,X}\alpha) = -(R(X,Y)Z) \lrcorner \alpha - Z \lrcorner \nabla_{T(X,Y)}\alpha \tag{30.6.15}
$$

for all vector fields $X, Y, Z \in \mathrm{Vect}(M)$.

Proof. It is helpful to use the Leibniz rule

$$
\nabla_X(Z \lrcorner \alpha) = (\nabla_X Z) \lrcorner \alpha + Z \lrcorner \nabla_X\alpha\,, \tag{30.6.16}
$$

as well as its second application

$$
\nabla_X\nabla_Y(Z \lrcorner \alpha) = (\nabla_X\nabla_Y Z) \lrcorner \alpha + (\nabla_Y Z) \lrcorner \nabla_X\alpha + (\nabla_X Z) \lrcorner \nabla_Y\alpha + Z \lrcorner \nabla_X\nabla_Y\alpha\,. \tag{30.6.17}
$$

Then we can calculate

$$
\begin{aligned}
Z \lrcorner (\nabla^2_{X,Y}\alpha - \nabla^2_{Y,X}\alpha) &= Z \lrcorner (\nabla_X\nabla_Y\alpha - \nabla_{\nabla_X Y}\alpha - \nabla_Y\nabla_X\alpha + \nabla_{\nabla_Y X}\alpha)\\
&= \nabla_X\nabla_Y(Z \lrcorner \alpha) - (\nabla_X\nabla_Y Z) \lrcorner \alpha - (\nabla_Y Z) \lrcorner \nabla_X\alpha - (\nabla_X Z) \lrcorner \nabla_Y\alpha\\
&\quad - \nabla_Y\nabla_X(Z \lrcorner \alpha) + (\nabla_Y\nabla_X Z) \lrcorner \alpha + (\nabla_X Z) \lrcorner \nabla_Y\alpha + (\nabla_Y Z) \lrcorner \nabla_X\alpha\\
&\quad - \nabla_{\nabla_X Y}(Z \lrcorner \alpha) + (\nabla_{\nabla_X Y}Z) \lrcorner \alpha + \nabla_{\nabla_Y X}(Z \lrcorner \alpha) - (\nabla_{\nabla_Y X}Z) \lrcorner \alpha\\
&= \nabla^2_{X,Y}(Z \lrcorner \alpha) - \nabla^2_{Y,X}(Z \lrcorner \alpha) - (\nabla^2_{X,Y}Z - \nabla^2_{Y,X}Z) \lrcorner \alpha\\
&= -\nabla_{T(X,Y)}(Z \lrcorner \alpha) - [R(X,Y)Z - \nabla_{T(X,Y)}Z] \lrcorner \alpha\\
&= -(R(X,Y)Z) \lrcorner \alpha - Z \lrcorner \nabla_{T(X,Y)}\alpha\,. \quad \blacksquare
\end{aligned}
\tag{30.6.18}
$$

In the proof given above, the vector field Z plays an auxiliary role, as it allows us to make use of the theorems 30.6.2 and 30.6.3 which we have proven earlier. If we calculate in coordinates, we can omit this vector field (as we have also omitted X and Y before) and replace it by a free index. Then we obtain

$$
\begin{aligned}
2\nabla_{[\mu}\nabla_{\nu]}\alpha_\rho &= \nabla_\mu(\partial_\nu\alpha_\rho - \Gamma^\sigma{}_{\rho\nu}\alpha_\sigma) - \nabla_\nu(\partial_\mu\alpha_\rho - \Gamma^\sigma{}_{\rho\mu}\alpha_\sigma)\\
&= \partial_\mu(\partial_\nu\alpha_\rho - \Gamma^\sigma{}_{\rho\nu}\alpha_\sigma) - \Gamma^\omega{}_{\nu\mu}(\partial_\omega\alpha_\rho - \Gamma^\sigma{}_{\rho\omega}\alpha_\sigma) - \Gamma^\omega{}_{\rho\mu}(\partial_\nu\alpha_\omega - \Gamma^\sigma{}_{\omega\nu}\alpha_\sigma)\\
&\quad - \partial_\nu(\partial_\mu\alpha_\rho - \Gamma^\sigma{}_{\rho\mu}\alpha_\sigma) + \Gamma^\omega{}_{\mu\nu}(\partial_\omega\alpha_\rho - \Gamma^\sigma{}_{\rho\omega}\alpha_\sigma) + \Gamma^\omega{}_{\rho\nu}(\partial_\mu\alpha_\omega - \Gamma^\sigma{}_{\omega\mu}\alpha_\sigma)\\
&= -\partial_\mu\Gamma^\sigma{}_{\rho\nu}\alpha_\sigma + \Gamma^\sigma{}_{\rho\nu}\partial_\mu\alpha_\sigma - \Gamma^\omega{}_{\nu\mu}(\partial_\omega\alpha_\rho - \Gamma^\sigma{}_{\rho\omega}\alpha_\sigma) - \Gamma^\omega{}_{\rho\mu}(\partial_\nu\alpha_\omega - \Gamma^\sigma{}_{\omega\nu}\alpha_\sigma)\\
&\quad + \partial_\nu\Gamma^\sigma{}_{\rho\mu}\alpha_\sigma - \Gamma^\sigma{}_{\rho\mu}\partial_\nu\alpha_\sigma + \Gamma^\omega{}_{\mu\nu}(\partial_\omega\alpha_\rho - \Gamma^\sigma{}_{\rho\omega}\alpha_\sigma) + \Gamma^\omega{}_{\rho\nu}(\partial_\mu\alpha_\omega - \Gamma^\sigma{}_{\omega\mu}\alpha_\sigma)\\
&= -(\partial_\mu\Gamma^\sigma{}_{\rho\nu} - \partial_\nu\Gamma^\sigma{}_{\rho\mu} + \Gamma^\sigma{}_{\omega\mu}\Gamma^\omega{}_{\rho\nu} - \Gamma^\sigma{}_{\omega\nu}\Gamma^\omega{}_{\rho\mu})\alpha_\sigma\\
&\quad + (\Gamma^\omega{}_{\mu\nu} - \Gamma^\omega{}_{\nu\mu})(\partial_\omega\alpha_\rho - \Gamma^\sigma{}_{\rho\omega}\alpha_\sigma)\\
&= -R^\sigma{}_{\rho\mu\nu}\alpha_\sigma - T^\omega{}_{\mu\nu}\nabla_\omega\alpha_\rho\,.
\end{aligned}
\tag{30.6.19}
$$

By comparing the results for the scalar, vector and covector case, we can already deduce a pattern, which we may expect to hold also for tensors of higher rank. To see this, it is helpful to recall that the curvature $R(X,Y)$ contracted with two vector fields is a tensor field of rank $(1,1)$. We can thus write the statement of theorem 30.6.3 as

$$\nabla^2_{X,Y}V - \nabla^2_{Y,X}V = \operatorname{tr}^1_1(V \otimes R(X,Y)) - \nabla_{T(X,Y)}V\,, \tag{30.6.20}$$

while the statement of theorem 30.6.4 becomes

$$\nabla^2_{X,Y}\alpha - \nabla^2_{Y,X}\alpha = -\operatorname{tr}^1_1(\alpha \otimes R(X,Y)) - \nabla_{T(X,Y)}\alpha\,. \tag{30.6.21}$$

▶...◀

Theorem 30.6.5. *For a tensor field $A \in \Gamma(T^r_sM)$ of rank (r,s), the commutator of second order covariant derivatives satisfies*

$$\blacktriangleright \ldots \blacktriangleleft \tag{30.6.22}$$

Proof. Since a Koszul connection is linear by definition, and any tensor field of rank (r,s) can be expressed as a sum of tensor products of r vector fields and s covector fields, it is sufficient to prove the statement for such tensor products. First, note that for any tensor fields A, B the Leibniz rule yields

$$\nabla^2_{X,Y}(A \otimes B) - \nabla^2_{Y,X}(A \otimes B) = (\nabla^2_{X,Y}A - \nabla^2_{Y,X}A) \otimes B + A \otimes (\nabla^2_{X,Y}B - \nabla^2_{Y,X}B)\,, \tag{30.6.23}$$

which can be shown in analogy to the proof of theorem 30.6.4. Let $V_1, \ldots, V_r \in \operatorname{Vect}(M)$ be vector fields and $\alpha_1, \ldots, \alpha_s \in \Omega^1(M)$ be covector fields, and set

$$A = V_1 \otimes \ldots \otimes V_r \otimes \alpha_1 \otimes \ldots \otimes \alpha_s\,. \tag{30.6.24}$$

By repeated application of the Leibniz rule, as well as theorems 30.6.3 and 30.6.4, we have

$$\begin{aligned}
\nabla^2_{X,Y}A - \nabla^2_{Y,X}A &= (\nabla^2_{X,Y}V_1 - \nabla^2_{Y,X}V_1) \otimes \ldots \otimes V_r \otimes \alpha_1 \otimes \ldots \otimes \alpha_s\\
&\quad + \ldots\\
&\quad + V_1 \otimes \ldots \otimes (\nabla^2_{X,Y}V_r - \nabla^2_{Y,X}V_r) \otimes \alpha_1 \otimes \ldots \otimes \alpha_s\\
&\quad + V_1 \otimes \ldots \otimes V_r \otimes (\nabla^2_{X,Y}\alpha_1 - \nabla^2_{Y,X}\alpha_1) \otimes \ldots \otimes \alpha_s\\
&\quad + \ldots\\
&\quad + V_1 \otimes \ldots \otimes V_r \otimes \alpha_1 \otimes \ldots \otimes (\nabla^2_{X,Y}\alpha_s - \nabla^2_{Y,X}\alpha_s)\\
&= [\operatorname{tr}^1_1(V_1 \otimes R(X,Y)) - \nabla_{T(X,Y)}V_1] \otimes \ldots \otimes V_r \otimes \alpha_1 \otimes \ldots \otimes \alpha_s\\
&\quad + \ldots\\
&\quad + V_1 \otimes \ldots \otimes [\operatorname{tr}^1_1(V_r \otimes R(X,Y)) - \nabla_{T(X,Y)}V_r] \otimes \alpha_1 \otimes \ldots \otimes \alpha_s\\
&\quad - V_1 \otimes \ldots \otimes V_r \otimes [\operatorname{tr}^1_1(\alpha_1 \otimes R(X,Y)) + \nabla_{T(X,Y)}\alpha_1] \otimes \ldots \otimes \alpha_s\\
&\quad - \ldots\\
&\quad - V_1 \otimes \ldots \otimes V_r \otimes \alpha_1 \otimes \ldots \otimes [\operatorname{tr}^1_1(\alpha_s \otimes R(X,Y)) + \nabla_{T(X,Y)}\alpha_s]\\
&= \operatorname{tr}^1_1(V_1 \otimes R(X,Y)) \otimes \ldots \otimes V_r \otimes \alpha_1 \otimes \ldots \otimes \alpha_s\\
&\quad + \ldots\\
&\quad + V_1 \otimes \ldots \otimes \operatorname{tr}^1_1(V_r \otimes R(X,Y)) \otimes \alpha_1 \otimes \ldots \otimes \alpha_s\\
&\quad - V_1 \otimes \ldots \otimes V_r \otimes \operatorname{tr}^1_1(\alpha_1 \otimes R(X,Y)) \otimes \ldots \otimes \alpha_s\\
&\quad - \ldots\\
&\quad - V_1 \otimes \ldots \otimes V_r \otimes \alpha_1 \otimes \ldots \otimes \operatorname{tr}^1_1(\alpha_s \otimes R(X,Y))\\
&\quad - \nabla_{T(X,Y)}(V_1 \otimes \ldots \otimes V_r \otimes \alpha_1 \otimes \ldots \otimes \alpha_s)\,.
\end{aligned} \tag{30.6.25}$$

By linearity of the affine connection, it follows that this holds also if A is any tensor fields of rank (r,s). ■

While the statement above is rather cumbersome to write (but nevertheless straightforward to calculate) without using coordinates, it can be written compactly in coordinates as

$$\begin{aligned}2\nabla_{[\mu}\nabla_{\nu]}A^{\rho_1\cdots\rho_r}{}_{\sigma_1\cdots\sigma_s} &= R^{\rho_1}{}_{\omega\mu\nu}A^{\omega\rho_2\cdots\rho_r}{}_{\sigma_1\cdots\sigma_s} + \ldots + R^{\rho_r}{}_{\omega\mu\nu}A^{\rho_1\cdots\rho_{r-1}\omega}{}_{\sigma_1\cdots\sigma_s} \\ &\quad - R^{\omega}{}_{\sigma_1\mu\nu}A^{\rho_1\cdots\rho_r}{}_{\omega\sigma_2\cdots\sigma_s} - \ldots - R^{\omega}{}_{\sigma_s\mu\nu}A^{\rho_1\cdots\rho_r}{}_{\sigma_1\cdots\sigma_{s-1}\omega} \\ &\quad - T^{\omega}{}_{\mu\nu}\nabla_{\omega}A^{\rho_1\cdots\rho_r}{}_{\sigma_1\cdots\sigma_s}\,,\end{aligned} \tag{30.6.26}$$

where we omit the calculation, as it becomes significantly more lengthy.

30.7 Autoparallel curves

Recall from section 26.7 that a connection on a fiber bundle $\pi : E \to M$ defines a notion of horizontal curves in the total space E of the bundle, and in particular the notion of the horizontal lift of a curve from the base manifold M to the total space E. In the case of an affine connection, this bundle is the tangent bundle $\tau : TM \to M$. From section 7.3 we know another way of lifting curves from the base manifold into the tangent bundle, known as the canonical lift. One may thus wonder whether there is any relation between this notions. While this is not the case in general, there exists a class of curves which is distinguished by the fact that these notions agree, and which we denote as follows.

Definition 30.7.1 (Autoparallel curve)**.** Let M be a manifold equipped with a connection ∇. An *autoparallel curve* is a curve $\gamma \in C^\infty(\mathbb{R}, M)$ such that its canonical lift $\dot{\gamma} \in C^\infty(\mathbb{R}, TM)$ is horizontal.

The name originates from the fact that for an autoparallel, the tangent vector at any point of the curve is obtained via parallel transport from any other point, following section 26.8. Using the coordinate expression for the connection in terms of connection coefficients, one can derive a coordinate expression for autoparallel curves as follows. ►. . . ◄

Theorem 30.7.1. *For every tangent vector $v \in TM$, an affine connection defines a unique autoparallel curve $\gamma : \mathbb{R} \to M$ such that $\dot{\gamma}(0) = v$.*

Proof. ►. . . ◄ ■

30.8 Affine bundle of connections

30.9 Pullback and Lie derivative

We have seen in the previous section 30.8 that an affine connection on a manifold M can be seen as a section of an affine bundle modeled over the tensor bundle T^1_2M. Since the latter is a natural bundle, which means that there exists a natural lift of diffeomorphisms from the base manifold M into the total space of the bundle, one may expect that the same holds also for the bundle of affine connections. We will now show that this is the case, and discuss a few of its properties, using the different approaches to affine connections encountered so far. We start with the following definition, which makes use of the representation of an affine connection as a principal connection on the tangent frame bundle.

Definition 30.9.1 (Pullback of a connection in the tangent frame bundle)**.** Let N be a manifold of dimension $\dim N = n$ equipped with a principal $\mathfrak{gl}(n,\mathbb{R})$-connection $\vartheta \in \Omega^1(FN, \mathfrak{gl}(n,\mathbb{R}))$ on its frame bundle FN and $\varphi : M \to N$ a diffeomorphism. The *pullback* of ϑ along φ is the principal $\mathfrak{gl}(n,\mathbb{R})$-connection $\varphi_\circ^*\vartheta \in \Omega^1(FM, \mathfrak{gl}(n,\mathbb{R}))$ on FM.

It is instructive to study this also in coordinates. First, recall from section ▶Construct this in tangent frame bundle section…◀ We thus find that the pullback of the connection can be expressed in coordinates as

$$\begin{aligned}
\vartheta'(x,p) &= (\vartheta'^i{}_{j\mu}(x,p)\mathrm{d}x^\mu + \bar{\vartheta}^i{}_{j\mu}{}^k\mathrm{d}p^\mu{}_k) \otimes \mathcal{H}_i{}^j \\
&= \left[\vartheta^i{}_{j\mu}(x',p')\left(\frac{\partial x'^\mu}{\partial x^\nu}\mathrm{d}x^\nu + \frac{\partial x'^\mu}{\partial p^\nu{}_k}\mathrm{d}p^\nu{}_k\right) + p'^{-1\,i}{}_\mu\left(\frac{\partial p'^\mu{}_j}{\partial x^\nu}\mathrm{d}x^\nu + \frac{\partial p'^\mu{}_j}{\partial p^\nu{}_k}\mathrm{d}p^\nu{}_k\right)\right] \otimes \mathcal{H}_i{}^j \\
&= \left[\vartheta^i{}_{j\mu}(x',p')\frac{\partial x'^\mu}{\partial x^\nu}\mathrm{d}x^\nu + p^{-1\,i}{}_\rho\frac{\partial x^\rho}{\partial x'^\mu}\left(p^\sigma{}_j\frac{\partial x'^\mu}{\partial x^\sigma\partial x^\nu}\mathrm{d}x^\nu + \frac{\partial x'^\mu}{\partial x^\nu}\delta^k_j\mathrm{d}p^\nu{}_k\right)\right] \otimes \mathcal{H}_i{}^j \\
&= \left[\left(\vartheta^i{}_{j\nu}(x',p')\frac{\partial x'^\nu}{\partial x^\mu} + p^{-1\,i}{}_\rho p^\sigma{}_j\frac{\partial x^\rho}{\partial x'^\nu}\frac{\partial x'^\nu}{\partial x^\sigma\partial x^\mu}\right)\mathrm{d}x^\mu + p^{-1\,i}{}_\mu\mathrm{d}p^\mu{}_j\right] \otimes \mathcal{H}_i{}^j\,,
\end{aligned} \tag{30.9.1}$$

and so we can read off the coefficients given by

$$\vartheta'^i{}_{j\mu}(x,p) = \vartheta^i{}_{j\nu}(x',p')\frac{\partial x'^\nu}{\partial x^\mu} + p^{-1\,i}{}_\rho p^\sigma{}_j\frac{\partial x^\rho}{\partial x'^\nu}\frac{\partial x'^\nu}{\partial x^\sigma\partial x^\mu} = p^{-1\,i}{}_\nu p^\rho{}_j\Gamma'^\nu{}_{\rho\mu}(x)\,, \tag{30.9.2}$$

where the last equality is obtained by expressing the pullback of the connection in terms of the transformed connection coefficients $\Gamma'^\mu{}_{\nu\rho}$. Solving for the latter yields

$$\begin{aligned}
\Gamma'^\mu{}_{\nu\rho}(x) &= p^\mu{}_i p^{-1\,j}{}_\nu\vartheta'^i{}_{j\mu}(x,p) \\
&= p^\mu{}_i p^{-1\,j}{}_\nu\left[\vartheta^i{}_{j\sigma}(x',p')\frac{\partial x'^\sigma}{\partial x^\rho} + p^{-1\,i}{}_\omega p^\tau{}_j\frac{\partial x^\omega}{\partial x'^\sigma}\frac{\partial x'^\sigma}{\partial x^\tau\partial x^\rho}\right] \\
&= p^\mu{}_i p^{-1\,j}{}_\nu\left[p'^{-1\,i}{}_\lambda p'^\psi{}_j\Gamma^\lambda{}_{\psi\sigma}(x')\frac{\partial x'^\sigma}{\partial x^\rho} + p^{-1\,i}{}_\omega p^\tau{}_j\frac{\partial x^\omega}{\partial x'^\sigma}\frac{\partial x'^\sigma}{\partial x^\tau\partial x^\rho}\right] \\
&= p^\mu{}_i p^{-1\,j}{}_\nu\left[p^{-1\,i}{}_\omega p^\tau{}_j\frac{\partial x^\omega}{\partial x'^\lambda}\frac{\partial x'^\psi}{\partial x^\tau}\Gamma^\lambda{}_{\psi\sigma}(x')\frac{\partial x'^\sigma}{\partial x^\rho} + p^{-1\,i}{}_\omega p^\tau{}_j\frac{\partial x^\omega}{\partial x'^\sigma}\frac{\partial x'^\sigma}{\partial x^\tau\partial x^\rho}\right] \\
&= \frac{\partial x^\mu}{\partial x'^\lambda}\frac{\partial x'^\psi}{\partial x^\nu}\frac{\partial x'^\sigma}{\partial x^\rho}\Gamma^\lambda{}_{\psi\sigma}(x') + \frac{\partial x^\mu}{\partial x'^\sigma}\frac{\partial x'^\sigma}{\partial x^\nu\partial x^\rho}\,,
\end{aligned} \tag{30.9.3}$$

which confirms the expectation that these components can only depend on the base manifold coordinates x^μ, but not on the fiber coordinates. Note the appearance the second, inhomogeneous term, which indicates that the connection coefficients do not transform as the components of a tensor field, for which no such term appears.

Recall from section 30.2 that we have equivalently expressed as affine connection as a Koszul connection on the tangent bundle, and hence by the corresponding covariant derivative acting on vector fields. Recalling from section 12.1 that we can also pull back vector fields along diffeomorphisms, we arrive at the following alternative and independent definition.

Definition 30.9.2 (Pullback of a linear connection in the tangent bundle)**.** Let N be a manifold equipped with an affine connection ∇ and $\varphi : M \to N$ a diffeomorphism. The *pullback* of ∇ along φ is the affine connection $\varphi^*\nabla$ on M defined such that

$$(\varphi^*\nabla)_X Y = \varphi^*[\nabla_{\varphi^{-1*}X}(\varphi^{-1*}Y)] = \varphi_*^{-1} \circ \nabla_{\varphi_*\circ X\circ\varphi^{-1}}(\varphi_* \circ Y \circ \varphi^{-1}) \circ \varphi \tag{30.9.4}$$

for all $X, Y \in \mathrm{Vect}(M)$.

Also using this formulation we can derive a coordinate expression. Denoting the pullbacks of the vector fields as $X' = \varphi^{-1*}X$ and $Y' = \varphi^{-1*}Y$, these are given in coordinates by

$$X'(x') = X^{\mu}(x)\frac{\partial x'^{\nu}}{\partial x^{\mu}}\partial'_{\nu}\,, \quad Y'(x') = Y^{\mu}(x)\frac{\partial x'^{\nu}}{\partial x^{\mu}}\partial'_{\nu}\,. \tag{30.9.5}$$

Now we can calculate their covariant derivative with respect to the connection ∇ at the point x', from which we find

$$\begin{aligned}
[\nabla_{\varphi^{-1*}X}(\varphi^{-1*}Y)](x') &= X^{\mu}(x)\frac{\partial x'^{\sigma}}{\partial x^{\mu}}\left[\partial'_{\sigma}\left(Y^{\nu}(x)\frac{\partial x'^{\omega}}{\partial x^{\nu}}\right) + \Gamma^{\omega}{}_{\tau\sigma}(x')Y^{\nu}(x)\frac{\partial x'^{\tau}}{\partial x^{\nu}}\right]\partial'_{\omega}\\
&= X^{\mu}(x)\frac{\partial x'^{\sigma}}{\partial x^{\mu}}\left[\frac{\partial x^{\rho}}{\partial x'^{\sigma}}\partial_{\rho}Y^{\nu}(x)\frac{\partial x'^{\omega}}{\partial x^{\nu}} + Y^{\nu}(x)\frac{\partial x^{\rho}}{\partial x'^{\sigma}}\frac{\partial x'^{\omega}}{\partial x^{\rho}\partial x^{\nu}} + \Gamma^{\omega}{}_{\tau\sigma}(x')Y^{\nu}(x)\frac{\partial x'^{\tau}}{\partial x^{\nu}}\right]\partial'_{\omega}\\
&= X^{\mu}(x)\left[\partial_{\mu}Y^{\nu}(x)\frac{\partial x'^{\omega}}{\partial x^{\nu}} + Y^{\nu}(x)\frac{\partial x'^{\omega}}{\partial x^{\mu}\partial x^{\nu}} + \Gamma^{\omega}{}_{\tau\sigma}(x')Y^{\nu}(x)\frac{\partial x'^{\tau}}{\partial x^{\nu}}\frac{\partial x'^{\sigma}}{\partial x^{\mu}}\right]\partial'_{\omega}\,.
\end{aligned} \tag{30.9.6}$$

Taking the pullback to the point x, we thus find

$$\begin{aligned}
[(\varphi^{*}\nabla)_X Y](x) &= X^{\mu}(x)\left[\partial_{\mu}Y^{\nu}(x)\frac{\partial x'^{\omega}}{\partial x^{\nu}} + Y^{\nu}(x)\frac{\partial x'^{\omega}}{\partial x^{\mu}\partial x^{\nu}} + \Gamma^{\omega}{}_{\tau\sigma}(x')Y^{\nu}(x)\frac{\partial x'^{\tau}}{\partial x^{\nu}}\frac{\partial x'^{\sigma}}{\partial x^{\mu}}\right]\frac{\partial x^{\lambda}}{\partial x'^{\omega}}\partial_{\lambda}\\
&= X^{\mu}(x)\left[\partial_{\mu}Y^{\nu}(x) + Y^{\rho}(x)\frac{\partial x^{\nu}}{\partial x'^{\omega}}\left(\frac{\partial x'^{\omega}}{\partial x^{\mu}\partial x^{\rho}} + \Gamma^{\omega}{}_{\tau\sigma}(x')\frac{\partial x'^{\tau}}{\partial x^{\rho}}\frac{\partial x'^{\sigma}}{\partial x^{\mu}}\right)\right]\partial_{\nu}\\
&= X^{\mu}(x)\left[\partial_{\mu}Y^{\nu}(x) + \Gamma'^{\nu}{}_{\rho\mu}(x)Y^{\rho}(x)\right]\partial_{\nu}\,,
\end{aligned} \tag{30.9.7}$$

so that we can finally read off the connection coefficients

$$\Gamma'^{\nu}{}_{\rho\mu}(x) = \frac{\partial x^{\nu}}{\partial x'^{\omega}}\left(\frac{\partial x'^{\omega}}{\partial x^{\mu}\partial x^{\rho}} + \Gamma^{\omega}{}_{\tau\sigma}(x')\frac{\partial x'^{\tau}}{\partial x^{\rho}}\frac{\partial x'^{\sigma}}{\partial x^{\mu}}\right)\,. \tag{30.9.8}$$

Hence, we find the same coordinate expression as from the principal connection approach. This result suggests that both approaches are equivalent, which can be formulated as follows.

Theorem 30.9.1. *The Koszul connection $\varphi^{*}\nabla$ is induced by the principal connection $\varphi_{\circ}^{*}\vartheta$ and vice versa.*

Proof. ▶...◀ ■

Once again recalling that affine connections form sections of an affine bundle modeled over T_2^1M, it follows that the difference $\varphi^{*}\nabla - \nabla$ is a section of T_2^1M. This remains true if instead of a single diffeomorphism φ we consider a one-parameter group $t \mapsto \varphi_t$ of diffeomorphisms, generated by a vector field $X = X^{\mu}\partial_{\mu}$. In analogy to the case of tensor fields, we can thus define the Lie derivative of an affine connection as follows.

Definition 30.9.3 (Lie derivative of an affine connection)**.** Let ∇ be an affine connection and $X \in \mathrm{Vect}(M)$ a vector field on a manifold M. Let $\phi : \mathbb{R} \times M \supseteq U \to M$ be the flow of X. The *Lie derivative* of ∇ with respect to X is the $(1,2)$-tensor field defined by

$$\mathcal{L}_X\nabla = \lim_{t\to 0}\frac{\phi_t^{*}\nabla - \nabla}{t}\,. \tag{30.9.9}$$

It is straightforward to derive a coordinate expression, in analogy to the Lie derivative of a tensor field. For this purpose, recall that the components of the tensor field we aim to calculate are given by the difference of the connection coefficients of the two connections, and so we can write

$$(\mathcal{L}_X\nabla)^\rho{}_{\mu\nu}(x) = \lim_{t\to 0}\frac{\Gamma'^\rho_t{}_{\mu\nu}(x) - \Gamma^\rho{}_{\mu\nu}(x)}{t} = \frac{\mathrm{d}}{\mathrm{d}t}\Gamma'^\rho_t{}_{\mu\nu}(x)\bigg|_{t=0}\,, \tag{30.9.10}$$

where we wrote

$$\Gamma'^\rho_t{}_{\mu\nu}(x) = \frac{\partial x^\rho}{\partial x'^\omega_t}(x'_t(x))\left(\frac{\partial x'^\omega_t}{\partial x^\mu\partial x^\nu}(x) + \Gamma^\omega{}_{\tau\sigma}(x'_t(x))\frac{\partial x'^\tau_t}{\partial x^\rho}(x)\frac{\partial x'^\sigma_t}{\partial x^\mu}(x)\right)\,. \tag{30.9.11}$$

As for a tensor field, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\Gamma^\omega{}_{\tau\sigma}(x'_t(x))\bigg|_{t=0} = X^\lambda(x)\partial_\lambda\Gamma^\omega{}_{\tau\sigma}(x)\,, \tag{30.9.12}$$

where we used the fact that the vector field is related to the flow by

$$X^\mu(x) = \frac{\mathrm{d}}{\mathrm{d}t}x'^\mu_t(x)\bigg|_{t=0}\,. \tag{30.9.13}$$

Further using the relations

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial x'^\nu_t}{\partial x^\mu}(x)\bigg|_{t=0} = \partial_\mu X^\nu(x)\,, \quad \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial x^\mu}{\partial x'^\nu_t}(x'_t(x))\bigg|_{t=0} = -\partial_\nu X^\mu(x)\,, \tag{30.9.14}$$

which we have derived in section 16.2, we are only left with the term

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial x'^\rho_t}{\partial x^\mu\partial x^\nu}(x)\bigg|_{t=0} = \partial_\mu\partial_\nu X^\rho(x). \tag{30.9.15}$$

Putting all terms together, we thus finally arrive at the expression

$$(\mathcal{L}_X\nabla)^\rho{}_{\mu\nu} = X^\sigma\partial_\sigma\Gamma^\rho{}_{\mu\nu} - \partial_\sigma X^\rho\Gamma^\sigma{}_{\mu\nu} + \partial_\mu X^\sigma\Gamma^\rho{}_{\sigma\nu} + \partial_\nu X^\sigma\Gamma^\rho{}_{\mu\sigma} + \partial_\mu\partial_\nu X^\rho\,. \tag{30.9.16}$$

Note that this result differs from that for a tensor field by the appearance of the last term, which originates from the inhomogeneous term in the transformation of the connection coefficients, which contains the second derivative of the coordinate transformation. By construction, we know that this result must constitute the components of a tensor field. We will thus derive another expression, from which this conclusion will be obvious.

Theorem 30.9.2. *The Lie derivative of an affine connection satisfies*

$$(\mathcal{L}_X\nabla)(Y,Z) = \nabla^2_{Z,Y}X - R(Z,X)Y - \nabla_Z(T(Y,X)) + T(\nabla_Z Y, X) \tag{30.9.17}$$

for all vector fields $X, Y, Z \in \mathrm{Vect}(M)$ *.*

Proof. We need to consider the vector field

$$(\varphi_t^*\nabla)_Z Y = \varphi_t^*[\nabla_{\varphi_t^{-1}{}^*Z}(\varphi_t^{-1}{}^*Y)]\,, \tag{30.9.18}$$

which approaches $\nabla_Z Y$ for $t \to 0$, in order to calculate

$$\begin{aligned}(\mathcal{L}_X\nabla)(Y,Z) &= \lim_{t\to 0}\frac{(\varphi_t^*\nabla)_Z Y - \nabla_Z Y}{t}\\ &= \frac{\mathrm{d}}{\mathrm{d}t}(\varphi_t^*\nabla)_Z Y\bigg|_{t=0}\\ &= [X, \nabla_Z Y] - \nabla_{[X,Z]}Y - \nabla_Z[X,Y]\\ &= \nabla_X\nabla_Z Y - \nabla_{\nabla_Z Y}X - T(X,\nabla_Z Y) - \nabla_X\nabla_Z Y + \nabla_Z\nabla_X Y\\ &\quad + R(X,Z)Y + \nabla_Z[T(X,Y) - \nabla_X Y + \nabla_Y X]\\ &= \nabla_Z\nabla_Y X - \nabla_{\nabla_Z Y}X + T(\nabla_Z Y, X) - R(Z,X)Y - \nabla_Z(T(Y,X))\\ &= \nabla^2_{Z,Y}X - R(Z,X)Y - \nabla_Z(T(Y,X)) + T(\nabla_Z Y, X)\,. \quad \blacksquare\end{aligned} \tag{30.9.19}$$

We also illustrate this result in coordinates. Making use of the coordinate expressions of the curvature, the torsion and the covariant derivative in terms of the connection coefficients, one has

$$\begin{aligned}(\mathcal{L}_X\nabla)(Y,Z) &= [Z^\nu Y^\mu \nabla_\nu\nabla_\mu X^\rho - R^\rho{}_{\mu\nu\sigma}Y^\mu Z^\nu X^\sigma - Z^\nu\nabla_\nu(T^\rho{}_{\mu\sigma}Y^\mu X^\sigma) + T^\rho{}_{\mu\sigma}Z^\nu\nabla_\nu Y^\mu X^\sigma]\partial_\rho\\ &= Y^\mu Z^\nu[\nabla_\nu\nabla_\mu X^\rho - R^\rho{}_{\mu\nu\sigma}X^\sigma - \nabla_\nu(X^\sigma T^\rho{}_{\mu\sigma})]\partial_\rho\\ &= Y^\mu Z^\nu\{\partial_\nu(\partial_\mu X^\rho + \Gamma^\rho{}_{\sigma\mu}X^\sigma) + \Gamma^\rho{}_{\omega\nu}(\partial_\mu X^\omega + \Gamma^\omega{}_{\sigma\mu}X^\sigma) - \Gamma^\omega{}_{\mu\nu}(\partial_\omega X^\rho + \Gamma^\rho{}_{\sigma\omega}X^\sigma)\\ &\quad - (\partial_\nu\Gamma^\rho{}_{\mu\sigma} - \partial_\sigma\Gamma^\rho{}_{\mu\nu} + \Gamma^\rho{}_{\omega\nu}\Gamma^\omega{}_{\mu\sigma} - \Gamma^\rho{}_{\omega\sigma}\Gamma^\omega{}_{\mu\nu})X^\sigma - \partial_\nu[X^\sigma(\Gamma^\rho{}_{\sigma\mu} - \Gamma^\rho{}_{\mu\sigma})]\\ &\quad + \Gamma^\rho{}_{\omega\nu}[X^\sigma(\Gamma^\omega{}_{\sigma\mu} - \Gamma^\omega{}_{\mu\sigma})] - \Gamma^\omega{}_{\mu\nu}[X^\sigma(\Gamma^\rho{}_{\sigma\omega} - \Gamma^\rho{}_{\omega\sigma})]\}\partial_\rho\\ &= Y^\mu Z^\nu(X^\sigma\partial_\sigma\Gamma^\rho{}_{\mu\nu} - \partial_\sigma X^\rho\Gamma^\sigma{}_{\mu\nu} + \partial_\mu X^\sigma\Gamma^\rho{}_{\sigma\nu} + \partial_\nu X^\sigma\Gamma^\rho{}_{\mu\sigma} + \partial_\mu\partial_\nu X^\rho)\partial_\rho\\ &= (\mathcal{L}_X\nabla)^\rho{}_{\mu\nu}Y^\mu Z^\nu\partial_\rho\,,\end{aligned} \tag{30.9.20}$$

which agrees with our previous result.

In the proof of theorem 30.9.2 we have encountered the Lie derivative of the covariant derivative of a vector field. It turns out that this is a special case of a more general result, which we show as follows.

Theorem 30.9.3. *The Lie derivative of the covariant derivative of a tensor field A of rank (r,s) satisfies*

$$\mathcal{L}_X(\nabla_Y A) = \nabla_Y(\mathcal{L}_X A) + \nabla_{[X,Y]}A + ▶ \ldots ◀ \tag{30.9.21}$$

and

$$\mathcal{L}_X(\nabla A) = \nabla(\mathcal{L}_X A) + ▶ \ldots ◀ \tag{30.9.22}$$

for all vector fields $X, Y \in \mathrm{Vect}(M)$.

Proof. ▶…◀ ■

In coordinates, this expression becomes ▶…◀

Finally, recall that we have constructed two tensor fields from the connection itself, namely the torsion and the curvature. It is natural to expect that also their Lie derivatives are fully expressed in terms of the Lie derivative of the affine connection, and hence ultimately in terms of covariant derivatives, as well as the curvature and torsion tensor. We now show that this is indeed the case, by proving the following statement.

Theorem 30.9.4. *The Lie derivatives of the curvature and the torsion satisfy*

$$\begin{aligned}(\mathcal{L}_V R)(X,Y)Z &= (\mathcal{L}_V\nabla)(X,\nabla_Y Z) - (\mathcal{L}_V\nabla)(Y,\nabla_X Z)\\ &\quad + \nabla_X[(\mathcal{L}_V\nabla)(Y,Z)] - \nabla_Y[(\mathcal{L}_V\nabla)(X,Z)] - (\mathcal{L}_V\nabla)([X,Y],Z)\end{aligned} \tag{30.9.23}$$

and

$$(\mathcal{L}_V T)(X,Y) = (\mathcal{L}_V\nabla)(X,Y) - (\mathcal{L}_V\nabla)(Y,X) \tag{30.9.24}$$

for all vector fields $V, X, Y, Z \in \mathrm{Vect}(M)$.

Proof. By direct calculation, making use of the Jacobi identity, one finds for the torsion

$$\begin{aligned}(\mathcal{L}_V T)(X,Y) &= [V,T(X,Y)] - T([V,X],Y) - T(X,[V,Y])\\ &= [V,\nabla_X Y - \nabla_Y X - [X,Y]] - \nabla_{[V,X]}Y + \nabla_Y[V,X] + [[V,X],Y]\\ &\quad - \nabla_X[V,Y] + \nabla_{[V,Y]}X + [X,[V,Y]]\\ &= [V,\nabla_X Y] - [V,\nabla_Y X] - \nabla_{[V,X]}Y + \nabla_Y[V,X] - \nabla_X[V,Y] + \nabla_{[V,Y]}X\\ &\quad + [[X,Y],V] + [[V,X],Y] + [[Y,V],X]\\ &= (\mathcal{L}_V\nabla)(X,Y) - (\mathcal{L}_V\nabla)(Y,X)\,,\end{aligned} \tag{30.9.25}$$

while for the curvature one obtains

$$
\begin{aligned}
(\mathcal{L}_V R)(X,Y)Z &= [V, R(X,Y)Z] - R([V,X],Y)Z - R(X,[V,Y])Z - R(X,Y)[V,Z] \\
&= [V, \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z] - \nabla_X\nabla_Y[V,Z] + \nabla_Y\nabla_X[V,Z] + \nabla_{[X,Y]}[V,Z] \\
&\quad - \nabla_{[V,X]}\nabla_Y Z + \nabla_Y\nabla_{[V,X]}Z + \nabla_{[[V,X],Y]}Z - \nabla_X\nabla_{[V,Y]}Z + \nabla_{[V,Y]}\nabla_X Z + \nabla_{[X,[V,Y]]}Z \\
&= [V, \nabla_X\nabla_Y Z] - [V, \nabla_Y\nabla_X Z] - [V, \nabla_{[X,Y]}Z] + \nabla_{[X,Y]}[V,Z] + \nabla_{[V,[X,Y]]}Z \\
&\quad - \nabla_X\nabla_Y[V,Z] + \nabla_Y\nabla_X[V,Z] - \nabla_{[V,X]}\nabla_Y Z + \nabla_Y\nabla_{[V,X]}Z - \nabla_X\nabla_{[V,Y]}Z + \nabla_{[V,Y]}\nabla_X Z \\
&= (\mathcal{L}_V\nabla)(X, \nabla_Y Z) + \nabla_X[V, \nabla_Y Z] - (\mathcal{L}_V\nabla)(Y, \nabla_X Z) - \nabla_Y[V, \nabla_X Z] \\
&\quad - \nabla_X\nabla_Y[V,Z] + \nabla_Y\nabla_X[V,Z] + \nabla_Y\nabla_{[V,X]}Z - \nabla_X\nabla_{[V,Y]}Z - (\mathcal{L}_V\nabla)([X,Y],Z) \\
&= (\mathcal{L}_V\nabla)(X, \nabla_Y Z) - (\mathcal{L}_V\nabla)(Y, \nabla_X Z) \\
&\quad + \nabla_X[(\mathcal{L}_V\nabla)(Y,Z)] - \nabla_Y[(\mathcal{L}_V\nabla)(X,Z)] - (\mathcal{L}_V\nabla)([X,Y],Z)\,. \quad \blacksquare
\end{aligned}
\tag{30.9.26}
$$

To illustrate this result in coordinates, it is most straightforward to start from the right hand sides derived above, so that intermediate terms can be canceled during the calculation. ►. . . ◄

Chapter 31

(Pseudo-)Riemannian geometry

31.1 Riemannian and pseudo-Riemannian metrics

One of the most important and most widely used geometric structures in physics is that of a *metric*. We first give the definition we will be using, before adding a number of clarifications.

Definition 31.1.1 (Pseudo-Riemannian metric). A *pseudo-Riemannian* (or *semi-Riemannian*) *metric* on a manifold M is a symmetric tensor field g of rank $(0,2)$, i.e., a section of the bundle $\mathrm{Sym}^2 T^*M$, such that at each point $p \in M$, it defines a non-degenerate, symmetric bilinear form g_p on T_pM with constant signature on M.

In the literature both terms pseudo-Riemannian and semi-Riemannian metric are in use for a metric whose signature is either unspecified or indefinite. Some authors also use the term *Riemannian* metric for the indefinite case. However, we reserve this term for the positive definite case, which we define as follows.

Definition 31.1.2 (Riemannian metric). A pseudo-Riemannian metric is called *Riemannian* if and only if it is positive definite.

Another case which is of particular importance in physics likewise deserves its own name.

Definition 31.1.3 (Lorentzian metric). A pseudo-Riemannian metric on a manifold of dimension n is called *Lorentzian* if and only if it is of signature $(1, n-1)$.

Here the *signature* refers to the signature of the symmetric bilinear form. Following Sylvester's law of inertia, a vector space equipped with a non-degenerate, symmetric, bilinear form $\langle \bullet, \bullet \rangle$ possesses a (non-unique) basis (ϵ_a), such that

$$\langle \epsilon_a, \epsilon_b \rangle = \eta(\mathbb{e}_a, \mathbb{e}_b) = \eta_{ab} = \begin{cases} -1 & a = b \in \{1, \ldots, k\}\,, \\ 1 & a = b \in \{k+1, \ldots, k+l\}\,, \\ 0 & \text{otherwise}, \end{cases} \tag{31.1.1}$$

with $k+l$ being the dimension of the vector space, where η is the canonical symmetric bilinear form of signature (k,l) on $\mathbb{R}^{k+l}$ with basis $(\mathbb{e}_a)$. Hence, we use the convention that a pseudo-Riemannian metric is said to be of signature (k,l) if the corresponding bilinear form in an orthonormal basis is given by

$$\eta = \operatorname{diag}(\underbrace{-1,\ldots,-1}_{k \text{ times}},\underbrace{1,\ldots,1}_{l \text{ times}})\,. \tag{31.1.2}$$

Note that we demanded in definition 31.1.1 that the signature is constant over the manifold M. We included this condition, since we allow manifolds with multiple connected components. If we considered only connected components, we could drop this condition and rely on the following theorem.

Theorem 31.1.1. *Let M be a manifold and g a symmetric tensor field g of rank $(0,2)$, i.e., a section of the bundle $\operatorname{Sym}^2 T^*M$, such that at each point $p \in M$, it defines a non-degenerate, symmetric bilinear form g_p on T_pM. Then the signature of g is constant on every connected component of M.*

Proof. ▶. . . ◀ ■

Given coordinates (x^μ), we can write a pseudo-Riemannian metric in the coordinate basis

$$g = g_{\mu\nu}\mathrm{d}x^\mu \otimes \mathrm{d}x^\nu\,, \tag{31.1.3}$$

where the components $g_{\mu\nu} = g_{(\mu\nu)}$ constitute a symmetric, non-degenerate matrix.

It should be noted that even though we represent a metric by a section of a vector bundle, namely the symmetric power of the cotangent bundle, the space of all metrics with signature (k,l) does *not* form a vector space, and is *not* in one-to-one correspondence with the sections of a vector bundle, due to the restrictions on the signature and that of non-degeneracy: we are only allowed to multiply a metric g by a positive (or nonzero if $k=l$) real number, and also the sum of two metrics of the same signature does not necessarily lead to another metric of the same signature. Nevertheless, it will turn out that the space of metrics of a given signature is still given by the sections of a fiber bundle, albeit not a vector bundle. We will reveal this bundle in section 31.4.

31.2 Inverse metric

From the condition that the metric is everywhere non-degenerate follows that it possesses an *inverse*, which can be defined in various, equivalent ways. Here we use the following definition.

Definition 31.2.1 (Inverse metric)**.** Let M be a manifold with a pseudo-Riemannian metric $g \in \Gamma(\operatorname{Sym}^2 T^*M)$. Its *inverse* is the unique tensor field $g^{-1} \in \Gamma(\operatorname{Sym}^2 TM)$ which satisfies $\operatorname{tr}_1^1(g \otimes g^{-1}) = \delta$, where $\delta \in \Gamma(T_1^1 M)$ is the unit section.

Here we essentially use the fact that g induces a vector bundle isomorphism $TM \to T^*M$ and define g^{-1} such that it induces the inverse vector bundle isomorphism $T^*M \to TM$; we will properly define these morphisms in section 31.3.

Using the coordinate expression (31.1.3) of the metric, as well as $\delta = \partial_\mu \otimes \mathrm{d}x^\mu = \delta^\mu_\nu \partial_\mu \otimes \mathrm{d}x^\nu$, we can express the inverse metric in the form

$$g^{-1} = g^{\mu\nu}\partial_\mu \otimes \partial_\nu\,, \quad g^{\rho\mu}g_{\rho\nu} = \delta^\mu_\nu\,. \tag{31.2.1}$$

Note that it is conventional to write the components of the inverse metric as $g^{\mu\nu}$, instead of the more cumbersome $(g^{-1})^{\mu\nu}$. Hence, if one expresses the components $g_{\mu\nu}$ of the metric in matrix form, this matrix is invertible, and the components $g^{\mu\nu}$ of the inverse metric are expressed by the inverse matrix.

Now the following is straightforward.

Theorem 31.2.1. *Let M be a manifold with a pseudo-Riemannian metric g of signature (k, l). Then the inverse metric g^{-1} induces a non-degenerate, symmetric, bilinear form with the same signature on every cotangent space.*

Proof. ►...◄ ■

31.3 Musical isomorphisms

Due to the fact that, by definition 31.1.1, a pseudo-Riemannian metric is non-degenerate, and hence possesses an inverse, it induces a bijective mapping from tangent vectors to covectors and vice versa. These mappings are commonly called "musical isomorphisms"[1], and can be defined as follows.

Definition 31.3.1 (Flat isomorphism)**.** Let M be a manifold equipped with a pseudo-Riemannian metric g. The *flat isomorphism* is the unique vector bundle isomorphism $\flat : TM \to T^*M$ such that $X^\flat = \flat \circ X \in \Omega^1(M)$ is the one-form defined by $X^\flat = g(X, \bullet)$ for all $X \in \mathrm{Vect}(M)$.

Definition 31.3.2 (Sharp isomorphism)**.** Let M be a manifold equipped with a pseudo-Riemannian metric g. The *sharp isomorphism* is the unique vector bundle isomorphism $\sharp : T^*M \to TM$ such that $\omega^\sharp = \sharp \circ \omega \in \mathrm{Vect}(M)$ is the vector field defined by $\omega^\sharp = g^{-1}(\omega, \bullet)$ for all $\omega \in \Omega^1(M)$.

In a coordinate basis, we can write these two operations as

$$X^\flat = g_{\mu\nu} X^\mu \, \mathrm{d}x^\nu = X_\nu \, \mathrm{d}x^\nu \tag{31.3.1}$$

and

$$\omega^\sharp = g^{\mu\nu} \omega_\mu \, \partial_\nu = \omega^\nu \, \partial_\nu \,, \tag{31.3.2}$$

where we also introduced the common notation of "lowering" and "raising" indices. The latter, together with the musical notation ♭ for lowering and ♯ for raising the pitch of a note, are the reason for the naming and notation of the musical isomorphisms.

The musical isomorphisms are sometimes useful to express the action of the metric g on vector fields and covector fields. Given vector fields $X, Y \in \mathrm{Vect}(M)$ and covector fields $\alpha, \beta \in \Omega^1(M)$ one immediately writes

$$g(X, Y) = X \lrcorner Y^\flat = Y \lrcorner X^\flat \,, \tag{31.3.3}$$

as well as

$$g^{-1}(\alpha, \beta) = \alpha^\sharp \lrcorner \beta = \beta^\sharp \lrcorner \alpha \,. \tag{31.3.4}$$

[1] As denoted by [Ber02, sec. 15.2]. The term is also used in [Lee13, ch. 13] and [Lee18, ch. 2]. See also [Fec11, sec. 2.4].

31.4 Orthonormal frame bundle

In section 22.6 we have encountered the tangent frame bundle FM as the bundle whose elements over $x \in M$ are linear bijections $f : \mathbb{R}^{\dim M} \to T_xM$. Given a pseudo-Riemannian metric, one may consider a particular set of frames, which relates the metric structure of the tangent spaces to a bilinear form η of the same signature on the vector space $\mathbb{R}^{\dim M}$. We define these frames as follows.

Definition 31.4.1 (Orthonormal frame bundle)**.** Let M be a manifold with a pseudo-Riemannian metric g of signature (k, l) and η the symmetric bilinear form on $\mathbb{R}^{k+l}$ with the same signature. An *orthonormal frame* at $x \in M$ is a bijective linear function $f : \mathbb{R}^{k+l} \to T_xM$ such that $\eta = f^*g$. The set of all orthonormal frames constitutes the *orthonormal frame bundle* $\mathrm{O}(M, g)$ with projection mapping $f : \mathbb{R}^{k+l} \to T_xM$ to $x \in M$.

The defining property which distinguishes an orthonormal frame is that it pulls the metric back to the canonical symmetric bilinear form η of signature (k, l) on $\mathbb{R}^{k+l}$. In other words, a frame f is orthonormal if and only if

$$g(f(v), f(w)) = \eta(v, w) \tag{31.4.1}$$

for all $v, w \in \mathbb{R}^{k+l}$. Using the coordinate expression (31.1.3), the basis expansion (31.1.1) and the basis expansion

$$f(v) = f(v^a \mathbb{e}_a) = v^a f(\mathbb{e}_a) = v^a f_a = f^\mu{}_a v^a \partial_\mu \tag{31.4.2}$$

of a frame f, we find that f is orthonormal if and only if

$$g_{\mu\nu} f^\mu{}_a f^\nu{}_b = \eta_{ab}\,. \tag{31.4.3}$$

Recall from example 15.1.3 that one may define the orthogonal group $\mathrm{O}(k, l)$ as the set of linear bijections $A : \mathbb{R}^{k+l} \to \mathbb{R}^{k+l}$ which leave the bilinear form η invariant, $\eta(Av, Aw) = \eta(v, w)$ for all $v, w \in \mathbb{R}^{k+l}$. This suggests a relation between the orthonormal frame bundle and the group $\mathrm{O}(k, l)$. We state this relation as follows.

Theorem 31.4.1. *The orthonormal frame bundle $\mathrm{O}(M, g)$ over a manifold M with pseudo-Riemannian metric g of signature (k, l) is a principal fiber bundle with structure group $\mathrm{O}(k, l)$, where the right action is given by $f \cdot A = f \circ A$ for $f \in \mathrm{O}(M, g)$ and $A \in \mathrm{O}(k, l)$.*

Proof. We show that $\mathrm{O}(k, l)$ acts freely and transitively on the fibers $F_x(M, g)$ of $\mathrm{O}(M, g)$ for $x \in M$:

- Let $f \in F_x(M, g)$ be an orthonormal frame at x and $A \in \mathrm{O}(k, l)$. Then

  $$g(f(Av), f(Aw)) = \eta(Av, Aw) = \eta(v, w) \tag{31.4.4}$$

  for all $v, w \in \mathbb{R}^{k+l}$, and so $f \circ A \in F_x(M, g)$. Hence, $\mathrm{O}(k, l)$ acts on the fibers $F_x(M, g)$.
- Let $f, f' \in F_x(M, g)$ be orthonormal frames at x. Then define $A = f^{-1} \circ f' : \mathbb{R}^{k+l} \to \mathbb{R}^{k+l}$ as the unique linear bijection satisfying $f' = f \circ A$. Observe that

  $$\eta(Av, Aw) = \eta(f^{-1}(f'(v)), f^{-1}(f'(w))) = g(f'(v), f'(w)) = \eta(v, w) \tag{31.4.5}$$

  for all $v, w \in \mathbb{R}^{k+l}$, and so $A \in \mathrm{O}(k, l)$. This element is unique, and thus the action is free and transitive.

To complete the proof, one still needs to construct the local trivializations of $\mathrm{O}(M, g)$ and show its bundle structure, i.e., that the total space $\mathrm{O}(M, g)$ carries the structure of a smooth manifold and that it defines a fiber bundle. We will not show this here, and remark that this can be done by using an atlas of M. ■

Obviously, the orthonormal frame bundle is canonically included in the general linear frame bundle, $\mathrm{O}(M,g) \hookrightarrow FM$. Recalling the construction of principal bundle reductions detailed in section 20.6, one arrives at the following property.

Theorem 31.4.2. *The orthonormal frame bundle together with the canonical inclusion $\mathrm{O}(M,g) \hookrightarrow FM$ is a $\mathrm{O}(k,l)$-reduction of the general linear frame bundle FM.*

Proof. It follows by construction that the inclusion $\mathrm{O}(M,g) \hookrightarrow FM$ is a principal bundle homomorphism covering the identity, so that the diagram (20.6.1) commutes. ∎

Using the preceding statement about the orthonormal frame bundle, we can now also get a deeper insight into the geometry of the bundle whose sections constitute metrics of a given signature. Recall from definition 31.1.1 that a pseudo-Riemannian metric is a section of the bundle $\mathrm{Sym}^2 T^*M$ with the additional constraints that it is non-degenerate and of constant signature. Hence, for a fixed signature (k,l), for $x \in M$ the fiber over x of the bundle we consider contains only those elements of $\mathrm{Sym}^2 T_x^*M$ which yield a symmetric bilinear form of signature (k,l). This space, in contrast to $\mathrm{Sym}^2 T_x^*M$, is not a vector space, since its zero element would be degenerate, and so the bundle of pseudo-Riemannian metrics is not a vector bundle. It does, however, have another structure, which we can see as follows.

Theorem 31.4.3. *The bundle of pseudo-Riemannian metrics of signature (k,l) is isomorphic to the associated bundle $FM \times_\rho G/H$ with $G = \mathrm{GL}(k+l,\mathbb{R})$, $H = \mathrm{O}(k,l)$ and ρ the canonical left action of G on the coset space G/H.*

Proof. The proof is carried out by the following steps:

1. Let $f \in FM$ with $\varpi(f) = x \in M$ and $A \in G$, so that $AH \in G/H$. This allows us to define a bilinear form $g = (A^{-1} \circ f^{-1})^*\eta$ on T_xM, which thus satisfies
$$g(v,w) = \eta(A^{-1}(f^{-1}(v)), A^{-1}(f^{-1}(w)))\,. \tag{31.4.6}$$
This is clearly of signature (k,l), since $A^{-1} \circ f^{-1} : T_xM \to \mathbb{R}^{k+l}$ is a vector space isomorphism and η is of signature (k,l). Further, it does not depend on the choice of the representative A of AH, since for any $A' = AB$ with $B \in H$ one has
$$\begin{aligned} g'(v,w) &= \eta(A'^{-1}(f^{-1}(v)), A'^{-1}(f^{-1}(w))) \\ &= \eta(B^{-1}(A^{-1}(f^{-1}(v))), B^{-1}(A^{-1}(f^{-1}(w)))) \\ &= \eta(A^{-1}(f^{-1}(v)), A^{-1}(f^{-1}(w))) = g(v,w)\,. \end{aligned} \tag{31.4.7}$$
Hence, it is fully determined by f and AH. Finally, it does not depend on the choice of the representative (f, AH) of $[f, AH]$, since for $B \in G$ and thus another representative
$$(f', A'H) = (f \circ B, B^{-1}AH) \in [f, AH] \tag{31.4.8}$$
one has
$$A'^{-1} \circ f'^{-1} = (B^{-1}A)^{-1} \circ (f \circ B)^{-1} = A^{-1} \circ B \circ B^{-1} \circ f^{-1} = A^{-1} \circ f^{-1}\,. \tag{31.4.9}$$
Hence, g depends only on the element $[f, AH] \in FM \times_\rho G/H$. We thus have a mapping from $FM \times_\rho G/H$ to the bundle of metrics of signature (k,l).

2. To construct the inverse mapping, let g be a non-degenerate bilinear form of signature (k,l) on T_xM and $f \in F_xM$ an orthonormal frame for $x \in M$, i.e., let f be chosen such that $f^*g = \eta$. Then consider the element $[f, eH] \in FM \times_\rho G/H$, where $e \in G$ is the unit element. Note that this does not depend on the choice of f, since for any other orthonormal frame $f' = f \cdot A$ with $A \in H$ we have
$$[f', eH] = [f \cdot A, eH] = [f, AH] = [f, eH]\,. \tag{31.4.10}$$
Hence, it is uniquely defined by the bilinear form g on T_xM. This defines a mapping from the space of bilinear forms of signature (k,l) on T_xM to $F_xM \times_\rho G/H$.

3. Finally, we have to show that the two mappings we have constructed are inverses of each other:
   (a) Starting from a bilinear form g, we choose an orthonormal frame f, to construct $[f, eH]$. Since f is orthonormal, it satisfies $f^*g = \eta$. Hence, $g = (f^{-1})^*\eta$ agrees with the metric constructed from $[f, eH]$.
   (b) Conversely, starting from $[f, AH] \in FM \times_\rho G/H$ we construct the metric $g = (A^{-1} \circ f^{-1})^*\eta$. By construction, it follows that $f \cdot A$ is an orthonormal frame. For the mapping in the converse direction, we can thus choose $f \cdot A$, and obtain the element $[f \cdot A, eH]$. From the action of G which defines the elements of $FM \times_\rho G/H$ we finally find $[f \cdot A, eH] = [f, AH]$.

   Hence, the two mappings we have constructed are inverses of each other, and we find a one-to-one correspondence between bilinear forms of signature (k, l) on T_xM and elements of $F_xM \times_\rho G/H$. Performing this for all $x \in M$, we obtain a fiber bundle isomorphism. ■

31.5 Twisted volume form

From a pseudo-Riemannian metric we can define a number of other interesting geometric objects. In this section we will discuss the following construction, which is most easily defined in terms of corresponding principal bundles.

Definition 31.5.1 (Twisted volume form of a Riemannian metric)**.** Let M be a manifold with a pseudo-Riemannian metric g. Its *twisted volume form* is the section $\mathrm{vol}_g \in \Gamma(D^-_{-1}(TM))$ defined such that for all $f \in \mathrm{O}(M, g)$ holds

$$[f, 1]_{\rho^-_{-1}} = \mathrm{vol}_g(\varpi(f)) \,, \tag{31.5.1}$$

where $\varpi : \mathrm{O}(M, g) \to M$ is the orthonormal frame bundle.

To see that this definition is independent of the representative f, let $f' = f \cdot A$ with $A \in \mathrm{O}(k, l)$ be another orthonormal frame over the same base point $\varpi(f) = \varpi(f')$. Then we have

$$[f', 1]_{\rho^-_{-1}} = [f \cdot A, 1]_{\rho^-_{-1}} = [f, \rho^-_{-1}(A, 1)]_{\rho^-_{-1}} = [f, |\det A|]_{\rho^-_{-1}} = [f, 1]_{\rho^-_{-1}} \,, \tag{31.5.2}$$

since $|\det A| = 1$ for an orthogonal matrix. In other words, we define the twisted volume form vol_g such that all g-orthonormal frames are vol_g-normalized. We can also state this as follows.

Theorem 31.5.1. *The normalized frame bundle* $\mathrm{SL}^{\pm}(M, \mathrm{vol}_g)$ *of the twisted volume form* vol_g *is given by*

$$\mathrm{SL}^{\pm}(M, \mathrm{vol}_g) = \{f \cdot A, f \in \mathrm{O}(M, g), A \in \mathrm{SL}^{\pm}(k, \mathbb{R})\} \,. \tag{31.5.3}$$

Proof. ►. . . ◄ ■

Theorem 31.5.2. *The orthonormal frame bundle* $\mathrm{O}(M, g)$ *of a pseudo-Riemannian metric with signature* (k, l) *is a* $\mathrm{O}(k, l)$ *reduction of the normalized frame bundle* $\mathrm{SL}^{\pm}(M, \mathrm{vol}_g)$ *of the twisted volume form* vol_g *induced by* g*.*

Proof. ►. . . ◄ ■

►Show relation with $\det g$ and derive coordinate expression.◄

Theorem 31.5.3. *The twisted volume form* $\mathrm{vol}_g \in \Gamma(D^-_{-1}(TM))$ *induced by a pseudo-Riemannian metric* g *of signature* (k, l) *satisfies*

$$\det g = (-1)^k \mathrm{vol}_g \otimes \mathrm{vol}_g \,. \tag{31.5.4}$$

Proof. ►...◄ ■

31.6 Differential forms on Riemannian manifolds

We have seen in section 31.2 that a metric, due to being non-degenerate, possesses an inverse, which is a non-degenerate bilinear form acting on covector fields, or one-forms. This bilinear form can be generalized also to differential forms of higher degree as follows.

Definition 31.6.1 (Riemannian bilinear form on differential forms)**.** Let M be a manifold with a pseudo-Riemannian metric g. The *induced bilinear form* on $\Omega^q(M)$ is the unique bilinear form $\langle \bullet, \bullet \rangle_g$ such that

$$\langle \alpha_1 \wedge \ldots \wedge \alpha_q, \beta_1 \wedge \ldots \wedge \beta_q \rangle_g = \det g^{-1}(\alpha_\bullet, \beta_\bullet) \tag{31.6.1}$$

for all $\alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_q \in \Omega^1(M)$.

This definition needs a few explanations. First note that $g^{-1}(\alpha_i, \beta_j) \in C^\infty(M)$ denotes the bilinear form g^{-1} acting on the one-forms α_i and β_j, in coordinates given by

$$g^{-1}(\alpha_i, \beta_j) = g^{\mu\nu} \alpha_{i\,\mu} \beta_{j\,\nu} \,, \tag{31.6.2}$$

where i, j label the one-forms $\alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_q \in \Omega^1(M)$, while a, b label the components of each one-form in the coordinate basis $\mathrm{d}x^a$. Calculating this for each pair i, j of labels, one obtains a matrix, and so one can take the determinant

$$\langle \alpha_1 \wedge \ldots \wedge \alpha_q, \beta_1 \wedge \ldots \wedge \beta_q \rangle_g = \det \begin{pmatrix} g^{-1}(\alpha_1, \beta_1) & \cdots & g^{-1}(\alpha_1, \beta_q) \\ \vdots & \ddots & \vdots \\ g^{-1}(\alpha_q, \beta_1) & \cdots & g^{-1}(\alpha_q, \beta_q) \end{pmatrix} . \tag{31.6.3}$$

Now recall that we can also write the determinant with the help of the totally antisymmetric Levi-Civita symbol as

$$\det g^{-1}(\alpha_\bullet, \beta_\bullet) = \frac{1}{q!} \epsilon^{i_1 \cdots i_q} \epsilon^{j_1 \cdots j_q} g^{-1}(\alpha_{i_1}, \beta_{j_1}) \cdots g^{-1}(\alpha_{i_q}, \beta_{j_q}) \,, \tag{31.6.4}$$

which in components becomes

$$\det g^{-1}(\alpha_\bullet, \beta_\bullet) = \frac{1}{q!} \epsilon^{i_1 \cdots i_q} \epsilon^{j_1 \cdots j_q} g^{\mu_1 \nu_1} \cdots g^{\mu_q \nu_q} \alpha_{i_1\,\mu_1} \cdots \alpha_{i_q\,\mu_q} \beta_{j_1\,\nu_1} \cdots \beta_{j_q\,\nu_q} \,. \tag{31.6.5}$$

From this expression we can recognize the exterior product

$$\alpha_1 \wedge \ldots \wedge \alpha_q = \alpha_{1\,\mu_1} \cdots \alpha_{q\,\mu_q} \mathrm{d}x^{\mu_1} \wedge \ldots \wedge \mathrm{d}x^{\mu_q} = \frac{1}{q!} \epsilon^{i_1 \cdots i_q} \alpha_{i_1\,\mu_1} \cdots \alpha_{i_q\,\mu_q} \mathrm{d}x^{\mu_1} \wedge \ldots \wedge \mathrm{d}x^{\mu_q} \,, \tag{31.6.6a}$$

$$\beta_1 \wedge \ldots \wedge \beta_q = \beta_{1\,\nu_1} \cdots \beta_{q\,\nu_q} \mathrm{d}x^{\nu_1} \wedge \ldots \wedge \mathrm{d}x^{\nu_q} = \frac{1}{q!} \epsilon^{j_1 \cdots j_q} \beta_{j_1\,\nu_1} \cdots \beta_{j_q\,\nu_q} \mathrm{d}x^{\nu_1} \wedge \ldots \wedge \mathrm{d}x^{\nu_q} \,, \tag{31.6.6b}$$

where we made use of the fact that the exterior product is totally antisymmetric with respect to reordering of the factors. Defining

$$\alpha = \frac{1}{q!}\alpha_{\mu_1\cdots\mu_q}\mathrm{d}x^{\mu_1} \wedge \ldots \wedge \mathrm{d}x^{\mu_q}\,, \quad \beta = \frac{1}{q!}\beta_{\nu_1\cdots\nu_q}\mathrm{d}x^{\nu_1} \wedge \ldots \wedge \mathrm{d}x^{\nu_q}\,, \tag{31.6.7}$$

we thus have

$$\alpha_{\mu_1\cdots\mu_q} = \epsilon^{i_1\cdots i_q}\alpha_{i_1\,\mu_1}\cdots\alpha_{i_q\,\mu_q}\,, \quad \beta_{\nu_1\cdots\nu_q} = \epsilon^{j_1\cdots j_q}\beta_{j_1\,\nu_1}\cdots\beta_{j_q\,\nu_q}\,. \tag{31.6.8}$$

Since a bilinear form is by definition bilinear, we thus have the general expression

$$\langle\alpha,\beta\rangle_g = \frac{1}{q!}g^{\mu_1\nu_1}\cdots g^{\mu_q\nu_q}\alpha_{\mu_1\cdots\mu_q}\beta_{\nu_1\cdots\nu_q}\,. \tag{31.6.9}$$

In particular, we see that it reduces to g^{-1} for one-forms in the case $q = 1$. Also the following is now straightforward.

Theorem 31.6.1. *The bilinear form $\langle\bullet,\bullet\rangle_g$ induced by a pseudo-Riemannian metric is non-degenerate.*

Proof. At each point $x \in M$, we can pick an orthogonal basis (ϵ^a) of the cotangent space T_x^*M, so that $g^{-1}(\epsilon^a,\epsilon^b) = \eta^{ab}$. This generates a basis

$$(\epsilon^{a_1} \wedge \ldots \wedge \epsilon^{a_q}, 1 \leq a_1 < \ldots < a_q \leq \dim M)\,. \tag{31.6.10}$$

Calculating the bilinear form on two basis vectors yields

$$\langle\epsilon^{a_1} \wedge \ldots \wedge \epsilon^{a_q}, \epsilon^{b_1} \wedge \ldots \wedge \epsilon^{b_q}\rangle_g = \eta^{a_1b_1}\cdots\eta^{a_qb_q}\,, \tag{31.6.11}$$

where we used the fact that any term in the matrix $g^{-1}(\epsilon^{a_i},\epsilon^{b_j})$ appearing in the definition of the bilinear form is equal to $\eta^{a_ib_j}$ and thus non-zero if and only if $a_i = b_j$, which, together with the fact that the indices a_i and b_j are in strictly ascending order, shows that this matrix is non-degenerate only if $a_i = b_i$ for all $1 \leq i \leq q$, in which case it becomes $\mathrm{diag}(\eta^{a_1b_1},\ldots,\eta^{a_qb_q})$. Hence, the basis (31.6.10) is orthonormal, and so the bilinear form $\langle\bullet,\bullet\rangle_g$ is non-degenerate. ■

The bilinear form has a number of properties which are often useful in calculations. We thus show the following helpful formula.

Theorem 31.6.2. *The bilinear form $\langle\bullet,\bullet\rangle_g$ induced by a pseudo-Riemannian metric satisfies*

$$\langle\alpha,\gamma\wedge\beta\rangle_g = \langle\gamma^\sharp \lrcorner\, \alpha,\beta\rangle_g \tag{31.6.12}$$

for all q-forms $\alpha \in \Omega^q(M)$, $(q-1)$-forms $\beta \in \Omega^{q-1}(M)$ and one-forms $\gamma \in \Omega^1(M)$.

Proof. It is sufficient to show this for differential forms

$$\alpha = \alpha_1 \wedge \ldots \alpha_q\,, \quad \beta = \beta_1 \wedge \ldots \beta_{q-1}\,, \tag{31.6.13}$$

and conclude on the general case by linearity. In this case we have

$$\begin{aligned}
\gamma^\sharp \lrcorner\, \alpha &= (\gamma^\sharp \lrcorner\, \alpha_1) \wedge \alpha_2 \wedge \ldots \wedge \alpha_q - \alpha_1 \wedge (\gamma^\sharp \lrcorner\, \alpha_2) \wedge \ldots \wedge \alpha_q + \ldots \\
&= \sum_{i=1}^q (-1)^{q-1}(\gamma^\sharp \lrcorner\, \alpha_i) \wedge \alpha_1 \wedge \ldots \wedge \widehat{\alpha_i} \wedge \ldots \wedge \alpha_q \\
&= \sum_{i=1}^q (-1)^{q-1}g^{-1}(\gamma,\alpha_i)\alpha_1 \wedge \ldots \wedge \widehat{\alpha_i} \wedge \ldots \wedge \alpha_q\,.
\end{aligned} \tag{31.6.14}$$

Now we find that

$$
\begin{aligned}
\langle \gamma^\sharp \lrcorner \alpha, \beta\rangle_g &= g^{-1}(\gamma, \alpha_1) \det \begin{pmatrix} g^{-1}(\alpha_2, \beta_1) & \cdots & g^{-1}(\alpha_2, \beta_{q-1}) \\ \vdots & \ddots & \vdots \\ g^{-1}(\alpha_q, \beta_1) & \cdots & g^{-1}(\alpha_q, \beta_{q-1}) \end{pmatrix} \\
&\quad - g^{-1}(\gamma, \alpha_2) \det \begin{pmatrix} g^{-1}(\alpha_1, \beta_1) & \cdots & g^{-1}(\alpha_1, \beta_{q-1}) \\ g^{-1}(\alpha_3, \beta_1) & \cdots & g^{-1}(\alpha_3, \beta_{q-1}) \\ \vdots & \ddots & \vdots \\ g^{-1}(\alpha_q, \beta_1) & \cdots & g^{-1}(\alpha_q, \beta_{q-1}) \end{pmatrix} + \cdots \\
&= \det \begin{pmatrix} g^{-1}(\alpha_1, \gamma) & g^{-1}(\alpha_1, \beta_1) & \cdots & g^{-1}(\alpha_1, \beta_{q-1}) \\ \vdots & \vdots & \ddots & \vdots \\ g^{-1}(\alpha_q, \gamma) & g^{-1}(\alpha_q, \beta_1) & \cdots & g^{-1}(\alpha_q, \beta_{q-1}) \end{pmatrix} \\
&= \langle \alpha, \gamma \wedge \beta\rangle_g \,. \quad \blacksquare
\end{aligned} \tag{31.6.15}
$$

In the definition 31.6.1 we have considered the case that α and β are differential forms, and bilinearity implies that if one multiplies any of these forms by a function $f \in C^\infty(M, \mathbb{R})$, the result is the same as if one multiplied the whole expression instead,

$$
\langle f\alpha, \beta\rangle_g = \langle \alpha, f\beta\rangle_g = f\langle \alpha, \beta\rangle_g \,. \tag{31.6.16}
$$

One can straightforwardly extend the definition of the bilinear form by demanding that this also holds for pseudoscalars. We thus define the following.

Definition 31.6.2 (Riemannian bilinear form on twisted differential forms)**.** Let M be a manifold with a pseudo-Riemannian metric g. The *induced bilinear form* on $\Omega^q(M)$ is extended to the space $\bar{\Omega}^q(M)$ of twisted differential forms by demanding that

$$
\langle \alpha \otimes \mathfrak{a}, \beta\rangle_g = \langle \alpha, \beta \otimes \mathfrak{a}\rangle_g = \langle \alpha, \beta\rangle_g \otimes \mathfrak{a} \,, \quad \langle \alpha \otimes \mathfrak{a}, \beta \otimes \mathfrak{b}\rangle_g = \langle \alpha, \beta\rangle_g \otimes \mathfrak{a} \otimes \mathfrak{b} \tag{31.6.17}
$$

for all $\alpha, \beta \in \Omega^q(M)$ and $\mathfrak{a}, \mathfrak{b} \in \bar{\Omega}^0(M)$.

Note that if one of the two arguments is a twisted differential form, then the (extended) bilinear form yields a pseudoscalar, while for two twisted differential forms one obtains a scalar function. It is straightforward to check that the properties we have proven for the bilinear form $\langle \bullet, \bullet\rangle_g$ hold also for its extension to twisted forms, since all operations used in their proofs are linear also with respect to multiplication by pseudoscalars.

31.7 Hodge dual

Using the results from the previous section, we can now show the following.

Theorem 31.7.1. *Let $\beta \in \Omega^q(M)$ be a q-form on a pseudo-Riemannian manifold (M, g) of dimension $\dim M = n$. Then there exists a unique twisted $(n-q)$-form $\star\beta \in \bar{\Omega}^{n-q}(M)$ such that*

$$
\alpha \wedge \star\beta = \langle \alpha, \beta\rangle_g \operatorname{vol}_g \tag{31.7.1}
$$

for all $\alpha \in \Omega^q(M)$.

Proof. ▶...◀ ■

This leads to the following definition.

Definition 31.7.1 (Hodge operator)**.** Let (M, g) be a pseudo-Riemannian manifold of dimension $\dim M = n$. The *Hodge operator* is the operator $\star : \Omega^q(M) \to \bar{\Omega}^{n-q}(M)$ constructed in theorem 31.7.1, and is extended to twisted differential forms by demanding

$$\star(\alpha \otimes \mathfrak{a}) = (\star\alpha) \otimes \mathfrak{a} \tag{31.7.2}$$

for all $\alpha \in \Omega^q(M)$ and $\mathfrak{a} \in \bar{\Omega}^0(M)$.

While theorem 31.7.1 uniquely defines the Hodge operator, it is useful to consider a more constructive approach. Recalling that any differential q-form can be obtained as an exterior product of q one-forms, one would like to have a formula for the Hodge operator applied to such a product. We can arrive at this construction in two steps. We start with the following.

Theorem 31.7.2. *The Hodge operator satisfies* $\star 1 = \mathrm{vol}_g$*.*

Proof. This is simply a special case of theorem 31.7.1 with $\alpha = \beta = 1$. ∎

Now we continue with the following.

Theorem 31.7.3. *The Hodge operator satisfies*

$$\star(\beta \wedge \gamma) = \gamma^\sharp \lrcorner \star\beta \tag{31.7.3}$$

for all q-forms $\beta \in \Omega^q(M)$ and one-forms $\gamma \in \Omega^1(M)$.

Proof. For all $(q+1)$-forms $\alpha \in \Omega^{q+1}(M)$ we have

$$\begin{aligned} \alpha \wedge \star(\beta \wedge \gamma) &= \langle \alpha, \beta \wedge \gamma \rangle_g \, \mathrm{vol}_g \\ &= (-1)^q \langle \gamma^\sharp \lrcorner \alpha, \beta \rangle_g \, \mathrm{vol}_g \\ &= (-1)^q (\gamma^\sharp \lrcorner \alpha) \wedge \star\beta \\ &= \alpha \wedge (\gamma^\sharp \lrcorner \star\beta)\,, \end{aligned} \tag{31.7.4}$$

where in the last line we used the identity

$$0 = \gamma^\sharp \lrcorner (\alpha \wedge \star\beta) = (\gamma^\sharp \lrcorner \alpha) \wedge \star\beta - (-1)^q \alpha \wedge (\gamma^\sharp \lrcorner \star\beta)\,, \tag{31.7.5}$$

which follows from the fact that $\alpha \wedge \star\beta$ is a differential form of rank $n+1$ on a manifold of dimension $\dim M = n$ and thus vanishes identically. Since this holds for all α, and the Hodge operator is unique, the proposition follows. ∎

From a repeated application of the theorem above now follows immediately that we can express the Hodge operator applied to the q-fold exterior product of one-forms as

$$\begin{aligned} \star(\beta_1 \wedge \ldots \wedge \beta_q) &= \beta_q^\sharp \lrcorner \star(\beta_1 \wedge \ldots \wedge \beta_{q-1}) \\ &= \beta_q^\sharp \lrcorner \beta_{q-1}^\sharp \lrcorner \star(\beta_1 \wedge \ldots \wedge \beta_{q-2}) \\ &= \ldots \\ &= \beta_q^\sharp \lrcorner \ldots \lrcorner \beta_1^\sharp \lrcorner \star 1 \\ &= \beta_q^\sharp \lrcorner \ldots \lrcorner \beta_1^\sharp \lrcorner \mathrm{vol}_g\,, \end{aligned} \tag{31.7.6}$$

and so we have an explicit formula. We can use this formula to show the following.

Theorem 31.7.4. *The Hodge dual of the volume form on a pseudo-Riemannian manifold (M, g) of signature (k, l) is given by*

$$\star \operatorname{vol}_g = (-1)^k \,. \tag{31.7.7}$$

Proof. ▶. . . ◀ ■

Theorem 31.7.5. *The bilinear form and Hodge operator on a pseudo-Riemannian manifold (M, g) of signature (k, l) satisfy*

$$\langle \star\alpha, \star\beta \rangle_g = (-1)^k \langle \alpha, \beta \rangle_g \tag{31.7.8}$$

for all q-forms $\alpha, \beta \in \Omega^q(M)$.

Proof. ▶. . . ◀ ■

Theorem 31.7.6. *The Hodge operator on a pseudo-Riemannian manifold (M, g) of signature (k, l) with dimension $\dim M = n = k + l$ satisfies*

$$\star\star\alpha = (-1)^{q(n-q)+k} \alpha \tag{31.7.9}$$

for all q-forms $\alpha \in \Omega^q(M)$.

Proof. ▶. . . ◀ ■

31.8 Codifferential

Definition 31.8.1 (Codifferential). Let (M, g) be a pseudo-Riemannian manifold of dimension $\dim M = n$. The *codifferential* is the unique operator $\delta : \Omega^q(M) \to \Omega^{q-1}(M)$ such that

$$\int_M \langle \alpha, \delta\beta \rangle_g \operatorname{vol}_g = \int_M \langle \mathrm{d}\alpha, \beta \rangle_g \operatorname{vol}_g \tag{31.8.1}$$

for all $q \in \{1, \ldots, n\}$ and compactly supported $\alpha \in \Omega^{q-1}(M)$ and $\beta \in \Omega^q(M)$.

Theorem 31.8.1. *The codifferential of a q-form $\beta \in \Omega^q(M)$ on a pseudo-Riemannian (M, g) manifold of signature(k, l) with dimension $\dim M = n = k + l$ is given by*

$$\delta\beta = -(-1)^{n(q-1)+k} \star \mathrm{d} \star \beta \,. \tag{31.8.2}$$

Proof. By Stokes' theorem we have

$$\begin{aligned}
0 &= \int_M \mathrm{d}(\alpha \wedge \star\beta) \\
&= \int_M (\mathrm{d}\alpha \wedge \star\beta + (-1)^q \alpha \wedge \mathrm{d}\star\beta) \\
&= \int_M (\langle \mathrm{d}\alpha, \beta \rangle_g + (-1)^q \langle \alpha, \star^{-1} \mathrm{d}\star\beta \rangle_g) \operatorname{vol}_g \\
&= \int_M (\langle \mathrm{d}\alpha, \beta \rangle_g + (-1)^{q+q(n-q)+k} \langle \alpha, \star\mathrm{d}\star\beta \rangle_g) \operatorname{vol}_g \\
&= \text{▶ . . . ◀}
\end{aligned} \tag{31.8.3}$$

▶. . . ◀ ■

Theorem 31.8.2. *The codifferential satisfies $\delta^2 = 0$.*

Proof. ▶. . . ◀ ■

31.9 Laplace-de Rham operator

Definition 31.9.1 (Laplace-de Rham operator)**.** Let (M, g) be a pseudo-Riemannian manifold. The *Laplace-de Rham operator* is the operator $\hat{\triangle} = \mathrm{d}\delta + \delta\mathrm{d} : \Omega^q(M) \to \Omega^q(M)$.

Theorem 31.9.1. *The Laplace-de Rham operator commutes with the hodge operator,* $\star\hat{\triangle}\alpha = \hat{\triangle}\star\alpha$ *for all* $\alpha \in \Omega^q(M)$*.*

Proof. ►. . . ◄ ■

Theorem 31.9.2. *The Laplace-de Rham operator on a pseudo-Riemannian manifold* (M, g) *satisfies*

$$\int_M \langle\alpha, \hat{\triangle}\beta\rangle_g \,\mathrm{vol}_g = \int_M \langle\hat{\triangle}\alpha, \beta\rangle_g \,\mathrm{vol}_g \tag{31.9.1}$$

for all compactly supported q*-forms* $\alpha, \beta \in \Omega^q(M)$*.*

Proof. ►. . . ◄ ■

Definition 31.9.2 (Harmonic differential form)**.** A differential q-form $\alpha \in \Omega^q(M)$ is called *harmonic* if and only if $\hat{\triangle}\alpha = 0$. The space of all harmonic q-forms is denoted $\mathcal{H}^q_{\triangle}(M)$.

Theorem 31.9.3. *If a differential form* $\alpha \in \Omega^q(M)$ *is harmonic, then also* $\star\alpha$ *is harmonic, and vice versa.*

Proof. ►. . . ◄ ■

Theorem 31.9.4. *If* (M, g) *is a Riemannian manifold (i.e., the metric* g *is positive definite), then a differential form* $\alpha \in \Omega^q(M)$ *is harmonic,* $\hat{\triangle}\alpha = 0$*, if and only if* $\mathrm{d}\alpha = 0$ *and* $\delta\alpha = 0$*.*

Proof. ►. . . ◄ ■

31.10 Levi-Civita connection

Given a pseudo-Riemannian metric, one may define another geometric object which is naturally compatible with the metric. Since the metric is a tensor field, one may ask for a connection which leaves this tensor field invariant. Indeed one finds that such connections exist. Moreover, it turns out that there exists a unique connection which is singled out by the additional demand that its torsion vanishes. This we formulate as follows.

Theorem 31.10.1. *Let* M *be a manifold with a pseudo-Riemannian metric* g*. Then there exists a unique linear connection* ∇ *on* TM *such that the metric is covariantly constant,* $\nabla g = 0$*, and the torsion of* ∇ *vanishes,* $\nabla_X Y - \nabla_Y X - [X, Y] = 0$ *for all* $X, Y \in \mathrm{Vect}(M)$*.*

Proof. Let $X, Y, Z \in \mathrm{Vect}(M)$ be vector fields. Recall that the metric, being a section of the tensor product bundle $\mathrm{Sym}^2 T^*M$, is covariantly constant if and only if

$$0 = (\nabla_Z g)(X, Y) = Z(g(X, Y)) - g(\nabla_Z X, Y) - g(X, \nabla_Z Y)\,. \tag{31.10.1}$$

Consider the linear combination

$$X(g(Y,Z))+Y(g(X,Z))-Z(g(X,Y)) = g(\nabla_X Y+\nabla_Y X,Z)+g(\nabla_X Z-\nabla_Z X,Y)+g(\nabla_Y Z-\nabla_Z Y,X)\,, \tag{31.10.2}$$

where we used the fact that g is symmetric and bilinear. Assuming that ∇ is torsion-free, the right hand side evaluates to

$$2g(\nabla_X Y,Z) - g([X,Y],Z) + g([X,Z],Y) + g([Y,Z],X)\,. \tag{31.10.3}$$

Solving this yields the *Koszul formula*

$$g(\nabla_X Y,Z) = \frac{1}{2}\left\{X(g(Y,Z)) + Y(g(X,Z)) - Z(g(X,Y)) + g([X,Y],Z) - g([X,Z],Y) - g([Y,Z],X)\right\}\,. \tag{31.10.4}$$

Since Z is arbitrary and g is non-degenerate, this uniquely defines $\nabla_X Y$ for all vector fields $X,Y \in \mathrm{Vect}(M)$. ■

This connection also has its own name, and we define:

Definition 31.10.1 (Levi-Civita connection)**.** The unique metric compatible, torsion-free connection from theorem 31.10.1 is called the *Levi-Civita connection* of the metric g.

From the Koszul formula (31.10.4) one may easily derive a coordinate expression for the connection coefficients of the Levi-Civita connection. For arbitrary vector fields $X,Y,Z \in \mathrm{Vect}(M)$, we can write the left hand side in the form

$$g(\nabla_X Y,Z) = g(\nabla_{X^\mu\partial_\mu} Y^\nu\partial_\nu, Z^\rho\partial_\rho) = g_{\mu\nu}X^\rho(\partial_\rho Y^\mu + \Gamma^\mu{}_{\sigma\rho}Y^\sigma)Z^\nu\,. \tag{31.10.5}$$

Using the Koszul formula, this is equal to

$$\begin{aligned} g(\nabla_X Y,Z) &= \frac{1}{2}\big\{X^\mu\partial_\mu(g_{\nu\rho}Y^\nu Z^\rho) + Y^\mu\partial_\mu(g_{\nu\rho}X^\nu Z^\rho) - Z^\mu\partial_\mu(g_{\nu\rho}X^\nu Y^\rho) \\ &\quad + g_{\mu\nu}\left[(X^\rho\partial_\rho Y^\mu - Y^\rho\partial_\rho X^\mu)Z^\nu - (X^\rho\partial_\rho Z^\mu - Z^\rho\partial_\rho X^\mu)Y^\nu - (Y^\rho\partial_\rho Z^\mu - Z^\rho\partial_\rho Y^\mu)X^\nu\right]\big\} \\ &= \frac{1}{2}\partial_\mu g_{\nu\rho}(X^\mu Y^\nu Z^\rho + Y^\mu X^\nu Z^\rho - Z^\mu X^\nu Y^\rho) + g_{\mu\nu}X^\rho\partial_\rho Y^\mu Z^\nu \\ &= g_{\mu\nu}X^\rho\left[\partial_\rho Y^\mu + \frac{1}{2}g^{\mu\sigma}(\partial_\omega g_{\sigma\rho} + \partial_\rho g_{\omega\sigma} - \partial_\sigma g_{\omega\rho})Y^\omega\right]Z^\nu\,, \end{aligned} \tag{31.10.6}$$

where most of the derivatives appearing in the first step cancel. By comparison of the last two coordinate expressions one reads off the formula for the connection coefficients

$$\Gamma^\mu{}_{\nu\rho} = \frac{1}{2}g^{\mu\sigma}(\partial_\nu g_{\sigma\rho} + \partial_\rho g_{\nu\sigma} - \partial_\sigma g_{\nu\rho})\,, \tag{31.10.7}$$

which, in the particular case of the Levi-Civita connection, are also called *Christoffel symbols*. With this formula one now easily verifies also in components the metric compatibility,

$$\nabla_\mu g_{\nu\rho} = \partial_\mu g_{\nu\rho} - \Gamma^\sigma{}_{\nu\mu}g_{\sigma\rho} - \Gamma^\sigma{}_{\rho\mu}g_{\nu\sigma} = 0\,, \tag{31.10.8}$$

and vanishing torsion,

$$T^\mu{}_{\nu\rho} = \Gamma^\mu{}_{\rho\nu} - \Gamma^\mu{}_{\nu\rho} = 0\,. \tag{31.10.9}$$

31.11 Laplace-Beltrami operator

31.12 Curvature tensors

Recall from section 30.3 that the curvature tensor of an affine connection can be written in the form

$$R^{\mu}{}_{\nu\rho\sigma}\partial_{\mu} \otimes \mathrm{d}x^{\nu} \otimes \mathrm{d}x^{\rho} \otimes \mathrm{d}x^{\sigma} , \tag{31.12.1}$$

and is antisymmetric in the last two indices. In the case of the Levi-Civita connection defined in the previous section, we have a number of further identities, which we list in the following, and which can be used in order to transform expressions involving the curvature tensor and its derivatives. It is helpful to lower the first index with the metric,

$$R_{\mu\nu\rho\sigma} = g_{\mu\omega}R^{\omega}{}_{\nu\rho\sigma} , \tag{31.12.2}$$

and use the metric also for further lowering and raising of indices, as usual. Here we list the following identities:

1. Symmetries of the Riemann tensor:
$$R_{\mu\nu\rho\sigma} = -R_{\nu\mu\rho\sigma} = -R_{\mu\nu\sigma\rho} = R_{\rho\sigma\mu\nu} . \tag{31.12.3}$$
2. First Bianchi identity:
$$R_{\mu[\nu\rho\sigma]} = 0 \quad \Leftrightarrow \quad R_{\mu\nu\rho\sigma} + R_{\mu\rho\sigma\nu} + R_{\mu\sigma\nu\rho} = 0 . \tag{31.12.4}$$
3. Second Bianchi identity:
$$R_{\mu\nu[\rho\sigma;\omega]} = 0 \quad \Leftrightarrow \quad R_{\mu\nu\rho\sigma;\omega} + R_{\mu\nu\sigma\omega;\rho} + R_{\mu\nu\omega\rho;\sigma} = 0 . \tag{31.12.5}$$

Recall from definition 30.3.2 that an affine connection defines a tensor field

Theorem 31.12.1. *The Ricci tensor of the Levi-Civita connection of a pseudo-Riemannian metric is symmetric.*

Proof. ▶ . . . ◀ ■

Definition 31.12.1 (Ricci scalar)**.** Let M be a manifold with a pseudo-Riemannian metric g. The *Ricci scalar* (or *scalar curvature*) is the metric trace of the Ricci curvature tensor,

$$\mathcal{R} = \mathrm{tr}_1^1 \, \mathrm{tr}_2^2 \, g^{-1} \otimes R\!\!\!/ . \tag{31.12.6}$$

$$\mathcal{R} = g^{\mu\nu}R_{\mu\nu} \tag{31.12.7}$$

31.13 Ricci decomposition

31.14 Geodesics

31.15 Isometries

Given a pseudo-Riemannian manifold (M, g), one is often interested in diffeomorphisms $\varphi : M \to M$ which leave the metric invariant. These carry their own name:

Definition 31.15.1 (Isometry)**.** Let (M, g) be a pseudo-Riemannian manifold. A diffeomorphism $\varphi : M \to M$ is called an *isometry* if and only if $\varphi^* g = g$.

Most often it is sufficient to consider infinitesimal diffeomorphisms, i.e., one-parameter groups of diffeomorphisms generated by a vector field X. Also these carry their own name, given in the following definition.

Definition 31.15.2 (Killing vector field)**.** Let (M, g) be a pseudo-Riemannian manifold. A vector field $X \in \mathrm{Vect}(M)$ on M is called a *Killing vector field* if and only if $\mathcal{L}_X g = 0$.

Using the fact that one can rewrite the Lie derivative with the help of the Levi-Civita covariant derivative of the metric, and that this derivative vanishes by definition, one can also find another expression for the condition $\mathcal{L}_X g = 0$. The most common coordinate-free form is the following.

Theorem 31.15.1 (Killing equation)**.** *A vector field $X \in \mathrm{Vect}(M)$ on a manifold M with pseudo-Riemannian metric g is a Killing vector field if and only if*

$$g(\nabla_Y X, Z) + g(\nabla_Z X, Y) = 0 \tag{31.15.1}$$

for all vector fields $Y, Z \in \mathrm{Vect}(M)$.

Proof. We use the Koszul formula for the Levi-Civita connection to obtain

$$\begin{aligned} g(\nabla_Y X, Z) + g(\nabla_Z X, Y) &= \frac{1}{2}\Big\{Y(g(X, Z)) + X(g(Y, Z)) - Z(g(Y, X)) \\ &\quad + g([Y, X], Z) - g([Y, Z], X) - g([X, Z], Y) \\ &\quad + Z(g(X, Y)) + X(g(Z, Y)) - Y(g(Z, X)) \\ &\quad + g([Z, X], Y) - g([Z, Y], X) - g([X, Y], Z)\Big\} \\ &= X(g(Z, Y)) - g([X, Y], Z) - g([X, Z], Y) \\ &= (\mathcal{L}_X g)(Y, Z)\,, \end{aligned} \tag{31.15.2}$$

which vanishes for all $Y, Z \in \mathrm{Vect}(M)$ if and only if X is a Killing vector field. ■

In the theorem and its proof above the vector fields Y, Z are used as auxiliary quantities to contract the free indices of the tensor field $\mathcal{L}_X g$, so that the contracted expression vanishes for all Y, Z if and only if the tensor field itself vanishes. To derive the equivalent condition in coordinates, they can be omitted, and we can calculate the Lie derivative of the metric as follows:

$$\begin{aligned} (\mathcal{L}_X g)_{\mu\nu} &= X^\rho \partial_\rho g_{\mu\nu} + \partial_\mu X^\rho g_{\rho\nu} + \partial_\nu X^\rho g_{\mu\rho} \\ &= X^\rho(\nabla_\rho g_{\mu\nu} + \Gamma^\sigma{}_{\mu\rho} g_{\sigma\nu} + \Gamma^\sigma{}_{\nu\rho} g_{\mu\sigma}) + (\nabla_\mu X^\rho - \Gamma^\rho{}_{\sigma\mu} X^\sigma) g_{\rho\nu} + (\nabla_\nu X^\rho - \Gamma^\rho{}_{\sigma\nu} X^\sigma) g_{\mu\rho} \\ &= X^\rho[\nabla_\rho g_{\mu\nu} + (\Gamma^\sigma{}_{\mu\rho} - \Gamma^\sigma{}_{\rho\mu}) g_{\sigma\nu} + (\Gamma^\sigma{}_{\nu\rho} - \Gamma^\sigma{}_{\rho\nu}) g_{\mu\sigma}] + \nabla_\mu X^\rho g_{\rho\nu} + \nabla_\nu X^\rho g_{\mu\rho} \\ &= \nabla_\mu X_\nu + \nabla_\nu X_\mu\,. \end{aligned} \tag{31.15.3}$$

Here we have used the fact that the covariant derivative of the metric and the torsion of the Levi-Civita connection vanish in order to omit these terms, and then used the former again to contract the metric under the covariant derivative, hence lowering the indices of the vector field X. This is the most commonly encountered coordinate form of the Killing equation.

From the theorem above we have seen that there is a close relationship between the Lie derivative of the metric and its Levi-Civita connection. Recalling from section 30.9 that we have also defined the Lie derivative of an affine connection, one may expect that also the latter is defined from that of the metric. We can show this as follows.

Theorem 31.15.2. *The Lie derivative of the Levi-Civita connection is given by*

$$(\mathcal{L}_V\nabla)(X,Y) = \blacktriangleright \dots \blacktriangleleft \tag{31.15.4}$$

Proof.

$$\begin{aligned}
g((\mathcal{L}_V\nabla)(X,Y),Z) &= g([V,\nabla_Y X] - \nabla_{[V,Y]}X - \nabla_Y[V,X],Z) \\
&= Vg(\nabla_Y X,Z) - (\mathcal{L}_V g)(\nabla_Y X,Z) - g(\nabla_Y X,[V,Z]) - g(\nabla_{[V,Y]}X,Z) - g(\nabla_Y[V,X],Z) \\
&= \frac{1}{2}V\Big\{Y(g(X,Z)) + X(g(Y,Z)) - Z(g(Y,X)) \\
&\quad + g([Y,X],Z) - g([Y,Z],X) - g([X,Z],Y)\Big\} - (\mathcal{L}_V g)(\nabla_Y X,Z) \\
&\quad - \frac{1}{2}\Big\{Y(g(X,[V,Z])) + X(g(Y,[V,Z])) - [V,Z](g(Y,X)) \\
&\quad + g([Y,X],[V,Z]) - g([Y,[V,Z]],X) - g([X,[V,Z]],Y)\Big\} \\
&\quad - \frac{1}{2}\Big\{[V,Y](g(X,Z)) + X(g([V,Y],Z)) - Z(g([V,Y],X)) \\
&\quad + g([[V,Y],X],Z) - g([[V,Y],Z],X) - g([X,Z],[V,Y])\Big\} \\
&\quad - \frac{1}{2}\Big\{Y(g([V,X],Z)) + [V,X](g(Y,Z)) - Z(g(Y,[V,X])) \\
&\quad + g([Y,[V,X]],Z) - g([Y,Z],[V,X]) - g([[V,X],Z],Y)\Big\} \\
&=
\end{aligned} \tag{31.15.5}$$

∎

Chapter 32

Metric-affine geometry

32.1 Nonmetricity

Definition 32.1.1 (Nonmetricity)**.** For a metric-affine

32.2 Connection decomposition

Theorem 32.2.1. *Let M be a manifold with a pseudo-Riemannian metric g. Then any affine connection ∇ on M is uniquely determined by its torsion T and nonmetricity Q and vice versa.*

Proof. We proceed in analogy to the proof of theorem 31.10.1. Let $X, Y, Z \in \mathrm{Vect}(M)$ be vector fields. From the definition 32.1.1 of the nonmetricity follows

$$Q(Z, X, Y) = (\nabla_Z g)(X, Y) = Z(g(X, Y)) - g(\nabla_Z X, Y) - g(X, \nabla_Z Y)\,. \tag{32.2.1}$$

Consider the linear combination

$$\begin{aligned} Q(Z, X, Y) - Q(X, Z, Y) - Q(Y, X, Z) &= Z(g(X, Y)) + g(\nabla_X Y + \nabla_Y X, Z) \\ &\quad - X(g(Y, Z)) + g(\nabla_X Z - \nabla_Z X, Y) \\ &\quad - Y(g(X, Z)) + g(\nabla_Y Z - \nabla_Z Y, X) \\ &= Z(g(X, Y)) - X(g(Y, Z)) - Y(g(X, Z)) \\ &\quad - g([X, Y], Z) + g([X, Z], Y) + g([Y, Z], X) \\ &\quad - g(T(X, Y), Z) + g(T(X, Z), Y) + g(T(Y, Z), X) \\ &\quad + 2g(\nabla_X Y, Z)\,, \end{aligned} \tag{32.2.2}$$

where we used the fact that g is symmetric and bilinear, and inserted the definition of the torsion. Solving this yields the formula

$$
\begin{aligned}
g(\nabla_X Y, Z) &= \frac{1}{2}\big\{X(g(Y,Z)) + Y(g(X,Z)) - Z(g(X,Y)) \\
&\quad + g([X,Y],Z) - g([X,Z],Y) - g([Y,Z],X) \\
&\quad + g(T(X,Y),Z) - g(T(X,Z),Y) - g(T(Y,Z),X) \\
&\quad + Q(Z,X,Y) - Q(X,Z,Y) - Q(Y,X,Z)\big\} \\
&= \frac{1}{2}\big\{g(T(X,Y),Z) - g(T(X,Z),Y) - g(T(Y,Z),X) \\
&\quad + Q(Z,X,Y) - Q(X,Z,Y) - Q(Y,X,Z)\big\} + g(\mathring{\nabla}_X Y, Z)\,.
\end{aligned}
\tag{32.2.3}
$$

Since Z is arbitrary and g is non-degenerate, this uniquely defines $\nabla_X Y$ for all vector fields $X, Y \in \mathrm{Vect}(M)$. ■

Chapter 33

Weyl geometry

33.1 Orthogonal frame bundle

Chapter 34

Weitzenböck geometry

Chapter 35

Symplectic geometry

35.1 Symplectic forms

In chapter 31 we have introduced the pseudo-Riemannian metric and studied it in subsequent chapters. It constitutes a non-degenerate, symmetric bilinear form on each tangent space, and hence induces a bijective mapping between tangent and cotangent spaces. We now come to a structure which has a similar property, but instead of being symmetric, it is *antisymmetric*, and which is crucial for the Hamilton theory we discuss in chapter 53. Here we mostly follow the treatment and conventions of [Ber01]. We start our discussion with the following definition.

Definition 35.1.1 (Almost symplectic form)**.** An *almost symplectic form* on a manifold M is a non-degenerate differential 2-form ω, i.e., at each point $p \in M$, ω defines a non-degenerate, antisymmetric, bilinear form ω_p on T_pM.

For a pseudo-Riemannian metric, it follows from Sylvester's law of inertia that there are pseudo-Riemannian metrics of arbitrary signature (k, l) on manifolds of dimension $k + l$. For almost symplectic forms, a similar statement exists, which comes from the properties of antisymmetric, non-degenerate bilinear forms. However, this statement is far more restrictive.

Theorem 35.1.1. *A manifold M endowed with an almost symplectic form ω is necessarily even-dimensional.*

Proof. This follows from the *Jacobi's theorem*, which states that an antisymmetric matrix A of odd dimension n satisfies

$$\det A = \det A^t = \det(-A) = (-1)^n \det A = -\det A \,, \tag{35.1.1}$$

and hence $\det A = 0$, so that it must be degenerate. ∎

Further, one can show that for every non-degenerate, antisymmetric, bilinear form there exists a basis in which it is represented by the matrix

$$\Omega = \begin{pmatrix} 0 & \mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{pmatrix} . \tag{35.1.2}$$

While various of the statements we discuss in the following sections hold for almost symplectic forms, there are some which require an additional condition, which can be regarded as an *integrability condition*. Imposing this condition yields the following definition.

Definition 35.1.2 (Symplectic form)**.** An almost symplectic form ω is called a *symplectic form* if and only if it is closed.

Since a symplectic form ω is closed, there exists locally (and if ω is exact also globally) an object which we define as follows.

Definition 35.1.3 (Symplectic potential)**.** A one-form $\theta \in \Omega^1(M)$, such that $\omega = \mathrm{d}\theta$ is a symplectic form, is called a *symplectic potential* of ω.

In the following, we will encounter several examples of symplectic geometries.

35.2 Symplectic structure on the cotangent bundle

An important example for a manifold which carries a *canonical* symplectic structure is the cotangent bundle, and we will study some of its properties in this section.

Definition 35.2.1 (Tautological one-form)**.** Let M be a manifold and $\bar{\tau} : T^*M \to M$ its cotangent bundle. For $\alpha \in T^*M$, define $\theta_\alpha \in T^*_\alpha T^*M$ by

$$\begin{array}{rcccc} \theta_\alpha & : & T_\alpha T^*M & \to & \mathbb{R} \\ & & \xi & \mapsto & \langle \bar{\tau}_*(\xi), \alpha \rangle \end{array} . \tag{35.2.1}$$

The *tautological one-form* $\theta \in \Omega^1(T^*M)$ is the map defined by

$$\begin{array}{rcccc} \theta & : & T^*M & \to & T^*T^*M \\ & & \alpha & \mapsto & \theta_\alpha \end{array} . \tag{35.2.2}$$

This construction can be understood as follows. For $\xi \in T_\alpha T^*M$, there exists the pushforward $\bar{\tau}_*(\xi) \in T_{\bar{\tau}(\alpha)}M$. Further, we have $\alpha \in T^*_{\bar{\tau}(\alpha)}M$. Via the canonical pairing $\langle \bullet, \bullet \rangle$ of the tangent and cotangent space, we thus have $\langle \bar{\tau}_*(\xi), \alpha \rangle \in \mathbb{R}$. This expression is linear in ξ, and so it defines a linear map $\theta_\alpha : T_\alpha T^*M \to \mathbb{R}$, hence an element $\theta_\alpha \in T^*_\alpha T^*M$. Doing this for every $\alpha \in T^*M$ yields a section of T^*T^*M, which is a one-form $\theta \in \Omega^1(T^*M)$. It has a few interesting properties, as we will show in the following.

Theorem 35.2.1. *The tautological one-form is homogeneous of order* 1*.*

Proof. Let $\chi : \mathbb{R} \times T^*M \to T^*M$ denote the dilatations, given in definition 19.9.1, and $\mathbf{c} \in \mathrm{Vect}(T^*M)$ their generating vector field, the Liouville vector field as in definition 19.9.2. For $\alpha \in T^*M$, $\xi \in T_\alpha T^*M$ and $\lambda \in \mathbb{R}$ we have

$$\begin{aligned} \langle \xi, (\chi_\lambda^*\theta)_\alpha \rangle &= \langle \chi_{\lambda *}(\xi), \theta_{\chi_\lambda(\alpha)} \rangle \\ &= \langle \bar{\tau}_*(\chi_{\lambda *}(\xi)), \chi_\lambda(\alpha) \rangle \\ &= \langle (\bar{\tau} \circ \chi_\lambda)_*(\xi), e^\lambda \alpha \rangle \\ &= e^\lambda \langle \bar{\tau}_*(\xi), \alpha \rangle \\ &= e^\lambda \langle \xi, \theta_\alpha \rangle \,, \end{aligned} \tag{35.2.3}$$

and so $\chi_\lambda^*\theta = e^\lambda\theta$. Hence, θ is 1-homogeneous. ■

Another useful property is the following.

Theorem 35.2.2. *For every one-form $\sigma \in \Omega^1(M)$ holds $\sigma^*(\theta) = \sigma$.*

Proof. First note that the left hand side indeed defines a one-form on M, since θ is a one-form on T^*M and $\sigma : M \to T^*M$ is a map relating these manifolds. Consider $x \in M$ and $v \in T_xM$. Then we have

$$\begin{aligned}\langle v, (\sigma^*\theta)_x\rangle &= \langle \sigma_*(v), \theta_{\sigma(x)}\rangle \\ &= \langle \bar{\tau}_*(\sigma_*(v)), \sigma(x)\rangle \\ &= \langle (\bar{\tau} \circ \sigma)_*(v), \sigma(x)\rangle \\ &= \langle v, \sigma(x)\rangle\,.\end{aligned} \tag{35.2.4}$$

Since this holds for all $x \in M$ and $v \in T_xM$, it follows that $\sigma^*(\theta) = \sigma$. ■

With this statement in place, one can now prove the following.

Theorem 35.2.3. *Let $\sigma \in \Omega^1(M)$ be a one-form, and define*

$$\begin{array}{cccc}\Sigma & : & T^*M & \to & T^*M \\ & & \alpha & \mapsto & \alpha + \sigma(\bar{\tau}(\alpha))\end{array}. \tag{35.2.5}$$

Then $\Sigma^*\theta = \theta + \bar{\tau}^*(\sigma)$.

Proof. Let $\alpha \in T^*M$ and $\xi \in T_\alpha T^*M$. Then we have

$$\begin{aligned}\langle \xi, (\Sigma^*\theta)_\alpha\rangle &= \langle \Sigma_*(\xi), \theta_{\Sigma(\alpha)}\rangle \\ &= \langle \bar{\tau}_*(\Sigma_*(\xi)), \Sigma(\alpha)\rangle \\ &= \langle (\bar{\tau} \circ \Sigma)_*(\xi), \alpha + \sigma(\bar{\tau}(\alpha))\rangle \\ &= \langle \bar{\tau}_*(\xi), \alpha + \sigma(\bar{\tau}(\alpha))\rangle \\ &= \langle \bar{\tau}_*(\xi), \alpha\rangle + \langle \bar{\tau}_*(\xi), \sigma(\bar{\tau}(\alpha))\rangle \\ &= \langle \xi, \theta_\alpha\rangle + \langle \xi, (\bar{\tau}^*\sigma)_\alpha\rangle \\ &= \langle \xi, (\theta + \bar{\tau}^*\sigma)_\alpha\rangle\,.\end{aligned} \tag{35.2.6}$$

Since this holds for all $\alpha \in T^*M$ and $\xi \in T_\alpha T^*M$, the proposition follows. ■

One now easily shows the following.

Theorem 35.2.4. *The tautological one-form θ is a symplectic potential for the symplectic form $\omega = \mathrm{d}\theta$*

Proof. Obviously, ω is closed, $\mathrm{d}\omega = \mathrm{d}\mathrm{d}\theta = 0$. It remains to show that ω is non-degenerate, which will be done in several steps. Since θ is 1-homogeneous as of theorem 35.2.1, one has

$$\theta = \mathcal{L}_{\mathbf{c}}\theta = \iota_{\mathbf{c}}\mathrm{d}\theta + \mathrm{d}\iota_{\mathbf{c}}\theta\,. \tag{35.2.7}$$

Here the second term vanishes, since $\mathbf{c}$ is vertical, and hence $\bar{\tau}_* \circ \mathbf{c} = 0$, and so we have

$$\theta = \iota_{\mathbf{c}}\omega\,. \tag{35.2.8}$$

Now consider a one-form $\sigma \in \Omega^1(M)$. ▶. . . ◀ ■

We finally derive coordinate expressions for the tautological one-form and its symplectic form. It is conventional to denote the coordinates on M by (q^a), and to introduce coordinates (q^a, p_a) on T^*M to denote covectors as $p_a\,\mathrm{d}q^a \in T^*_qM$. Given $\alpha = (q^a, p_a) \in T^*M$, we can express $\xi \in T_\alpha T^*M$ as

$$\xi = \xi^a \frac{\partial}{\partial q^a} + \bar{\xi}_a \frac{\partial}{\partial p_a} \in T_\alpha T^*M \tag{35.2.9}$$

in the coordinate basis of TT^*M. Its pushforward is given by

$$\bar{\tau}_*(\xi) = \xi^a \frac{\partial}{\partial q^a} \in T_{\bar{\tau}(\alpha)}M \tag{35.2.10}$$

in the coordinate basis of TM. The canonical pairing then yields

$$\langle \bar{\tau}_*(\xi), \alpha \rangle = p_a \xi^a \,. \tag{35.2.11}$$

Hence, the tautological one-form is given by

$$\theta = p_a\,\mathrm{d}q^a \in \Omega^1(T^*M)\,. \tag{35.2.12}$$

Its exterior derivative yields the symplectic form

$$\omega = \mathrm{d}\theta = \mathrm{d}p_a \wedge \mathrm{d}q^a\,. \tag{35.2.13}$$

35.3 Hamiltonian vector field

In section 31.3 we have introduced the musical isomorphisms, which relate the tangent and cotangent bundles of a manifold equipped with a pseudo-Riemannian metric. In particular, we used the map $\sharp$ to obtain a vector field from a one-form. We can perform a similar operation using a symplectic form, and combine this with the total differential acting on functions. This yields the following definition.

Definition 35.3.1 (Hamiltonian vector field)**.** Let M be a manifold equipped with an almost symplectic form ω and $f \in C^\infty(M, \mathbb{R})$. The *Hamiltonian vector field* of f is the unique vector field $X_f \in \mathrm{Vect}(M)$ such that $\iota_{X_f}\omega = \mathrm{d}f$.

Note that there are different sign conventions used in the literature. Here we follow [Ber01, def. 3.2]. We also remark that for the definition above we only need ω to be non-degenerate, but not necessarily closed. This is also enough to show the following properties.

Theorem 35.3.1. *The assignment $X_\bullet : C^\infty(M, \mathbb{R}) \to \mathrm{Vect}(M)$ is a linear function which satisfies the Leibniz rule,*

$$X_{fg} = fX_g + gX_f\,, \tag{35.3.1}$$

for $f, g \in C^\infty(M, \mathbb{R})$.

Proof. We make use of the properties 8.4.1 of the total differential. Linearity follows from

$$\iota_{X_{\mu f+\nu g}}\omega = \mathrm{d}(\mu f + \nu g) = \mu\,\mathrm{d}f + \nu\,\mathrm{d}g = \mu\iota_{X_f}\omega + \nu\iota_{X_g}\omega = \iota_{\mu X_f + \nu X_g}\omega \tag{35.3.2}$$

for $f, g \in C^\infty(M, \mathbb{R})$ and $\mu, \nu \in \mathbb{R}$. Similarly, one shows the Leibniz rule

$$\iota_{X_{fg}}\omega = \mathrm{d}(fg) = f\,\mathrm{d}g + g\,\mathrm{d}f = f\iota_{X_g}\omega + g\iota_{X_f}\omega = \iota_{fX_g+gX_f}\omega \tag{35.3.3}$$

for $f, g \in C^\infty(M, \mathbb{R})$. ■

Also the following property holds for almost symplectic forms.

Theorem 35.3.2. *Every function $f \in C^\infty(M, \mathbb{R})$ is constant along its Hamiltonian vector field, $X_f f = 0$.*

Proof. This follows immediately from

$$X_f f = \iota_{X_f} \mathrm{d}f = \iota_{X_f} \iota_{X_f} \omega = 0\,, \tag{35.3.4}$$

due to the antisymmetry of the interior product. ■

Given an arbitrary vector field, one may now pose the question whether this vector field happens to be the Hamiltonian vector field X_f of some function f. In other words, we are looking for a possibility to characterize all Hamiltonian vector fields. Looking at the definition 35.3.1, we see that the existence of f is equivalent to stating that $\iota_{X_f}\omega$ is exact. This leads us to the following definition.

Definition 35.3.2 (Hamiltonian vector fields)**.** Let M be a manifold equipped with an almost symplectic form ω. A vector field $Y \in \mathrm{Vect}(M)$ is called:

1. *Hamiltonian*, $Y \in \mathrm{Ham}(M, \omega)$, if $\iota_Y \omega$ is exact,
2. *locally Hamiltonian*, $Y \in \mathrm{Ham}_0(M, \omega)$, if $\iota_Y \omega$ is closed.

Obviously, every Hamiltonian vector field is also locally Hamiltonian, and hence $\mathrm{Ham}(M, \omega) \subseteq \mathrm{Ham}_0(M, \omega)$. Further, we see that $\mathrm{Ham}(M, \omega)$ and $\mathrm{Ham}_0(M, \omega)$ are vector spaces, and $X_f \in \mathrm{Ham}(M, \omega)$ for all $f \in C^\infty(M, \mathbb{R})$. We can find a few more interesting relations, which we will now discuss, starting with the following.

Theorem 35.3.3. *Let M be a manifold equipped with an almost symplectic form ω, and $n \in \mathbb{N}$ be the number of connected components of M. There exists an exact sequence*

$$0 \longrightarrow \mathbb{R}^n \hookrightarrow C^\infty(M, \mathbb{R}) \xrightarrow{X_\bullet} \mathrm{Ham}(M, \omega) \longrightarrow 0 \tag{35.3.5}$$

of vector spaces, where $\mathbb{R}^n \hookrightarrow C^\infty(M, \mathbb{R})$ is the canonical inclusion which assigns to $c = (c_1, \ldots, c_n) \in \mathbb{R}^n$ the constant function $\bullet \mapsto c_i$ on the i'th connected component of M.

Proof. The map $X_\bullet$ is surjective, since, by definition, a vector field is Hamiltonian if and only if it is the image of a function $f \in C^\infty(M, \mathbb{R})$ under $X_\bullet$. The kernel is given by those functions for which $\mathrm{d}f = 0$, since ω is non-degenerate. These are the functions which are constant on the connected components of M. ■

Recall that the vector fields $\mathrm{Vect}(M)$ on a manifold M, equipped with the commutator, form a Lie algebra. One may wonder whether this holds also for the subspace of (locally) Hamiltonian vector fields. It turns out that this is the case only for *symplectic* (and hence closed) forms ω. We show this in several steps. First, we give another characteristic property of locally Hamiltonian vector fields.

Theorem 35.3.4. *Let M be a manifold equipped with a symplectic form ω and $Y \in \mathrm{Vect}(M)$. Then Y is locally Hamiltonian if and only if ω is constant along Y, i.e., $\mathcal{L}_Y \omega = 0$.*

Proof. By direct calculation we have, if and only if $\iota_Y \omega$ is closed,

$$\mathcal{L}_Y \omega = \mathrm{d}\iota_Y \omega + \iota_Y \mathrm{d}\omega = 0\,, \tag{35.3.6}$$

where the second term vanishes due to $\mathrm{d}\omega = 0$. ■

Now we can state the property for locally Hamiltonian vector fields.

Theorem 35.3.5. *Let M be a manifold equipped with a symplectic form ω. Then the locally Hamiltonian vector fields* $\mathrm{Ham}_0(M, \omega)$ *form a Lie subalgebra of* $\mathrm{Vect}(M)$.

Proof. Let $Y, Z \in \mathrm{Ham}_0(M, \omega)$. Then we have

$$\mathcal{L}_{[Y,Z]}\omega = \mathcal{L}_Y \mathcal{L}_Z \omega - \mathcal{L}_Z \mathcal{L}_Y \omega = 0\,, \tag{35.3.7}$$

and so $[Y, Z] \in \mathrm{Ham}_0(M, \omega)$. ■

This holds also for Hamiltonian vector fields.

Theorem 35.3.6. *Let M be a manifold equipped with a symplectic form ω. Then the Hamiltonian vector fields* $\mathrm{Ham}(M, \omega)$ *form a Lie subalgebra of* $\mathrm{Vect}(M)$.

Proof. We defer this proof to the proof of theorem 35.4.3, where we give an explicit formula for the commutator of two Hamilton vector fields, and show that it is again Hamiltonian. ■

35.4 Poisson bracket

A symplectic form ω equips a manifold with a number of other interesting objects, which are relevant for physical application. A particularly important example will be discussed in this section. We start with the following definition.

Definition 35.4.1 (Poisson structure)**.** Let M be a manifold. A *Poisson structure* is a $\mathbb{R}$-bilinear mapping $\{\bullet, \bullet\} : C^\infty(M, \mathbb{R}) \times C^\infty(M, \mathbb{R}) \to C^\infty(M, \mathbb{R})$ which satisfies for all $f, g, h \in C^\infty(M, \mathbb{R})$:

1. Antisymmetry: $\{f, g\} = -\{g, f\}$,
2. Jacobi identity: $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$,
3. Leibniz rule: $\{fg, h\} = f\{g, h\} + \{f, h\}g$.

The first two properties state that the Poisson bracket equips the set $C^\infty(M, \mathbb{R})$ of real functions on M with a Lie algebra structure. The Leibniz rule states that for $f \in C^\infty(M, \mathbb{R})$ and $x \in M$, the Poisson bracket defines a derivation

$$\begin{array}{rcccc} \{\bullet, f\}_x & : & C^\infty(M, \mathbb{R}) & \to & \mathbb{R} \\ & & g & \mapsto & \{g, f\}(x) \end{array} \tag{35.4.1}$$

at x, since

$$\{gh, f\}_x = \{gh, f\}(x) = g(x)\{h, f\}(x) + \{g, f\}(x)h(x) = g(x)\{h, f\}_x + \{g, f\}_x h(x)\,, \tag{35.4.2}$$

and hence a tangent vector $\{\bullet, f\}_x \in T_xM$. Doing this for every $x \in M$, we thus obtain a vector field. Note that this is reminiscent of the Hamiltonian vector field from definition 35.3.1. To see that these are indeed related, we first show how to obtain a Poisson bracket from a symplectic form.

Theorem 35.4.1. *Let M be a manifold equipped with a symplectic form ω. Then*

$$\begin{array}{rcccc} \{\bullet, \bullet\} & : & C^\infty(M, \mathbb{R}) \times C^\infty(M, \mathbb{R}) & \to & C^\infty(M, \mathbb{R}) \\ & & (f, g) & \mapsto & \{f, g\} = \omega(X_f, X_g) \end{array} \tag{35.4.3}$$

defines a Poisson structure.

Proof. We have to show that $\{\bullet, \bullet\}$ satisfies the properties given in definition 35.4.1. Note first that $\{\bullet, \bullet\}$ is bilinear, which follows from the linearity 35.3.1 of the Hamiltonian vector fields and the bilinear form ω. We further show:

1. Antisymmetry:
$$\{f, g\} = \omega(X_f, X_g) = -\omega(X_g, X_f) = -\{g, f\}, \tag{35.4.4}$$
due to the antisymmetry of ω.

2. By explicit calculation, using theorem 9.4.3, we have
$$\begin{aligned}\{h, \{f, g\}\} &= \omega(X_h, \omega(X_f, X_g))\\ &= -\iota_{X_h} \mathrm{d}\omega(X_f, X_g)\\ &= -\mathcal{L}_{X_h}\omega(X_f, X_g)\\ &= -(\mathcal{L}_{X_h}\omega)(X_f, X_g) - \omega(\mathcal{L}_{X_h}X_f, X_g) - \omega(X_f, \mathcal{L}_{X_h}X_g)\\ &= -(\iota_{X_h}\mathrm{d}\omega)(X_f, X_g) - (\mathrm{d}\iota_{X_h}\omega)(X_f, X_g) - \omega([X_h, X_f], X_g) - \omega(X_f, [X_h, X_g])\\ &= -\iota_{X_g}\iota_{X_f}\iota_{X_h}\mathrm{d}\omega - \iota_{X_g}\iota_{X_f}\mathrm{d}\iota_{X_h}\omega - \iota_{X_g}\mathrm{d}\iota_{X_h}\iota_{X_f}\omega + \iota_{X_f}\mathrm{d}\iota_{X_h}\iota_{X_g}\omega\\ &\quad - \iota_{X_g}\iota_{X_h}\mathrm{d}\iota_{X_f}\omega + \iota_{X_f}\iota_{X_h}\mathrm{d}\iota_{X_g}\omega + 2\iota_{X_g}\iota_{X_f}\mathrm{d}\iota_{X_h}\omega + 2\iota_{X_g}\iota_{X_f}\iota_{X_h}\mathrm{d}\omega\\ &= -\iota_{X_g}\mathrm{d}\omega(X_f, X_h) + \iota_{X_f}\mathrm{d}\omega(X_g, X_h)\\ &\quad - \iota_{X_g}\iota_{X_h}\mathrm{dd}f + \iota_{X_f}\iota_{X_h}\mathrm{dd}g + \iota_{X_g}\iota_{X_f}\mathrm{dd}h + \mathrm{d}\omega(X_h, X_f, X_g)\\ &= \omega(X_g, \omega(X_f, X_h)) - \omega(X_f, \omega(X_g, X_h))\\ &= -\{g, \{h, f\}\} - \{f, \{g, h\}\},\end{aligned} \tag{35.4.5}$$
where we used $\mathrm{d}\omega = 0$.

3. From the properties 35.3.1 of the Hamilton vector fields follows
$$\begin{aligned}\{fg, h\} &= \omega(X_{fg}, X_h)\\ &= \omega(fX_g + gX_f, X_h)\\ &= f\omega(X_g, X_h) + g\omega(X_f, X_h)\\ &= f\{g, h\} + \{f, h\}g\,. \quad ■\end{aligned} \tag{35.4.6}$$

Note in particular that the validity of the Jacobi identity depends on the condition $\mathrm{d}\omega = 0$ that ω is closed, i.e., it does not hold if ω is only an almost symplectic form, but it must indeed be a symplectic form. With this knowledge, we can now return to our observation that, due to the Leibniz rule, $\{\bullet, f\}$ constitutes a vector field, by its action on functions. We now show that this vector field is already familiar.

Theorem 35.4.2. *Let M be a manifold equipped with a symplectic form ω and $\{,\}$ the induced Poisson structure. For $f \in C^\infty(M, \mathbb{R})$, the Hamilton vector field $X_f \in \mathrm{Vect}(M)$ of f is the unique vector field such that $X_f g = \{g, f\}$ for all $g \in C^\infty(M, \mathbb{R})$.*

Proof. By definition of the Hamiltonian vector field and the Poisson bracket we have
$$X_f g = \iota_{X_f}\mathrm{d}g = \iota_{X_f}\iota_{X_g}\omega = \omega(X_g, X_f) = \{g, f\}\,. \tag{35.4.7}$$

Further, X_f is unique, since g is arbitrary and a vector field is uniquely defined by its actions on functions g. ■

Observe further that the Poisson bracket, due to being bilinear and antisymmetric and satisfying the Jacobi identity, equips $C^\infty(M, \mathbb{R})$ with the structure of a Lie algebra. One may therefore ask whether $X_\bullet : C^\infty(M, \mathbb{R}) \to \mathrm{Vect}(M)$ preserves this Lie algebra structure, since also vector fields form a Lie algebra. We now show that this is the case.

Theorem 35.4.3. *The Poisson bracket, the commutator of vector fields and the Hamiltonian vector fields are related by*

$$X_{\{f,g\}} = -[X_f, X_g] \tag{35.4.8}$$

for $f, g \in C^\infty(M, \mathbb{R})$ *.*

Proof. We show two different proofs. Using the Jacobi identity of the Poisson bracket we have

$$\begin{aligned} X_{\{f,g\}}h &= \{h, \{f, g\}\} \\ &= -\{f, \{g, h\}\} - \{g, \{h, f\}\} \\ &= -X_f X_g h + X_g X_f h \\ &= -[X_f, X_g]h \end{aligned} \tag{35.4.9}$$

for arbitrary $f, g, h \in C^\infty(M, \mathbb{R})$, and $X_{\{f,g\}}$ is uniquely defined by its action, since h is arbitrary.

Alternatively, using theorem 9.4.3 we have

$$\begin{aligned} \iota_{X_{\{f,g\}}}\omega &= \mathrm{d}\{f, g\} \\ &= \mathrm{d}\iota_{X_g}\iota_{X_f}\omega \\ &= -\iota_{[X_f,X_g]}\omega + \iota_{X_f}\mathrm{d}\iota_{X_g}\omega - \iota_{X_g}\mathrm{d}\iota_{X_f}\omega - \iota_{X_g}\iota_{X_f}\mathrm{d}\omega \\ &= -\iota_{[X_f,X_g]}\omega + \iota_{X_f}\mathrm{dd}g - \iota_{X_g}\mathrm{dd}f \\ &= -\iota_{[X_f,X_g]}\omega\,, \end{aligned} \tag{35.4.10}$$

where we used $\mathrm{d}\omega = 0$. We remark that this latter property also entered the proof of the Jacobi identity, hence this holds only if ω is a symplectic form. ■

Note that the negative sign in the relation implies that not $X_\bullet$, but $-X_\bullet : C^\infty(M, \mathbb{R}) \to \mathrm{Vect}(M)$ is a Lie algebra homomorphism. This fact depends on the sign convention chosen when defining $X_\bullet$, and where we follow [Ber01]. If one uses the opposite sign convention, then the sign in (35.4.8) changes and $X_\bullet$ becomes a Lie algebra homomorphism instead.

35.5 Moment map

35.6 Symplectic frame bundle

Definition 35.6.1 (Symplectic frame bundle). Let M be a manifold of dimension $2n$ equipped with an almost symplectic form ω and Ω the canonical antisymmetric bilinear form on $\mathbb{R}^{2n}$. A *symplectic frame* at $x \in M$ is a bijective linear function $p : \mathbb{R}^{2n} \to T_xM$ such that $\Omega = p^*\omega$. The set of all symplectic frames constitutes the *symplectic frame bundle* $\mathrm{Sp}(M, \omega)$ with projection mapping $p : \mathbb{R}^{2n} \to T_xM$ to $x \in M$.

Theorem 35.6.1. *The symplectic frame bundle* $\mathrm{Sp}(M, \omega)$ *over a manifold* M *of dimension* $2n$ *with almost symplectic form* ω *is a principal fiber bundle with structure group* $\mathrm{Sp}(2n, \mathbb{R})$*, where the right action is given by* $p \cdot A = p \circ A$ *for* $p \in \mathrm{Sp}(M, \omega)$ *and* $A \in \mathrm{Sp}(2n, \mathbb{R})$*.*

Proof. ▶. . . ◀ ■

Theorem 35.6.2. *The symplectic frame bundle together with the canonical inclusion* $\mathrm{Sp}(M, \omega) \hookrightarrow FM$ *is a* $\mathrm{Sp}(2n, \mathbb{R})$*-reduction of the general linear frame bundle* FM*.*

Proof. ▶ . . . ◀ ∎

Theorem 35.6.3. *The bundle of almost symplectic forms on a manifold M of dimension $2n$ is isomorphic to the associated bundle $FM \times_\rho G/H$ with $G = \mathrm{GL}(2n, \mathbb{R})$, $H = \mathrm{Sp}(2n, \mathbb{R})$ and ρ the canonical left action of G on the coset space G/H.*

Proof. ▶ . . . ◀ ∎

35.7 Symplectomorphisms

Chapter 36

Contact geometry

36.1 Contact forms

Definition 36.1.1 (Contact form)**.** A *contact form* on a manifold M of odd dimension $2k+1$ is a differential one-form α, such that

$$\alpha \wedge \underbrace{\mathrm{d}\alpha \wedge \ldots \wedge \mathrm{d}\alpha}_{k \text{ times}} \tag{36.1.1}$$

is nowhere vanishing.

36.2 Reeb vector field

Chapter 37

Non-linear connections in the tangent bundle

37.1 Distributions in the double tangent bundle

Since the tangent bundle is a vector bundle, one most often considers linear connections on the tangent bundle, as defined for a general vector bundle in chapter 28, and discussed in the particular case of the tangent bundle in section 30. Here we relax this condition and consider general, non-linear connections on the tangent bundle, and different possibilities to specify these connections. These are most conveniently defined via horizontal distributions, as a special case of definition 26.1.1, as follows.

Definition 37.1.1 (Non-linear tangent bundle connection)**.** A *non-linear connection* on the tangent bundle TM of a manifold M is a horizontal distribution HTM over TM.

As discussed in section 26.1, the choice of a horizontal distribution allows a unique decomposition of a tangent vector $\psi \in TTM$, which lies in the tangent bundle over the total space TM of the bundle $\tau : TM \to M$, into horizontal and vertical parts. In the context of non-linear connections of the tangent bundle, one makes frequent use of these projectors, and it is most convenient to regard them as vector bundle endomorphisms on TTM, or equivalently as $(1,1)$ tensor fields on TM, as we did with the tangent structure in section 29.2. This leads us to the following definitions.

Definition 37.1.2 (Horizontal and vertical projectors)**.** For a horizontal distribution HTM on a manifold M, the *horizontal and vertical projectors* are the vector bundle homomorphisms $\mathbf{h} : TTM \to TTM$ and $\mathbf{v} : TTM \to TTM$ from $\varpi : TTM \to TM$ to itself, covering the identity on TM, which satisfy:

1. $\ker \mathbf{h} = \operatorname{im} \mathbf{v} = VTM$,
2. $\operatorname{im} \mathbf{h} = \ker \mathbf{v} = HTM$,
3. $\mathbf{h} + \mathbf{v} = \mathrm{id}_{TTM}$,
4. $\mathbf{h} \circ \mathbf{h} = \mathbf{h}$,

5. $\mathbf{v} \circ \mathbf{v} = \mathbf{v}$,

6. $\mathbf{h} \circ \mathbf{v} = \mathbf{v} \circ \mathbf{h} = 0$.

We have also discussed in sections 26.1 and 26.2 that a convenient coordinate expression for the projections onto the horizontal and vertical subbundles can be obtained by introducing a particular basis of the tangent bundle over the total space, which respects the split into horizontal and vertical parts. In the case of non-linear connections on the tangent bundle, this basis has a particular name, and we define it as follows.

Definition 37.1.3 (Berwald basis)**.** Let M be a manifold equipped with a non-linear connection. The *Berwald basis* of the tangent bundle is given by

$$(\delta_a = \partial_a - N^b{}_a \bar{\partial}_b, \bar{\partial}_a)\,, \tag{37.1.1}$$

while that of the cotangent bundle reads

$$(\mathrm{d}x^a, \delta\bar{x}^a = \mathrm{d}\bar{x}^a + N^a{}_b \mathrm{d}x^b)\,, \tag{37.1.2}$$

where (δ_a) is a basis of the horizontal tangent bundle HTM, $(\bar{\partial}_a)$ is a basis of the vertical tangent bundle VTM, $(\mathrm{d}x^a)$ is a basis of the horizontal cotangent bundle H^*TM and $(\delta\bar{x}^a)$ is a basis of the vertical cotangent bundle V^*TM, and $N^a{}_b = N^a{}_b(x, \bar{x})$ are the coefficients of the connection.

The connection coefficients, which are usually denoted $N^a{}_b$ for a non-linear tangent bundle connection, are simply the coefficients characterizing the connection form introduced in section 26.2. Indeed, writing the horizontal and vertical projectors in the Berwald basis, where they take the convenient form

$$\mathbf{h} = \delta_a \otimes \mathrm{d}x^a\,, \quad \mathbf{v} = \bar{\partial}_a \otimes \delta\bar{x}^a\,, \tag{37.1.3}$$

we see that the vertical projector is nothing but the connection form, for which we found the same coordinate expression (26.2.8).

We have discussed in chapter 26 that there are various different possibilities to specify connections on general fiber bundles. For the special case of the tangent bundle, there exist various additional and similarly useful possibilities, since it is canonically equipped with further geometric objects, in particular the tangent and cotangent structures. Given a non-linear connection, these turn out to obey a helpful set of rules, which we write as follows.

Theorem 37.1.1. *The horizontal and vertical projectors and the tangent structure satisfy the relations*

$$\mathbf{v} \circ J = J \circ \mathbf{h} = J\,, \quad \mathbf{h} \circ J = J \circ \mathbf{v} = 0\,. \tag{37.1.4}$$

Proof. Recall from theorem 29.2.3 that J vanishes on vertical vectors, which immediately implies $J \circ \mathbf{v} = 0$. From $\mathbf{v} + \mathbf{h} = \mathrm{id}_{TTM}$ then follows

$$J \circ \mathbf{h} = J \circ (\mathrm{id}_{TTM} - \mathbf{v}) = J - 0 = J\,. \tag{37.1.5}$$

Similarly, recall from theorem 29.2.2 that the image of J is vertical. From the fact that $\mathbf{h}$ vanishes on vertical vectors, while $\mathbf{v}$ restricts to the identity on VTM, then follows $\mathbf{h} \circ J = 0$ and $\mathbf{v} \circ J = J$. ■

From the previous statement we find another helpful relation.

Theorem 37.1.2. *The restriction of the tangent structure to the horizontal tangent bundle is a vector bundle isomorphism* $J|_{HTM} : HTM \to VTM$.

Proof. ►...◄ ∎

37.2 Characterizing tensors

Having shown that the horizontal and vertical subbundles are isomorphic, one naturally arrives at the question for the inverse isomorphism. We can construct such an isomorphism by demanding certain properties, which guarantee that it becomes the inverse of J, if it is properly restricted to VTM. These are given in the following definition.

Definition 37.2.1 (Adjoint structure). An *adjoint structure* on a manifold M is a vector bundle homomorphism $\Theta : TTM \to TTM$ from $\varpi : TTM \to TM$ to itself, covering the identity on TM, which satisfies $\Theta \circ \Theta = 0$ and $\Theta \circ J + J \circ \Theta = \mathrm{id}_{TTM}$.

The tangent and adjoint structures have in common that they are nilpotent, since both square to zero. However, there is also a fundamental difference. While the tangent structure, alongside with its image and kernel VTM, is canonically defined on the tangent bundle of any manifold, this is not the case for the adjoint structure, its image and kernel. In fact, it turns out that the latter are uniquely given by the horizontal distribution of a non-linear connection, which we state as follows.

Theorem 37.2.1. *There is a one-to-one correspondence between horizontal distributions* HTM *on a manifold* M *and adjoint structures* Θ *on* M*, which is given by* $\operatorname{im}\Theta = \ker\Theta = HTM$.

Proof. Let Θ be an adjoint structure, and define

$$\mathbf{h} = \Theta \circ J\,, \quad \mathbf{v} = J \circ \Theta\,. \tag{37.2.1}$$

To check that $\mathbf{h}$ is a horizontal projector, and hence $\mathbf{v}$ the corresponding vertical projector, note first that for all $\psi \in VTM$ one has $\mathbf{h}\psi = \Theta J\psi = 0$. Conversely, given $\psi \in TTM$ with $\mathbf{h}\psi = 0$, one has

$$\psi = \Theta J\psi + J\Theta\psi = J\Theta\psi \in \operatorname{im} J = VTM\,. \tag{37.2.2}$$

Hence, $\ker\mathbf{h} = VTM$. Further,

$$\mathbf{h} \circ \mathbf{h} = \Theta \circ J \circ \Theta \circ J = \Theta \circ (\mathrm{id}_{TTM} - \Theta \circ J) \circ J = \Theta \circ J = \mathbf{h}\,, \tag{37.2.3}$$

so that $\mathbf{h}$ is indeed a horizontal projector.

To prove the converse direction, let $TTM = HTM \oplus VTM$ be a non-linear connection on the tangent bundle and define Θ such that $\Theta|_{HTM} = 0$ and $\Theta|_{VTM}$ as the unique inverse of the vector bundle isomorphism $J|_{HTM} : HTM \to VTM$. One easily checks that this is an adjoint structure. ∎

From the image and the kernel of the adjoint structure one now easily derives the following helpful relations.

Theorem 37.2.2. *The horizontal and vertical projectors and the adjoint structure satisfy the relations*

$$\mathbf{h} \circ \Theta = \Theta \circ \mathbf{v} = \Theta\,, \quad \mathbf{v} \circ \Theta = \Theta \circ \mathbf{h} = 0\,. \tag{37.2.4}$$

Proof. The proof proceeds in full analogy to theorem 37.1.1, using the fact that $\operatorname{im}\Theta = \ker\Theta = HTM$. ■

It is now straightforward to calculate an explicit expression for the adjoint structure, using the relations given in definition 37.2.1 and theorem 37.2.2. This is most easily done using the Berwald basis, in which the adjoint structure takes the simple form

$$\Theta = \delta_a \otimes \delta\bar{x}^a \,. \tag{37.2.5}$$

Next, we proceed with another object, which can similarly be defined by a number of properties it is supposed to satisfy. These are given as follows.

Definition 37.2.2 (Almost product structure)**.** An *almost product structure* compatible with the tangent structure J on the tangent bundle TM a manifold M is a vector bundle homomorphism $\mathbb{P} : TTM \to TTM$ from $\varpi : TTM \to TM$ to itself, covering the identity on TM, which satisfies $J \circ \mathbb{P} = J$ and $\mathbb{P} \circ J = -J$.

Conventionally, the term "almost product structure" denotes a vector space (or vector bundle) endomorphism which squares to the identity, and thus has eigenvalues ± 1 on some respective subspaces or subbundles. Indeed this is the case also for the almost product structure $\mathbb{P}$ defined above, and we shall see in the following theorem that these subbundles are simply the horizontal and vertical bundles corresponding to a non-linear connection.

Theorem 37.2.3. *There is a one-to-one correspondence between horizontal distributions HTM on a manifold M and almost product structures $\mathbb{P}$ compatible with J on TM, which is given by $\mathbb{P}|_{HTM} = \mathrm{id}_{HTM}$ and $\mathbb{P}|_{VTM} = -\mathrm{id}_{VTM}$.*

Proof. Let $\mathbb{P}$ be an almost product structure compatible with J. Then we can construct a horizontal distribution as follows:

- First recall that $\operatorname{im} J = VTM$, so that we can write any element of VTM as $J\psi$ for some $\psi \in TTM$. By definition 37.2.2 we have $\mathbb{P}J\psi = -J\psi$. Hence, for all $\sigma \in VTM$ we have $\mathbb{P}\sigma = -\sigma$.
- Conversely, let $\psi \in TTM$ with $\mathbb{P}\psi = -\psi$. Acting with J, we find
$$-J\psi = J\mathbb{P}\psi = J\psi \,, \tag{37.2.6}$$
and hence $J\psi = 0$. It thus follows that $\psi \in VTM$.
- Taking the previous two items together, we find that $\mathbb{P}\psi = -\psi$ if and only if $\psi \in VTM$.
- Let $\psi \in TTM$ and $\sigma = \mathbb{P}\psi - \psi$. Note that $\sigma \in VTM$, since
$$J\sigma = J\mathbb{P}\psi - J\psi = J\psi - J\psi = 0 \,. \tag{37.2.7}$$
We then have
$$\mathbb{P}^2\psi = \mathbb{P}(\psi + \sigma) = \mathbb{P}\psi + \mathbb{P}\sigma = \psi + \sigma - \sigma = \psi \,, \tag{37.2.8}$$
and hence $\mathbb{P} \circ \mathbb{P} = \mathrm{id}_{TTM}$.
- Set
$$\mathbf{h} = \frac{\mathrm{id}_{TTM} + \mathbb{P}}{2} \,, \quad \mathbf{v} = \frac{\mathrm{id}_{TTM} - \mathbb{P}}{2} \,. \tag{37.2.9}$$
Clearly, $\mathbf{h}$ is a horizontal projector, since $\ker\mathbf{h} = VTM$ and
$$\mathbf{h} \circ \mathbf{h} = \frac{\mathrm{id}_{TTM} + 2\mathbb{P} + \mathbb{P} \circ \mathbb{P}}{4} = \frac{\mathrm{id}_{TTM} + \mathbb{P}}{2} = \mathbf{h} \,, \tag{37.2.10}$$
and $\mathbf{v}$ is the corresponding vertical projector.

To prove the converse direction, given a non-linear connection, set $\mathbb{P} = \mathbf{h} - \mathbf{v}$. One easily checks that this satisfies the conditions given in section 37.2.2. ∎

From the result obtained above one may now easily derive the following properties, which relate the almost product structure to the horizontal and vertical projectors.

Theorem 37.2.4. *The horizontal and vertical projectors and the almost product structure satisfy the relations*

$$\mathbf{h} \circ \mathbb{P} = \mathbb{P} \circ \mathbf{h} = \mathbf{h}\,, \quad \mathbf{v} \circ \mathbb{P} = \mathbb{P} \circ \mathbf{v} = -\mathbf{v}\,. \tag{37.2.11}$$

Proof. This follows directly from the formula $\mathbb{P} = \mathbf{h} - \mathbf{v}$ and the properties of the projectors. ∎

Finally, from the coordinate expression (37.1.3) of the horizontal and vertical projectors in the Berwald basis one immediately finds that the almost product structure can be written in the form

$$\mathbb{P} = \delta_a \otimes \mathrm{d}x^a - \bar{\partial}_a \otimes \delta\bar{x}^a\,. \tag{37.2.12}$$

The last tensor field we discuss in this section has a similar property compared to the almost product structure. While the latter squares to unity, we now consider a structure whose square (in the sense of function composition) is just the negative of unity. Such an object is usually known as an almost complex structure, and in the particular context we are studying here we define it as follows.

Definition 37.2.3 (Almost complex structure). An *almost complex structure* compatible with the tangent structure J on the tangent bundle TM a manifold M is a vector bundle homomorphism $\mathbb{F} : TTM \to TTM$ from $\varpi : TTM \to TM$ to itself, covering the identity on TM, which satisfies $\mathbb{F} \circ \mathbb{F} = -\mathrm{id}_{TTM}$ and $\mathbb{F} \circ J + J \circ \mathbb{F} = \mathrm{id}_{TTM}$.

We see that the almost complex structure shares the same relation with the tangent structure as the adjoint structure. This suggests that one may obtain a non-linear connection by following the same procedure as in the proof of theorem 37.2.1. To see that also the converse construction is possible, we prove the following statement.

Theorem 37.2.5. *There is a one-to-one correspondence between horizontal distributions HTM on a manifold M and almost complex structures $\mathbb{F}$ compatible with J on TM, which is given by* $\mathrm{im}(\mathbb{F} + J) = \ker(\mathbb{F} + J) = HTM$.

Proof. The proof proceeds similarly to the analogous theorem 37.2.1 for the adjoint structure. Given an almost complex structure $\mathbb{F}$ as in definition 37.2.3, define

$$\mathbf{h} = \mathbb{F} \circ J\,, \quad \mathbf{v} = J \circ \mathbb{F}\,. \tag{37.2.13}$$

To check that $\mathbf{h}$ is a horizontal projector, and hence $\mathbf{v}$ the corresponding vertical projector, note first that for all $\psi \in VTM$ one has $\mathbf{h}\psi = \mathbb{F}J\psi = 0$. Conversely, given $\psi \in TTM$ with $\mathbf{h}\psi = 0$, one has

$$\psi = \mathbb{F}J\psi + J\mathbb{F}\psi = J\mathbb{F}\psi \in \mathrm{im}\, J = VTM\,. \tag{37.2.14}$$

Hence, $\ker \mathbf{h} = VTM$. Further,

$$\mathbf{h} \circ \mathbf{h} = \mathbb{F} \circ J \circ \mathbb{F} \circ J = \mathbb{F} \circ (\mathrm{id}_{TTM} - \mathbb{F} \circ J) \circ J = \mathbb{F} \circ J = \mathbf{h}\,, \tag{37.2.15}$$

so that $\mathbf{h}$ is indeed a horizontal projector.

Conversely, given a non-linear connection with adjoint structure Θ, define $\mathbb{F} = \Theta - J$. This immediately yields

$$\mathbb{F} \circ J + J \circ \mathbb{F} = \Theta \circ J + J \circ \Theta = \mathrm{id}_{TTM} \tag{37.2.16}$$

and

$$\mathbb{F} \circ \mathbb{F} = (\Theta - J) \circ (\Theta - J) = -J \circ \Theta - \Theta \circ J = -\mathrm{id}_{TTM}\,, \tag{37.2.17}$$

so that $\mathbb{F}$ is an almost complex structure compatible with J. ■

Hence, we have another tensor field which equivalently characterizes a non-linear connection. Its properties are easily studied. First note the following relations with the projection operators.

Theorem 37.2.6. *The horizontal and vertical projectors and the almost complex structure satisfy the relations*

$$\mathbf{h} \circ \mathbb{F} = \mathbb{F} \circ \mathbf{v} = \Theta\,, \quad \mathbf{v} \circ \mathbb{F} = \mathbb{F} \circ \mathbf{h} = -J\,. \tag{37.2.18}$$

Proof. This follows directly from the formula $\mathbb{F} = \Theta - J$ and the properties of the tangent and adjoint structures given in theorems 37.1.1 and 37.2.2. ■

Also we can use the formula $\mathbb{F} = \Theta - J$ to derive the coordinate expression

$$\mathbb{F} = \delta_a \otimes \delta\bar{x}^a - \bar{\partial}_a \otimes dx^a \tag{37.2.19}$$

in the Berwald basis.

37.3 Horizontal lift

We have discussed in section 26.4 how general connections on fiber bundles give rise to the notion of horizontal lifts, thus allowing to lift certain objects from the base manifold to the total space. In particular, we have introduced in definition 26.6.1 the horizontal lift of a vector field. This can, of course, also be applied to non-linear connections on the tangent bundle. In analogy with the previous definitions in section 29.5 we introduce the following notation.

Definition 37.3.1 (Horizontal lift of a vector field). Let M be a manifold equipped with a non-linear connection on the tangent bundle and $X \in \mathrm{Vect}(M)$ a vector field. Its *horizontal lift* is the unique vector field $\overset{\mathrm{H}}{X} \in \mathrm{Vect}(TM)$ which satisfies $\mathbf{h}\overset{\mathrm{H}}{X} = \overset{\mathrm{H}}{X}$ and $\tau_* \circ \overset{\mathrm{H}}{X} = X \circ \tau$.

Indeed, we see that this is just a special case of the general definition 26.6.1 we gave earlier. It is worth mentioning again that the horizontal lift is indeed unique. This follows from the fact that under a non-linear connection the double tangent bundle splits in the form $TTM = HTM \oplus VTM$, so that every vector field $Y \in \mathrm{Vect}(TM)$ splits into $Y = \mathbf{h}Y + \mathbf{v}Y$. For $Y = \overset{\mathrm{H}}{X}$, we see that the first condition fixes the vertical part to vanish, $\mathbf{v}\overset{\mathrm{H}}{X} = 0$. The horizontal part is uniquely fixed by the second condition, since $\ker \tau_* = VTM$. In the Berwald basis, one thus finds the expression

$$\overset{\mathrm{H}}{X} = X^a \delta_a = X^a(\partial_a - N^b{}_a \bar{\partial}_b)\,. \tag{37.3.1}$$

In the particular case of the tangent bundle, there exist also the vertical and complete lift introduced in section 29.5. We find that the horizontal lift is related to them by a number of useful relations, such as the following.

Theorem 37.3.1. *The vertical and horizontal lifts of a vector field $X \in \mathrm{Vect}(M)$ are related by $\overset{H}{X} = \Theta\overset{V}{X}$ and $\overset{V}{X} = J\overset{H}{X}$.*

Proof. We start with the second proposition. Following the definitions 29.2.1 of the tangent structure and 29.5.1 of the vertical lift as tangent vectors of curves, we have for $v \in TM$:

$$(J\overset{\text{H}}{X})(v) = \frac{\mathrm{d}}{\mathrm{d}t}\left[v + t(\tau_* \circ \overset{\text{H}}{X})(v)\right]\bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\left[v + t(X \circ \tau)(v)\right]\bigg|_{t=0} = \overset{\text{V}}{X}(v)\,. \tag{37.3.2}$$

We then use this result to show the first proposition. Writing the horizontal projector as $\mathbf{h} = \Theta \circ J$, we have

$$\Theta\overset{\text{V}}{X} = \Theta J\overset{\text{H}}{X} = \mathbf{h}\overset{\text{H}}{X} = \overset{\text{H}}{X}\,, \tag{37.3.3}$$

since $\overset{\text{H}}{X}$ is horizontal by definition. ■

This relation can also be illustrated using the Berwald basis. Given a vector field $X \in \mathrm{Vect}(M)$, which we write in the form $X = X^a\partial_a$ using the coordinate basis of TM, its horizontal lift is simply given by $\overset{\text{H}}{X} = X^a\delta_a$ in the Berwald basis. Now one easily sees that

$$J\overset{\text{H}}{X} = X^a\bar{\partial}_a = \overset{\text{V}}{X}\,, \tag{37.3.4}$$

while conversely

$$\Theta\overset{\text{V}}{X} = X^a\delta_a = \overset{\text{H}}{X}\,. \tag{37.3.5}$$

Another helpful relation is the following.

Theorem 37.3.2. *The complete and horizontal lifts of a vector field $X \in \mathrm{Vect}(M)$ are related by $\overset{H}{X} = \mathbf{h}\overset{C}{X}$.*

Proof. Since $\mathbf{h}\overset{\text{H}}{X} = \overset{\text{H}}{X}$, it is sufficient to show that the difference $\overset{\text{H}}{X} - \overset{\text{C}}{X}$ is vertical for any vector field $X \in \mathrm{Vect}(M)$. This can be see using theorems 29.5.4 and 37.3.1, from which follows

$$J(\overset{\text{H}}{X} - \overset{\text{C}}{X}) = \overset{\text{V}}{X} - \overset{\text{V}}{X} = 0\,. \tag{37.3.6}$$

so that their difference lies in $\ker J = VTM$. Hence,

$$\mathbf{h}\overset{\text{C}}{X} = \mathbf{h}\overset{\text{H}}{X} = \overset{\text{H}}{X}\,. \quad ■ \tag{37.3.7}$$

Also this becomes obvious in coordinates, since

$$\mathbf{h}\overset{\text{C}}{X} = (\delta_c \otimes \mathrm{d}x^c)(X^a\partial_a + \bar{x}^b\partial_b X^a\bar{\partial}_a) = X^a\delta_a = \overset{\text{H}}{X}\,. \tag{37.3.8}$$

37.4 Homogeneous connections

A class of non-linear connections which is of particular interest is given by homogeneous connections. As we will see later in chapter 38, there exists a relation between homogeneous connections and sprays, and in chapter 40 we will see how to derive a particular homogeneous connection from a Finsler geometry. We start our discussion of homogeneous connections with the following definition.

Definition 37.4.1 (Homogeneous connection). A non-linear connection on the tangent bundle TM of a manifold M is called *homogeneous* if and only if $\chi_{\lambda *}\psi \in HTM$ for all $\lambda \in \mathbb{R}$ and $\psi \in HTM$, where $\chi : \mathbb{R} \times TM \to TM$ denotes the dilatations.

In other words, we demand that the horizontal distribution is invariant under dilatations. While this definition is the most intuitive from a geometric point of view, it is not necessarily the most practical in terms of proving the homogeneity or non-homogeneity of a given connection. It is simpler if instead of finite diffeomorphisms, we can study infinitesimal operations. Recalling that the dilatations are infinitesimally generated by the Liouville vector field, we come to the following statement.

Theorem 37.4.1. *A non-linear connection is homogeneous if and only if* $\mathbf{v}[\mathbf{c}, \mathbf{h}X] = 0$ *for all* $X \in \mathrm{Vect}(TM)$*.*

Proof. In case of a homogeneous non-linear connection, one has

$$\begin{aligned} \mathbf{v}[\mathbf{c}, \mathbf{h}X] &= \mathbf{v}\mathcal{L}_{\mathbf{c}}\mathbf{h}X \\ &= \mathbf{v} \lim_{\lambda \to 0} \frac{\chi_\lambda^* \mathbf{h}X - \mathbf{h}X}{\lambda} \\ &= \lim_{\lambda \to 0} \frac{\mathbf{v}\chi_\lambda^* \mathbf{h}X - \mathbf{v}\mathbf{h}X}{\lambda} \\ &= 0\,, \end{aligned} \tag{37.4.1}$$

where we used the fact that the pullback of a horizontal vector field along χ_λ, defined via the pushforward of every vector along the inverse $\chi_{-\lambda}$ following definition 12.1.1, is again horizontal for a homogeneous connection:

$$(\mathbf{v}\chi_\lambda^* \mathbf{h}X)(v) = (\mathbf{v} \circ \chi_{-\lambda *} \circ (\mathbf{h}X) \circ \chi_\lambda)(v) = 0 \tag{37.4.2}$$

for $v \in TM$. Conversely, let $\psi \in HTM$. ►...◄ ■

We can simplify this statement even further. Recall that we have characterized non-linear connections by a number of tensor fields in section 37.2, and so it would be most convenient if we could employ those in order to assess the homogeneity of a connection. We will now see that this is indeed the case. For this purpose, it is useful to introduce the following tensor field.

Definition 37.4.2 (Tension). Let M be a manifold equipped with a non-linear connection on the tangent bundle. Its *tension* is the vector-valued one-form $\mathbb{T} \in \Omega^1(TM, TTM)$ defined by the Lie derivative $\mathbb{T} = \frac{1}{2}\mathcal{L}_{\mathbf{c}}\mathbb{P}$, where $\mathbb{P}$ is the almost product structure.

We have seen in the preceding sections that the Lie derivative with respect to the Liouville vector field measures the order of homogeneity of tensor fields defined on the tangent bundle. This applies also here. In fact, it turns out that we can check the homogeneity of a non-linear connection as follows.

Theorem 37.4.2. *A non-linear connection is homogeneous if and only if it has vanishing tension,* $\mathbb{T} = 0$*, i.e., its almost product structure* $\mathbb{P}$ *is 0-homogeneous,* $\mathcal{L}_{\mathbf{c}}\mathbb{P} = 0$*.*

Proof. By direct calculation, we have for any vector field $X \in \mathrm{Vect}(TM)$:

$$\begin{aligned} \mathbb{T}X &= \frac{1}{2}(\mathcal{L}_{\mathbf{c}}\mathbb{P})X \\ &= \frac{1}{2}([\mathbf{c}, \mathbb{P}X] - \mathbb{P}[\mathbf{c}, X]) \\ &= \frac{1}{2}\{(\mathbf{h} + \mathbf{v})[\mathbf{c}, (\mathbf{h} - \mathbf{v})X] - (\mathbf{h} - \mathbf{v})[\mathbf{c}, (\mathbf{h} + \mathbf{v})X]\} \\ &= \mathbf{v}[\mathbf{c}, \mathbf{h}X] - \mathbf{h}[\mathbf{c}, \mathbf{v}X] \\ &= \mathbf{v}[\mathbf{c}, \mathbf{h}X]\,, \end{aligned} \tag{37.4.3}$$

where the last term vanishes since both $\mathbf{c}$ and $\mathbf{v}X$, and hence also their Lie bracket, is vertical. We thus find that $\mathbf{v}[\mathbf{c}, \mathbf{h}X] = 0$ for all $X \in \mathrm{Vect}(TM)$ if and only if the tension vanishes, $\mathbb{T} = 0$. Following theorem 37.4.1, the former is equivalent to the non-linear connection being homogeneous, and hence the same holds for the latter. ■

We now illustrate the calculations shown in this section by using coordinates. We start by calculating the tension. From the coordinate expression 37.2.12 for the almost product structure in the Berwald basis follows:

$$\begin{aligned}
\mathbb{T} &= \frac{1}{2}\mathcal{L}_{\mathbf{c}}\mathbb{P} \\
&= \frac{1}{2}\mathcal{L}_{\bar{x}^a\bar{\partial}_a}(\delta_b \otimes \mathrm{d}x^b - \bar{\partial}_b \otimes \delta\bar{x}^b) \\
&= \frac{1}{2}\Big\{[\bar{x}^a\bar{\partial}_a, \partial_b - N^c{}_b\bar{\partial}_c] \otimes \mathrm{d}x^b + \delta_b \otimes \mathrm{d}\iota_{\bar{x}^a\bar{\partial}_a}\mathrm{d}x^b + \iota_{\bar{x}^a\bar{\partial}_a}\mathrm{dd}x^b \\
&\quad - [\bar{x}^a\bar{\partial}_a, \bar{\partial}_b] \otimes \delta\bar{x}^b - \bar{\partial}_b \otimes \mathrm{d}\iota_{\bar{x}^a\bar{\partial}_a}(\mathrm{d}\bar{x}^b + N^b{}_c\mathrm{d}x^c) - \bar{\partial}_b \otimes \iota_{\bar{x}^a\bar{\partial}_a}\mathrm{d}(\mathrm{d}\bar{x}^b + N^b{}_c\mathrm{d}x^c)\Big\} \\
&= \frac{1}{2}\left(N^a{}_b\bar{\partial}_a \otimes \mathrm{d}x^b - \bar{x}^a\bar{\partial}_a N^c{}_b\bar{\partial}_c \otimes \mathrm{d}x^b + \bar{\partial}_b \otimes \delta\bar{x}^b - \bar{\partial}_b \otimes \mathrm{d}\bar{x}^b - \bar{x}^a\bar{\partial}_a N^b{}_c\bar{\partial}_b \otimes \mathrm{d}x^c\right) \\
&= (N^a{}_b - \bar{x}^c\bar{\partial}_c N^a{}_b)\bar{\partial}_a \otimes \mathrm{d}x^b\,.
\end{aligned} \tag{37.4.4}$$

We thus see that the non-linear connection is homogeneous if and only if its coefficients satisfy

$$\bar{x}^c\bar{\partial}_c N^a{}_b = N^a{}_b\,, \tag{37.4.5}$$

and hence are homogeneous functions of order 1. For comparison, let $X = X^a\delta_a + \bar{X}^a\bar{\partial}_a$ and calculate

$$\begin{aligned}
\mathbf{v}[\mathbf{c}, \mathbf{h}X] &= \mathbf{v}[\bar{x}^a\bar{\partial}_a, X^b(\partial_b - N^c{}_b\bar{\partial}_c)] \\
&= \mathbf{v}(\bar{x}^a\bar{\partial}_a X^b\delta_b - \bar{x}^a X^b\bar{\partial}_a N^c{}_b\bar{\partial}_c + X^b N^a{}_b\bar{\partial}_a) \\
&= (N^a{}_b - \bar{x}^c\bar{\partial}_c N^a{}_b)X^b\bar{\partial}_a\,,
\end{aligned} \tag{37.4.6}$$

in agreement with the coordinate-free calculation (37.4.3).

►Calculate pushforward and homogeneity in coordinates.◄

Theorem 37.4.3. *A non-linear connection on TM is homogeneous if and only the horizontal lift of every vector field from M to TM is a homogeneous vector field of order 0.*

Proof. For the complete lift of a vector field $X \in \mathrm{Vect}(M)$ we have

$$\begin{aligned}
\mathbb{T}\overset{\mathrm{C}}{X} &= \frac{1}{2}(\mathcal{L}_{\mathbf{c}}\mathbb{P})\overset{\mathrm{C}}{X} \\
&= \frac{1}{2}\left(\left[\mathbf{c}, \mathbb{P}\overset{\mathrm{C}}{X}\right] - \mathbb{P}\left[\mathbf{c}, \overset{\mathrm{C}}{X}\right]\right) \\
&= \frac{1}{2}\left[\mathbf{c}, 2\mathbf{h}\overset{\mathrm{C}}{X} - \overset{\mathrm{C}}{X}\right] \\
&= \left[\mathbf{c}, \mathbf{h}\overset{\mathrm{C}}{X}\right] \\
&= \left[\mathbf{c}, \overset{\mathrm{H}}{X}\right]\,.
\end{aligned} \tag{37.4.7}$$

This vanishes if and only if $\overset{\mathrm{H}}{X}$ is 0-homogeneous. ►Show that then also $\mathbb{T} = 0$.◄ ■

Also this can be illustrated in coordinates, using the relation (37.3.1) which states that the horizontal lift of a vector field $X \in \mathrm{Vect}(M)$ is given by $\overset{\mathrm{H}}{X} = X^a\delta_a$. We then have

$$\left[\mathbf{c}, \overset{\mathrm{H}}{X}\right] = [\bar{x}^c\bar{\partial}_c, X^a(\partial_a - N^b{}_a\bar{\partial}_b)] = X^a(N^b{}_a - \bar{x}^c\bar{\partial}_c N^b{}_a)\bar{\partial}_b = \mathbb{T}\overset{\mathrm{C}}{X}\,, \tag{37.4.8}$$

which vanishes for every vector field X if and only if the homogeneity condition (37.4.5) is satisfied.

37.5 Torsion

For a non-linear connection, different notions of torsion exist. Here we start with the following definition.

Definition 37.5.1 (Weak torsion)**.** Let M be a manifold equipped with a non-linear connection on the tangent bundle. Its *weak torsion* is the vector-valued two-form $\mathbf{t} \in \Omega^2(TM, TTM)$ defined by the Frölicher-Nijenhuis bracket $\mathbf{t} = [\![\mathbf{h}, J]\!]$, where $\mathbf{h}$ is the horizontal projector.

One can easily expand this expression by using the definition and properties of the Frölicher-Nijenhuis bracket as follows.

Theorem 37.5.1. *For any two vector fields $X, Y \in \mathrm{Vect}(TM)$, the weak torsion $\mathbf{t}$ of the non-linear connection satisfies*

$$\mathbf{t}(X,Y) = J[\mathbf{h}X, \mathbf{h}Y] - \mathbf{v}[JX, \mathbf{h}Y] - \mathbf{v}[\mathbf{h}X, JY] \,. \tag{37.5.1}$$

Proof. A direct calculation using theorem 17.6.5 yields

$$\begin{aligned}\mathbf{t}(X,Y) &= [\![\mathbf{h}, J]\!](X,Y)\\ &= J([\mathbf{h}X, Y] + [X, \mathbf{h}Y]) + \mathbf{h}([JX, Y] + [X, JY]) - [JX, \mathbf{h}Y] - [\mathbf{h}X, JY] - (J\mathbf{h} + \mathbf{h}J)[X,Y]\\ &= \mathbf{h}([JX, \mathbf{h}Y] + [\mathbf{h}X, JY]) - [JX, \mathbf{h}Y] - [\mathbf{h}X, JY] - J([X,Y] - [\mathbf{h}X, Y] - [X, \mathbf{h}Y])\\ &= J[\mathbf{h}X, \mathbf{h}Y] - \mathbf{v}[JX, \mathbf{h}Y] - \mathbf{v}[\mathbf{h}X, JY] \,,\end{aligned} \tag{37.5.2}$$

where we used the fact that the vertical distribution is integrable, so that $[\mathbf{v}X, \mathbf{v}Y]$ is vertical, and hence $J[\mathbf{v}X, \mathbf{v}Y] = 0$. ■

It thus follows in particular that the weak torsion vanishes on any vertical vector field, and that it is itself vertical, hence

$$\mathbf{t}(\mathbf{v}X, Y) = \mathbf{t}(X, \mathbf{v}Y) = \mathbf{h}\mathbf{t}(X,Y) = 0 \,. \tag{37.5.3}$$

Using the relation (37.5.1), one can now easily derive a coordinate expression for the weak torsion. For this purpose, we expand the vector fields $X = X^a\delta_a + \bar{X}^a\bar{\partial}_a$ and $Y = Y^a\delta_a + \bar{Y}^a\bar{\partial}_a$ in the Berwald basis. A straightforward calculation then shows that

$$\begin{aligned}\mathbf{t}(X,Y) &= J[\mathbf{h}X, \mathbf{h}Y] - \mathbf{v}[JX, \mathbf{h}Y] - \mathbf{v}[\mathbf{h}X, JY]\\ &= J[X^a\delta_a, Y^b\delta_b] - \mathbf{v}[X^a\bar{\partial}_a, Y^b\delta_b] - \mathbf{v}[X^a\delta_a, Y^b\bar{\partial}_b]\\ &= J[X^a(\partial_a - N^c{}_a\bar{\partial}_c), Y^b(\partial_b - N^d{}_b\bar{\partial}_d)]\\ &\quad - \mathbf{v}[X^a\bar{\partial}_a, Y^b(\partial_b - N^d{}_b\bar{\partial}_d)] - \mathbf{v}[X^a(\partial_a - N^c{}_a\bar{\partial}_c), Y^b\bar{\partial}_b]\\ &= J(X^a\delta_aY^b\partial_b - Y^b\delta_bX^a\partial_a)\\ &\quad - \mathbf{v}(X^a\delta_aY^b\bar{\partial}_b - Y^b\delta_bX^a\bar{\partial}_a - X^aY^b\bar{\partial}_aN^d{}_b\bar{\partial}_d + Y^bX^a\bar{\partial}_bN^c{}_a\bar{\partial}_c)\\ &= 2X^aY^b\bar{\partial}_{[a}N^c{}_{b]}\bar{\partial}_c \,,\end{aligned} \tag{37.5.4}$$

where we used the fact that J vanishes on vertical vectors, while $\mathbf{v}$ vanishes on horizontal vectors, to omit various terms which appear at intermediate steps. Hence, we can write the weak torsion as

$$\mathbf{t} = \bar{\partial}_{[a}N^c{}_{b]}\bar{\partial}_c \otimes \mathrm{d}x^a \wedge \mathrm{d}x^b \,. \tag{37.5.5}$$

In particular, one may consider non-linear connections whose weak torsion vanishes. These are given their own name:

Definition 37.5.2 (Symmetric non-linear connection)**.** A non-linear connection on a manifold M is called *symmetric* if and only if its weak torsion vanishes, $\mathbf{t} = 0$.

It is straightforward to see that a non-linear connection is symmetric if and only if

$$\bar{\partial}_{[a}N^c{}_{b]} = 0\,, \tag{37.5.6}$$

hence justifying the name.

Another type of torsion which we may define for a non-linear connection is the following.

Definition 37.5.3 (Strong torsion)**.** Let M be a manifold equipped with a non-linear connection on the tangent bundle. Its *strong torsion* is the vector-valued one-form $\mathbf{T} \in \Omega^1(TM, TTM)$ defined by $\mathbf{T} = \iota_{\Theta\mathbf{c}}\mathbf{t} + \mathbb{T}$.

Also for the strong torsion one easily derives a coordinate expression. Using the Berwald basis, one has $\Theta\mathbf{c} = \bar{x}^a\delta_a$. Using the coordinate expression (37.5.5) for the weak torsion, one thus has

$$\iota_{\Theta\mathbf{c}}\mathbf{t} = 2\bar{x}^c\bar{\partial}_{[c}N^a{}_{b]}\bar{\partial}_a \otimes \mathrm{d}x^b\,, \tag{37.5.7}$$

and therefore

$$\begin{aligned} \mathbf{T} &= \iota_{\Theta\mathbf{c}}\mathbf{t} + \mathbb{T} \\ &= \left[2\bar{x}^c\bar{\partial}_{[c}N^a{}_{b]} + (N^a{}_b - \bar{x}^c\bar{\partial}_c N^a{}_b)\right]\bar{\partial}_a \otimes \mathrm{d}x^b \\ &= (N^a{}_b - \bar{x}^c\bar{\partial}_b N^a{}_c)\bar{\partial}_a \otimes \mathrm{d}x^b\,. \end{aligned} \tag{37.5.8}$$

This can also be expressed as follows.

Theorem 37.5.2. *The strong torsion of a non-linear connection is given by*

$$\mathbf{T}X = J[\Theta\mathbf{c}, \mathbf{h}X] - \mathbf{v}[\Theta\mathbf{c}, JX]\,. \tag{37.5.9}$$

Proof. By theorem 37.5.1, the weak torsion satisfies

$$\begin{aligned} \iota_{\Theta\mathbf{c}}\mathbf{t}(X) &= \mathbf{t}(\Theta\mathbf{c}, X) \\ &= J[\mathbf{h}\Theta\mathbf{c}, \mathbf{h}X] - \mathbf{v}[J\Theta\mathbf{c}, \mathbf{h}X] - \mathbf{v}[\mathbf{h}\Theta\mathbf{c}, JX] \\ &= J[\Theta\mathbf{c}, \mathbf{h}X] - \mathbf{v}[\mathbf{c}, \mathbf{h}X] - \mathbf{v}[\Theta\mathbf{c}, JX]\,. \end{aligned} \tag{37.5.10}$$

Note that the second term is just the tension according to theorem 37.4.2. Hence, we find

$$\mathbf{T}X = \iota_{\Theta\mathbf{c}}\mathbf{t}X + \mathbb{T}X = J[\Theta\mathbf{c}, \mathbf{h}X] - \mathbf{v}[\Theta\mathbf{c}, JX]\,. \tag{37.5.11}$$

■

To derive this in coordinates, let $X = X^a\delta_a + \bar{X}^a\bar{\partial}_a$ and calculate

$$\begin{aligned} \mathbf{T}X &= J[\bar{x}^b\delta_b, X^a\delta_a] - \mathbf{v}[\bar{x}^b\delta_b, X^a\bar{\partial}_a] \\ &= \bar{\partial}_e \otimes \mathrm{d}x^e[\bar{x}^b(\partial_b - N^c{}_b\bar{\partial}_c), X^a(\partial_a - N^d{}_a\bar{\partial}_d)] - \bar{\partial}_e \otimes (\mathrm{d}\bar{x}^e + N^e{}_f\mathrm{d}x^f)[\bar{x}^b(\partial_b - N^c{}_b\bar{\partial}_c), X^a\bar{\partial}_a] \\ &= \left[\bar{x}^b(\partial_b - N^c{}_b\bar{\partial}_c)X^a + X^bN^a{}_b - \bar{x}^b(\partial_b - N^c{}_b\bar{\partial}_c)X^a - X^c\bar{x}^b\bar{\partial}_c N^a{}_b\right]\bar{\partial}_a \\ &= X^b(N^a{}_b - \bar{x}^c\bar{\partial}_b N^a{}_c)\bar{\partial}_a\,, \end{aligned} \tag{37.5.12}$$

in agreement with the previous result (37.5.8).

It is clear from the definition 37.5.3 of the strong torsion that if the weak torsion and the tension vanish, also the strong torsion vanishes. We can, however, find an even stronger result as follows.

Theorem 37.5.3. *The strong torsion of a non-linear connection vanishes,* $\mathbf{T} = 0$*, if and only if both its weak torsion and tension vanish,* $\mathbf{t} = 0$ *and* $\mathbb{T} = 0$.

Proof. We defer this proof to the proof of theorem 38.5.4. ■

The result is easily obtained using coordinates. If the strong torsion vanishes, $\mathbf{T} = 0$, we have

$$N^a{}_b = \bar{x}^c \bar{\partial}_b N^a{}_c \tag{37.5.13}$$

and thus

$$\bar{\partial}_{[a} N^c{}_{b]} = \bar{\partial}_{[a}(\bar{x}^d \bar{\partial}_{b]} N^c{}_d) = -\bar{\partial}_{[a} N^c{}_{b]}\,, \tag{37.5.14}$$

which must therefore vanish, and so $\mathbf{t} = 0$. It then further follows from (37.5.8) that also $\mathbb{T} = 0$. The converse statement is obvious from the definition of the strong torsion.

37.6 Curvature

Since any non-linear connection on the tangent bundle is in particular a connection, it also possesses a notion of curvature. In this case, one conventionally employs the following definition.

Definition 37.6.1 (Curvature of a non-linear connection)**.** Let M be a manifold equipped with a non-linear connection on the tangent bundle. Its *curvature* is the vector-valued two-form $\mathbf{R} \in \Omega^2(TM, TTM)$ defined by the Nijenhuis tensor $\mathbf{R} = -N_{\mathbf{h}}$, where $\mathbf{h}$ is the horizontal projector.

The obvious question arises how this notion is related to the definition 26.10.1. The following statement answers this question.

Theorem 37.6.1. *For any two vector fields* $X, Y \in \mathrm{Vect}(TM)$*, the curvature* $\mathbf{R}$ *of the non-linear connection satisfies*

$$\mathbf{R}(X, Y) = -\mathbf{v}[\mathbf{h}X, \mathbf{h}Y]\,. \tag{37.6.1}$$

Proof. A direct calculation using theorem 17.6.6 yields

$$\begin{aligned}\mathbf{R}(X, Y) &= -N_{\mathbf{h}}(X, Y)\\ &= \mathbf{h}\left([\mathbf{h}X, Y] + [X, \mathbf{h}Y]\right) - \mathbf{h}^2[X, Y] - [\mathbf{h}X, \mathbf{h}Y]\\ &= -\mathbf{h}[\mathbf{v}X, \mathbf{v}Y] - \mathbf{v}[\mathbf{h}X, \mathbf{h}Y]\\ &= -\mathbf{v}[\mathbf{h}X, \mathbf{h}Y]\,,\end{aligned} \tag{37.6.2}$$

where in the last line we used the fact that the vertical distribution is integrable, so that $[\mathbf{v}X, \mathbf{v}Y]$ is vertical. ■

37.7 Autoparallel curves

Definition 37.7.1 (Autoparallel curve)**.** Let M be a manifold equipped with a non-linear connection on the tangent bundle. An *autoparallel curve* is a curve $\gamma \in C^\infty(\mathbb{R}, M)$ such that its canonical lift $\dot{\gamma} \in C^\infty(\mathbb{R}, TM)$ is horizontal, $\mathbf{v} \circ \ddot{\gamma} = 0$.

Theorem 37.7.1. *For every tangent vector $v \in TM$, an non-linear connection defines a unique autoparallel curve $\gamma : \mathbb{R} \to M$ such that $\dot{\gamma}(0) = v$.*

Proof. ►. . . ◄ ■

37.8 Affine bundle of connections

37.9 Pullback and Lie derivative

37.10 Linear connections

Chapter 38

Sprays and semisprays

38.1 Semisprays

In the previous chapter on non-linear connections, we have been working intensively with vector fields defined over the tangent bundle TM of a manifold M. We now come to a particular class of such vector fields, which have particularly useful properties which make them suitable for the description of higher order differential equations, as we will discuss in section 48.2, and which are closely related to non-linear connections, as shown in section 38.3. To introduce these vector fields, we start with the following definition.

Definition 38.1.1 (Semispray)**.** Let M be a manifold. A vector field $X \in \mathrm{Vect}(TM)$ on the tangent bundle TM of M is called a *semispray* (or *second-order vector field*) if and only if $JX = \mathbf{c}$, where J is the tangent structure and $\mathbf{c}$ is the Liouville vector field.

Here we follow the definition given in [BM07, def. 4.1.1]. In contrast, the same definition is only called *second-order vector field* in [SLK14, def. 5.1.1], while demanding additional properties for a *semispray* in [SLK14, 5.1.23].

From this definition one can easily derive the most general form of a semispray in induced coordinates $(x^a, \bar{x}^a)$ on TM, as introduced in section 29.1. Recall that in these coordinates the tangent structure has the form (29.2.6), while the Liouville vector field is given by (19.9.2). Writing a general vector field $X \in \mathrm{Vect}(TM)$ as $X = X^a\partial_a + \bar{X}^a\bar{\partial}_a$, we thus have the condition

$$\bar{x}^a\bar{\partial}_a = \mathbf{c} = JX = (\bar{\partial}_b \otimes \mathrm{d}x^b)(X^a\partial_a + \bar{X}^a\bar{\partial}_a) = X^a\bar{\partial}_a\,. \tag{38.1.1}$$

Hence, X is a semispray if and only if $X^a = \bar{x}^a$. We thus find that a semispray has the general coordinate expression $X = \bar{x}^a\partial_a + \bar{X}^a\bar{\partial}_a$, and so it is fully characterized by the coefficients $\bar{X}^a$. In the following, we will show further, equivalent possibilities to come to this conclusion is by using the following statement.

Theorem 38.1.1. *A vector field $X \in \mathrm{Vect}(TM)$ on the tangent bundle TM of a manifold M is a semispray if and only if $\tau_* \circ X = \mathrm{id}_{TM}$, where $\tau : TM \to M$ is the projection map of the tangent bundle.*

Proof. Recall that a vector field on TM is a map $X : TM \to TTM$. Its composition with the differential $\tau_* : TTM \to TM$ of the tangent bundle projection τ is thus indeed a map from TM to itself. Let $v \in TM$ and $\psi = X(v)$, and so $v = \varpi(\psi)$, where $\varpi : TTM \to TM$ is the bundle projection of the double tangent bundle. Then recall from definition 29.2.1 that the tangent structure J assigns to ψ the tangent vector $\dot{\gamma}_\psi(0)$ of the curve

$$\begin{array}{rcccc} \gamma_\psi & : & \mathbb{R} & \to & TM \\ & & \lambda & \mapsto & \gamma_\psi(\lambda) = \varpi(\psi) + \lambda\tau_*(\psi) \end{array} . \tag{38.1.2}$$

Hence, $JX : TM \to TTM$ is the map that assigns to $v \in TM$ the tangent vector $\dot{\gamma}_{X(v)}(0)$.

We now compare this construction to the Liouville vector field $\mathbf{c} \in \operatorname{Vect}(TM)$. Recall that the latter is the generator of dilatations. This means that $\mathbf{c}(v) \in TTM$ is the tangent vector $\dot{\chi}_v(0)$ to the curve

$$\begin{array}{rcccc} \chi_v & : & \mathbb{R} & \to & TM \\ & & \lambda & \mapsto & \chi_v(\lambda) = e^\lambda v \end{array} . \tag{38.1.3}$$

for any $v \in TM$. One easily sees that $\chi_v(0) = v = \gamma_{X(v)}(0)$. Further, since both curves lie inside the vector space T_vTM, taking the derivative with respect to λ shows that their tangent vectors at $\lambda = 0$ agree, $\dot{\chi}_v(0) = \dot{\gamma}_{X(v)}(0)$, if and only if

$$v = \tau_*(\psi) = \tau_*(X(v)) \, . \tag{38.1.4}$$

It thus follows that this holds for all $v \in TM$, and hence $\mathbf{c} = JX$, so that X is a semispray, if and only if $\tau_* \circ X = \mathrm{id}_{TM}$. ■

We can now check that this leads to the same coordinate form of a semispray X, making use of the coordinate expression (29.7.1). Writing v in induced coordinates as $(x^a, \bar{x}^a)$, and $X(v)$ as $(x^a, \bar{x}^a, X^a, \bar{X}^a)$, we have

$$\tau_*(X(v)) = \tau_*(X^a\partial_a + \bar{X}^a\bar{\partial}_a) = X^a\partial_a \, , \tag{38.1.5}$$

and hence $v = \tau_*(X(v))$ if and only if $X^a = \bar{x}^a$.

Theorem 38.1.2. *A vector field $X \in \operatorname{Vect}(TM)$ on the tangent bundle TM of a manifold M is a semispray if and only if $\kappa \circ X = X$, where $\kappa : TTM \to TTM$ is the canonical involution of the double tangent bundle.*

Proof. ►. . . ◄ ■

Theorem 38.1.3. *A vector field $X \in \operatorname{Vect}(TM)$ on the tangent bundle TM of a manifold M is a semispray if and only if every integral curve, i.e., every curve $\Gamma : \mathbb{R} \to TM$ with $\dot{\Gamma} = X \circ \Gamma$, is the canonical lift of a curve on M.*

Proof. Recall from theorem 10.2.3 that a curve Γ on TM is a canonical lift if and only if $\tau_* \circ \dot{\Gamma} = \Gamma$. For an integral curve of X, this means that it is a canonical lift if and only if $\tau_* \circ X \circ \Gamma = \Gamma$. Hence, *every* integral curve is a canonical lift if and only if $\tau_* \circ X = \mathrm{id}_{TM}$, and hence if and only if X is a semispray according to theorem 38.1.1. ■

Having obtained several characterizations of semisprays, we now study their properties. An important result, which finds application in numerous other theorems, is the following.

Theorem 38.1.4 (Grifone's identity)**.** *For every semispray $X \in \operatorname{Vect}(TM)$ and arbitrary vector field $Y \in \operatorname{Vect}(TM)$ holds*

$$J[JY, X] = JY \, . \tag{38.1.6}$$

Proof. We make use theorem 29.2.7, which states that the Nijenhuis tensor N_J of J vanishes, as well as theorem 17.6.6 to compute

$$\begin{aligned}
0 &= N_J(X,Y) \\
&= \frac{1}{2}[\![J,J]\!](X,Y) \\
&= [JX,JY] + J^2[X,Y] - J([JX,Y] + [X,JY]) \\
&= [\mathbf{c},JY] - J[\mathbf{c},Y] - J[X,JY] \\
&= (\mathcal{L}_{\mathbf{c}}J)Y + J[JY,X] \\
&= -JY + J[JY,X]\,,
\end{aligned} \tag{38.1.7}$$

where we used the defining property $JX = \mathbf{c}$ of a semispray, the relation 16.6.1 for the Lie derivative of an endomorphism and the homogeneity of the tangent structure J stated in theorem 29.2.5. ∎

Given a semispray, we can also obtain another semispray as follows.

Theorem 38.1.5. *For any semispray $X \in \mathrm{Vect}(TM)$, also $[\mathbf{c},X]$ is a semispray.*

Proof. This follows from Grifone's identity 38.1.4, since

$$J[\mathbf{c},X] = J[JX,X] = JX = \mathbf{c}\,. \quad ∎ \tag{38.1.8}$$

We may also regard a semispray as a vector-valued zero-form on the tangent bundle, and apply the theory of graded derivations detailed in chapter 17, to derive a few more useful properties of semisprays, which directly follow from their definition. This allows us to state the following.

Theorem 38.1.6. *If $X \in \mathrm{Vect}(TM)$ is a semispray, then it satisfies*

$$[\iota_X,\iota_J] = \iota_{\mathbf{c}}\,, \quad [\iota_X,\mathcal{L}_J] = \mathcal{L}_{\mathbf{c}} + \iota_{[\![J,X]\!]}\,. \tag{38.1.9}$$

Proof. The first relation can be shown using the explicit expression for the Nijenhuis-Richardson bracket given in theorem 17.4.4, from which follows

$$[X,J]^{\wedge} = \iota_X J - (-1)^{1\cdot 2}\iota_J X = \iota_X J = JX = \mathbf{c}\,, \tag{38.1.10}$$

and thus

$$[\iota_X,\iota_J] = \iota_{[X,J]^{\wedge}} = \iota_{\mathbf{c}}\,. \tag{38.1.11}$$

For the second relation, we use theorem 17.7.2, from which follows

$$[\iota_X,\mathcal{L}_J] = \mathcal{L}_{\iota_X J} - (-1)^{1\cdot 1}\iota_{[\![J,X]\!]} = \mathcal{L}_{\mathbf{c}} + \iota_{[\![J,X]\!]}\,. \tag{38.1.12}$$

∎

In the last statement, the Frölicher-Nijenhuis bracket $[\![J,X]\!] \in \Omega^1(TM,TTM)$ between the tangent structure and the semispray X appeared. This object has a number of interesting properties. Note that by definition it is a tensor field of rank $(1,1)$ on TM, and so it can be applied to vector fields. By theorem 17.6.7 it is given by

$$[\![J,X]\!] = -\mathcal{L}_X J\,. \tag{38.1.13}$$

This relation is helpful in proving the following identities.

Theorem 38.1.7. *For every semispray $X \in \mathrm{Vect}(TM)$, the Frölicher-Nijenhuis bracket $[\![J,X]\!]$ is an involution,*

$$[\![J,X]\!] \circ [\![J,X]\!] = \mathrm{id}_{TM}\,. \tag{38.1.14}$$

Proof. We defer this proof to theorem 38.3.1, where it will be shown that $[\![J, X]\!]$ is an almost product structure. Further, in theorem 37.2.3 it was shown that an almost product structure is an involution. ■

We may now use the fact that an involution on a vector space, and hence also on a vector bundle, always has eigenvalues ± 1. One may thus ask what are the eigenspaces corresponding to these eigenvalues. We find the following statement about the negative one.

Theorem 38.1.8. *For every semispray $X \in \text{Vect}(TM)$ and vector field $Y \in \text{Vect}(TM)$, the Frölicher-Nijenhuis bracket satisfies*

$$[\![J, X]\!]Y = -Y \tag{38.1.15}$$

if and only if Y is vertical.

Proof. If $[\![J, X]\!]Y = -Y$ holds, then

$$\begin{aligned} 0 &= J([\![J, X]\!]Y + Y) \\ &= JY - J[X, JY] + J^2[X, Y] \\ &= 2JY \end{aligned} \tag{38.1.16}$$

using Grifone's identity, and so Y is vertical. Conversely, if Y is vertical, it can be written as $Y = JZ$, and we find

$$JZ - [X, J^2Z] + J[X, JZ] = JZ + J[X, JZ] = 0\,, \tag{38.1.17}$$

once again using Grifone's identity. ■

We see that this eigenspace is independent of the choice of the semispray X. This is not the case for the positive eigenspace. We will explicitly construct it and discuss its properties in section 38.3.

Theorem 38.1.9. *For every semispray $X \in \text{Vect}(TM)$ and function $f \in C^\infty(M, \mathbb{R})$ the vertical and canonical lifts are related by*

$$X\overset{V}{f} = \overset{C}{f}\,. \tag{38.1.18}$$

Proof. Let $v \in TM$. Using theorem 38.1.1, which states that $\tau_* \circ X = \text{id}_{TM}$, as well as the definition 10.1.1 of the pushforward acting on a function, we have

$$\overset{C}{f}(v) = v(f) = \tau_*(X(v))(f) = X(v)(f \circ \tau) = X(v)\overset{V}{f} = \left(X\overset{V}{f}\right)(v)\,. \quad ■ \tag{38.1.19}$$

Theorem 38.1.10. *For every semispray $X \in \text{Vect}(TM)$ and vector field $Z \in \text{Vect}(M)$ the vertical and canonical lifts satisfy*

$$J\left[\overset{V}{Z}, X\right] = \overset{V}{Z}\,, \quad J\left[\overset{C}{Z}, X\right] = 0\,. \tag{38.1.20}$$

Proof. For the first equality, one can use the relation 29.5.4 between the vertical and complete lifts, as well as Grifone's identity 38.1.4, to show that

$$J\left[\overset{V}{Z}, X\right] = J\left[J\overset{C}{Z}, X\right] = J\overset{C}{Z} = \overset{V}{Z}\,. \tag{38.1.21}$$

For the second equality, we find

$$\begin{aligned} J\left[\overset{C}{Z}, X\right] &= J\left[\overset{C}{Z}, X\right] - \left[\overset{C}{Z}, \mathbf{c}\right] \\ &= J\left[\overset{C}{Z}, X\right] - \left[\overset{C}{Z}, JX\right] \\ &= -\left(\mathcal{L}_{\overset{C}{Z}}J\right)X \\ &= 0\,. \end{aligned} \tag{38.1.22}$$

Here we used theorem 29.5.3 that the complete lift of a vector field is 0-homogeneous to add a vanishing term in the first line, inserted the semispray via $JX = \mathbf{c}$ in the second line, used the Lie derivative 16.6.1 in the third line, and finally theorem 29.5.7 to show that the latter vanishes. ∎

38.2 Sprays

We now come to a class of semisprays, which are of particular interest. These are defined by demanding the following additional property.

Definition 38.2.1 (Spray)**.** A semispray $X \in \mathrm{Vect}(TM)$ on a manifold M is called a *spray* if and only if it is homogeneous of order 1.

Note that a spray cannot be homogeneous of any order $r \neq 1$. This follows from the defining property $JX = \mathbf{c}$ and the homogeneities of -1 for the tangent structure and 0 for the Liouville vector field, so that

$$\begin{aligned} 0 &= \mathcal{L}_{\mathbf{c}}\mathbf{c} \\ &= \mathcal{L}_{\mathbf{c}}(JX) \\ &= (\mathcal{L}_{\mathbf{c}}J)X + J(\mathcal{L}_{\mathbf{c}}X) \\ &= (r-1)JX\,, \end{aligned} \tag{38.2.1}$$

which can be true only for $JX = 0$, which would contradict the assumption $JX = \mathbf{c}$ that X is a semispray, or $r = 1$.

One also easily derives how the homogeneity condition restricts the coefficients of X in the induced coordinates. Writing $X = \bar{x}^a\partial_a + \bar{X}^a\bar{\partial}_a$, we have

$$\begin{aligned} \mathcal{L}_{\mathbf{c}}X &= [\mathbf{c}, X] \\ &= [\bar{x}^b\bar{\partial}_b, \bar{x}^a\partial_a + \bar{X}^a\bar{\partial}_a] \\ &= \bar{x}^b\delta^a_b\partial_a + \bar{x}^b\bar{\partial}_b\bar{X}^a\bar{\partial}_a - \bar{X}^a\delta^b_a\bar{\partial}_b \\ &= \bar{x}^a\partial_a + (\bar{x}^b\bar{\partial}_b\bar{X}^a - \bar{X}^a)\bar{\partial}_a\,. \end{aligned} \tag{38.2.2}$$

For a spray, this must again be equal to X, which is the case if and only if

$$\bar{x}^b\bar{\partial}_b\bar{X}^a = 2\bar{X}^a\,. \tag{38.2.3}$$

Hence, while the spray is a homogeneous vector field of order 1, the coefficients are homogeneous functions of order 2, due to the fact that the coordinate vector fields $\bar{\partial}_a$ are homogeneous of order -1.

One common possibility to measure by how much a semispray fails to be a spray is by defining the following vector field.

Definition 38.2.2 (Deviation)**.** Let $X \in \mathrm{Vect}\,TM$ be a semispray. Its *deviation* is the vector field

$$X^\star = [\mathbf{c}, X] - X\,. \tag{38.2.4}$$

Obviously we have $X^\star = 0$ if and only if X is 1-homogeneous, and hence a spray. Before we derive a coordinate expression, we prove the following statement.

Theorem 38.2.1. *The deviation of a semispray $X \in \mathrm{Vect}(TM)$ is vertical.*

Proof. We use theorem 38.1.5 that $[\mathbf{c}, X]$ is again a semispray, and hence $J[\mathbf{c}, X] = \mathbf{c}$, from which follows

$$J([\mathbf{c}, X] - X) = J[\mathbf{c}, X] - JX = \mathbf{c} - \mathbf{c} = 0\,. \tag{38.2.5}$$

Hence, the deviation lies in the kernel of J, and is thus vertical. ■

We finally calculate the deviation in coordinates. From the expression (38.2.2) we find

$$X^\star = (\bar{x}^b \bar{\partial}_b \bar{X}^a - 2\bar{X}^a)\bar{\partial}_a\,, \tag{38.2.6}$$

which is obviously vertical and vanishes when the homogeneity condition (38.2.3) is satisfied.

38.3 Non-linear connection induced by a semispray

As already mentioned at the beginning of this chapter, there exists a close relation between semisprays and non-linear connections on the tangent bundle, which we will now study in detail. First, we will see how every semispray induces a non-linear connection. Recall from section 37.2 that a non-linear connection can be described by various tensor fields on the tangent bundle TM. One of these tensor fields is the almost product structure given in definition 37.2.2. We can obtain an almost product structure from a semispray by virtue of the following statement.

Theorem 38.3.1. *Let X be a semispray and J the tangent structure. Then $\mathbb{P} = -\mathcal{L}_X J$ is an almost product structure.*

Proof. We first show that $J \circ \mathbb{P} = J$. For this purpose, we use theorem 16.6.1 for the Lie derivative to obtain

$$\mathbb{P}Y = -(\mathcal{L}_X J)Y = J[X, Y] - [X, JY] \tag{38.3.1}$$

for any $Y \in \mathrm{Vect}(TM)$. Applying J, the first term vanishes, since $J \circ J = 0$, and so we are left with

$$(J \circ \mathbb{P})Y = -J[X, JY] = JY\,, \tag{38.3.2}$$

using Grifone's identity 38.1.4. Since this holds for all $Y \in \mathrm{Vect}(TM)$, we find $J \circ \mathbb{P} = J$.

To show the opposite order $\mathbb{P} \circ J = -J$, observe that

$$(\mathbb{P} \circ J)Y = -(\mathcal{L}_X J)JY = J[X, JY] - [X, J^2Y] = J[X, JY] = -JY\,, \tag{38.3.3}$$

following the same arguments as above. Hence, we find that $\mathbb{P}$ indeed satisfies the properties of an almost product structure. ■

It is instructive to derive a coordinate expression for the almost product structure, and hence the connection coefficients, from a coordinate expression of the semispray. For this purpose, we will write the latter as $X = \bar{x}^a\partial_a + \bar{X}^a\bar{\partial}_a$. Together with the coordinate expression (29.2.6) of the tangent structure we then find the almost product structure

$$\begin{aligned}
\mathbb{P} &= -\mathcal{L}_X J \\
&= -\mathcal{L}_{\bar{x}^a\partial_a + \bar{X}^a\bar{\partial}_a}(\bar{\partial}_b \otimes \mathrm{d}x^b) \\
&= -[\bar{x}^a\partial_a + \bar{X}^a\bar{\partial}_a, \bar{\partial}_b] \otimes \mathrm{d}x^b - \bar{\partial}_b \otimes \mathrm{d}\iota_{\bar{x}^a\partial_a + \bar{X}^a\bar{\partial}_a}\mathrm{d}x^b \\
&= (\delta^a_b\partial_a + \bar{\partial}_b\bar{X}^a\bar{\partial}_a) \otimes \mathrm{d}x^b - \bar{\partial}_b \otimes \mathrm{d}\bar{x}^b \\
&= \bar{\partial}_b\bar{X}^a\bar{\partial}_a \otimes \mathrm{d}x^b + \partial_a \otimes \mathrm{d}x^a - \bar{\partial}_b \otimes \mathrm{d}\bar{x}^b
\end{aligned} \tag{38.3.4}$$

We can compare this with the coordinate expression (37.2.12) of the almost product structure, which reads, after expanding the Berwald basis,

$$\begin{aligned}\mathbb{P} &= (\partial_a - N^b{}_a\bar{\partial}_b) \otimes \mathrm{d}x^a - \bar{\partial}_a \otimes (\mathrm{d}\bar{x}^a + N^a{}_b\mathrm{d}x^b) \\ &= -2N^a{}_b\bar{\partial}_a \otimes \mathrm{d}x^b + \partial_a \otimes \mathrm{d}x^a - \bar{\partial}_b \otimes \mathrm{d}\bar{x}^b\,.\end{aligned} \tag{38.3.5}$$

We see that this is of the same form, and that the connection coefficients can be obtained from the coefficients of the semispray as

$$N^a{}_b = -\frac{1}{2}\bar{\partial}_b\bar{X}^a\,. \tag{38.3.6}$$

We also remark that the obtained non-linear connection is not arbitrary, but satisfies certain additional properties. In particular, we find the following relation.

Theorem 38.3.2. *The non-linear connection defined from a semispray X by the almost product structure $\mathbb{P} = -\mathcal{L}_X J$ is symmetric.*

Proof. For the proof we make use of the theory of graded derivations and the Frölicher-Nijenhuis bracket defined in section 17.6. Recall from definition 37.5.1 that the weak torsion is given by the Frölicher-Nijenhuis bracket $[\![\mathbf{h}, J]\!] = [\![J, \mathbf{h}]\!]$. Similarly, following theorem 17.6.7 we can write the almost product structure as $\mathbb{P} = -[\![X, J]\!] = [\![J, X]\!]$. This allows us to write the weak torsion as

$$\mathbf{t} = [\![J, \mathbf{h}]\!] = \frac{1}{2}[\![J, \mathrm{id}_{TTM} + \mathbb{P}]\!] = \frac{1}{2}[\![J, [\![J, X]\!]]\!]\,, \tag{38.3.7}$$

where the first part $[\![J, \mathrm{id}_{TTM}]\!] = 0$ vanishes following theorem 17.6.4. Using the graded Jacobi identity, we now have

$$[\![J, [\![J, X]\!]]\!] = [\![[\![J, J]\!], X]\!] - [\![J, [\![J, X]\!]]\!] = -[\![J, [\![J, X]\!]]\!]\,, \tag{38.3.8}$$

where the first term vanishes due to the vanishing of the Nijenhuis tensor $N_J = 0$ of the tangent structure. Hence, we find $\mathbf{t} = 0$, so that the induced connection is symmetric. ■

This can also be seen very easily from the coordinate expressions. Using the relation (37.5.6) for a non-linear connection one has

$$\bar{\partial}_{[a}N^c{}_{b]} = -\frac{1}{2}\bar{\partial}_{[a}\bar{\partial}_{b]}\bar{X}^c = 0\,, \tag{38.3.9}$$

which vanishes, since the appearing vertical derivatives commute.

38.4 Semispray induced by a non-linear connection

So far we have thus constructed a (symmetric) non-linear connection from a semispray. We now follow the inverse direction and construct a semispray from a non-linear connection. The following statement shows that this is indeed possible.

Theorem 38.4.1. *Let M be a manifold equipped with a non-linear connection on the tangent bundle. Then $\tilde{X} = \Theta\mathbf{c}$ is a semispray, where Θ is the adjoint structure and $\mathbf{c}$ is the Liouville vector field.*

Proof. This follows from the fact that the Liouville vector field is vertical, and hence

$$J\tilde{X} = (J \circ \Theta)\mathbf{c} = \mathbf{vc} = \mathbf{c}\,. \quad ■ \tag{38.4.1}$$

Again it is instructive to derive a coordinate expression for the semispray X from that of the adjoint structure. Using the relations (19.9.2) for the Liouville vector field and (37.2.5) for the adjoint structure in the Berwald basis follows

$$\tilde{X} = \Theta\mathbf{c} = (\delta_a \otimes \delta\bar{x}^a)y^b\bar{\partial}_b = \bar{x}^a\delta_a = \bar{x}^a(\partial_a - N^b{}_a\bar{\partial}_b)\,. \tag{38.4.2}$$

This obviously satisfies the condition (38.1.1) for a semispray, and we see that its coefficients are given by $\bar{\tilde{X}}^a = -\bar{x}^b N^a{}_b$.

The semispray induced by a non-linear connection has several interesting properties, which we will prove in the following. The first properly allows for an alternative characterization, and is stated as follows.

Theorem 38.4.2. *Let M be a manifold equipped with a non-linear connection on the tangent bundle. Then $\tilde{X} = \Theta\mathbf{c}$ is the unique horizontal semispray, and it satisfies*

$$\tilde{X} = \mathbf{h}[\mathbf{c}, \tilde{X}]\,. \tag{38.4.3}$$

Proof. For the first part of the proposition, let X be a horizontal semispray. Then we have

$$X = \mathbf{h}X = \Theta J X = \Theta\mathbf{c} = \tilde{X}\,, \tag{38.4.4}$$

and so $\tilde{X} = X$ is unique.

For the second statement, recall from theorem 38.1.5 that for any semispray $\tilde{X}$, also $[\mathbf{c}, \tilde{X}]$ is a semispray. Applying the first part of the proposition, where we have shown that $\mathbf{h}X = \tilde{X}$ for any semispray X, to the semispray $X = [\mathbf{c}, \tilde{X}]$, we find that its horizontal part is again the unique horizontal semispray $\tilde{X}$. ■

Also this can easily be seen from the coordinate expressions. From the Berwald basis one sees immediately that $\tilde{X} = \bar{x}^a\delta_a$ is horizontal. Further, any semispray $X = \bar{x}^a\partial_a + \bar{X}^a\bar{\partial}_a$ which is horizontal must satisfy

$$0 = \mathbf{v}X = (\bar{X}^a + \bar{x}^b N^a{}_b)\bar{\partial}_a\,, \tag{38.4.5}$$

and so its coefficients $\bar{X}^a = -\bar{x}^b N^a{}_b$ are uniquely determined. Finally, one may directly calculate

$$\begin{aligned}
\mathbf{h}[\mathbf{c}, \tilde{X}] &= \mathbf{h}[\bar{x}^c\bar{\partial}_c, \bar{x}^a(\partial_a - N^b{}_a\bar{\partial}_b)]\\
&= \mathbf{h}(\bar{x}^a\delta_a - \bar{x}^a\bar{x}^c\bar{\partial}_c N^b{}_a\bar{\partial}_b + \bar{x}^a N^b{}_a\bar{\partial}_b)\\
&= \bar{x}^a\delta_a\\
&= \tilde{X}\,.
\end{aligned} \tag{38.4.6}$$

38.5 Relation between mutual inductions

Since now we have two constructions for obtaining a non-linear connection from a semispray and vice versa, one may pose the question how these constructions are related to each other. The first step towards the answer to this question is the following statement.

Theorem 38.5.1. *Let X be a semispray, which induces a non-linear connection with almost product structure $\mathbb{P} = -\mathcal{L}_X J$ and adjoint structure Θ. Then the semispray $\tilde{X} = \Theta\mathbf{c}$ is given by*

$$\tilde{X} = \frac{1}{2}(X + [\mathbf{c}, X])\,. \tag{38.5.1}$$

Proof. Following theorem 38.4.2, $\tilde{X}$ is the horizontal part of X, and so

$$\begin{aligned}
\tilde{X} &= \Theta\mathbf{c} \\
&= \Theta JX \\
&= \mathbf{h}X \\
&= \frac{1}{2}(X + \mathbb{P}X) \\
&= \frac{1}{2}(X + J[X,X] - [X,JX]) \\
&= \frac{1}{2}(X + [\mathbf{c},X]) . \quad \blacksquare
\end{aligned} \tag{38.5.2}$$

Also this construction can nicely be illustrated using coordinates. Writing the semisprays as $X = \bar{x}^a\partial_a + \bar{X}^a\bar{\partial}_a$ and $\tilde{X} = \bar{x}^a\partial_a + \bar{\tilde{X}}^a\bar{\partial}_a$, we find from the coordinate expressions of the induced non-linear connection and of the induced semispray the relation

$$\bar{\tilde{X}}^a = -\bar{x}^b N^a{}_b = \frac{1}{2}\bar{x}^b\bar{\partial}_b\bar{X}^a . \tag{38.5.3}$$

Using theorem 38.5.1, we find the expression

$$\begin{aligned}
\tilde{X} &= \frac{1}{2}(X + [\mathbf{c},X]) \\
&= \frac{1}{2}\left(\bar{x}^a\partial_a + \bar{X}^a\bar{\partial}_a + [\bar{x}^b\bar{\partial}_b, \bar{x}^a\partial_a + \bar{X}^a\bar{\partial}_a]\right) \\
&= \frac{1}{2}\left(\bar{x}^a\partial_a + \bar{X}^a\bar{\partial}_a + \bar{x}^b\delta^a_b\partial_a + \bar{x}^b\bar{\partial}_b\bar{X}^a\bar{\partial}_a - \bar{X}^a\delta^b_a\bar{\partial}_b\right) \\
&= \bar{x}^a\partial_a + \frac{1}{2}\bar{x}^b\bar{\partial}_b\bar{X}^a\bar{\partial}_a ,
\end{aligned} \tag{38.5.4}$$

which obviously agrees with the result obtained by direct calculation.

One may wonder under which circumstances one obtains the same semispray, $X = \tilde{X}$, and how this is related to the properties of the connection. Also this question has a simple answer, as we shall see below.

Theorem 38.5.2. *Let X be a semispray. The following conditions are equivalent:*

1. *X is a spray.*
2. *X is horizontal with respect to its non-linear connection induced via the almost product structure $\mathbb{P} = -\mathcal{L}_X J$.*
3. *The semispray $\tilde{X} = \Theta\mathbf{c}$ obtained from the induced non-linear connection is equal to X.*

Further, from any of these conditions follows that the induced non-linear connection is homogeneous.

Proof. We prove the equivalence of the first three statements above using the following steps:

1. Recall that the horizontal projector is given by $\mathbf{h} = \frac{1}{2}(\mathbb{P} + \mathrm{id}_{TTM})$. Using theorem 38.3.1, as well as $JX = \mathbf{c}$ for a semispray, we have

$$\begin{aligned}
X - \mathbf{h}X &= \frac{1}{2}(X - \mathbb{P}X) \\
&= \frac{1}{2}(X - J[X,X] + [X,JX]) \\
&= \frac{1}{2}(X - [\mathbf{c},X]) \\
&= \frac{1}{2}(X - \mathcal{L}_{\mathbf{c}}X) ,
\end{aligned} \tag{38.5.5}$$

so that $X = \mathbf{h}X$ if and only if X is 1-homogeneous, i.e., a spray.

2. From the relation (38.5.1) we find

$$\tilde{X} - X = \frac{1}{2}([\mathbf{c}, X] - X) = \frac{1}{2}(\mathcal{L}_{\mathbf{c}}X - X)\,, \tag{38.5.6}$$

and so $\tilde{X} = X$ if and only if X is 1-homogeneous, i.e., a spray.

Finally, if X is a spray, and hence 1-homogeneous, it follows from theorem 19.9.6 that $\mathbb{P}$ is 0-homogeneous, so that the connection is homogeneous. ■

Using coordinates, the statement of the theorem becomes obvious. We see that the three equivalent conditions are expressed in coordinates as follows:

1. X is a spray, and hence 1-homogeneous:

$$\bar{x}^b\bar{\partial}_b\bar{X}^a = 2\bar{X}^a\,. \tag{38.5.7}$$

2. X is horizontal with respect to its induced non-linear connection:

$$0 = \mathbf{v}X = (\bar{X}^a + \bar{x}^b N^a{}_b)\bar{\partial}_a = \left(\bar{X}^a - \frac{1}{2}\bar{x}^b\bar{\partial}_b\bar{X}^a\right)\bar{\partial}_a\,. \tag{38.5.8}$$

3. The semispray $\tilde{X} = \Theta\mathbf{c}$ obtained from the induced non-linear connection is equal to X:

$$\bar{X}^a = \bar{\tilde{X}}^a = \frac{1}{2}\bar{x}^b\bar{\partial}_b\bar{X}^a\,. \tag{38.5.9}$$

This is obviously the same condition on the spray coefficients $\bar{X}^a$ in all three cases. Finally, the homogeneity of the induced non-linear connection follows from any of these, since in this case

$$\begin{aligned}\bar{x}^c\bar{\partial}_c N^a{}_b &= \bar{x}^c\bar{\partial}_c\bar{\partial}_b\bar{X}^a\\ &= \bar{\partial}_b(\bar{x}^c\bar{\partial}_c\bar{X}^a) - (\bar{\partial}_b\bar{x}^c)\bar{\partial}_c\bar{X}^a\\ &= 2\bar{\partial}_b\bar{X}^a - \bar{\partial}_b\bar{X}^a\\ &= N^a{}_b\,,\end{aligned} \tag{38.5.10}$$

which is simply the relation (37.4.5) for a homogeneous non-linear connection.

Note that the last property, the homogeneity follows from the other, equivalent properties, but the converse is not true: a homogeneous connection does not imply that X is a spray. To see this, let X be a spray, so that $\mathbb{P} = -\mathcal{L}_X J$ is the almost product structure of a homogeneous connection, and $Z \in \operatorname{Vect}(M)$. Then we have that $\overset{\mathrm{v}}{Z}$ is a -1-homogeneous vertical vector field satisfying $J\overset{\mathrm{v}}{Z} = 0$ and $\mathcal{L}_{\overset{\mathrm{v}}{Z}}J = 0$, following theorems 29.5.1 and 29.5.6. Hence, $X' = X + \overset{\mathrm{v}}{Z}$ is *not* 1-homogeneous, so that it is only a semispray, since it still satisfies $JX' = \mathbf{c}$, but $-\mathcal{L}_{X'}J = -\mathcal{L}_X J$ defines the same homogeneous connection as induced by X.

The latter can also be seen in coordinates. Let $X = \bar{x}^a\partial_a + \bar{X}^a\bar{\partial}_a$ be a spray and $Z = Z^a\partial_a \in \operatorname{Vect}(M)$ a vector field on M. Its vertical lift is given by $\overset{\mathrm{v}}{Z} = Z^a\bar{\partial}_a$. Hence, also

$$X' = X + \overset{\mathrm{v}}{Z} = \bar{x}^a\partial_a + (\bar{X}^a + Z^a)\bar{\partial}_a \tag{38.5.11}$$

is a semispray, but not a spray. Still the induced non-linear connection has coefficients

$$N^a{}_b = -\frac{1}{2}\bar{\partial}_b(\bar{X}^a + Z^a) = -\frac{1}{2}\bar{\partial}_b\bar{X}^a\,, \tag{38.5.12}$$

since $\bar{\partial}_b Z^a = 0$, and so is the same as that induced by X, and thus homogeneous.

Finally, we also consider the case in which we start from a non-linear connection with coefficients $N^a{}_b$, construct its induced semispray $\tilde{X}$, and from the latter again the induced connection, whose coefficients we denote by $\tilde{N}^a{}_b$. Again one may pose the question under which conditions the resulting non-linear connection agrees with the original one. The answer is given by the following theorem.

Theorem 38.5.3. *Let M be a manifold equipped with a non-linear connection on the tangent bundle, $\tilde{X} = \Theta\mathbf{c}$ the induced semispray, and $\tilde{\mathbb{P}} = -\mathcal{L}_{\tilde{X}}J$ the almost product structure of the non-linear connection induced by $\tilde{X}$. Then the difference of the corresponding almost product structures is given by the strong torsion of the original connection,*

$$\mathbb{P} - \tilde{\mathbb{P}} = -\mathbf{T}\,. \tag{38.5.13}$$

Proof. Using horizontal and vertical projectors, we have for any vector field $X \in \mathrm{Vect}(TM)$ the almost product structures

$$\mathbb{P}X = \mathbf{h}X - \mathbf{v}X \tag{38.5.14}$$

and

$$\tilde{\mathbb{P}}X = -(\mathcal{L}_{\Theta\mathbf{c}}J)X = J[\Theta\mathbf{c}, X] - [\Theta\mathbf{c}, JX]\,. \tag{38.5.15}$$

Together with the formula from theorem 37.5.2 we then have

$$\mathbb{P}X - \tilde{\mathbb{P}}X + \mathbf{T}X = \mathbf{h}X - \mathbf{v}X - J[\Theta\mathbf{c}, \mathbf{v}X] + \mathbf{h}[\Theta\mathbf{c}, JX]\,. \tag{38.5.16}$$

Note that the first and the last term are horizontal and vanish if X is vertical, while the remaining two terms are vertical and vanish if X is horizontal. It thus suffices to check them individually. Let first X be vertical and write $X = JY$ for some vector field $Y \in \mathrm{Vect}(TM)$. Then

$$JY + J[\Theta\mathbf{c}, JY] = 0 \tag{38.5.17}$$

due to Grifone's identity, since $\Theta\mathbf{c}$ is a semispray. This shows that the vertical terms cancel. For the horizontal terms, we apply J and find

$$JX + J[\Theta\mathbf{c}, JX] = 0\,, \tag{38.5.18}$$

which vanishes for the same reasons as above. Hence, also the horizontal terms cancel. ■

Also this can be seen easily in coordinates. Using the previously derived coordinate expressions we have

$$\tilde{N}^a{}_b = -\frac{1}{2}\bar{\partial}_b\bar{\tilde{X}}^a = \frac{1}{2}\bar{\partial}_b(\bar{x}^c N^a{}_c) = \frac{1}{2}(N^a{}_b + \bar{x}^c\bar{\partial}_b N^a{}_c)\,. \tag{38.5.19}$$

The almost product structures of the non-linear connections are thus given by

$$\mathbb{P} = \delta_a \otimes \mathrm{d}x^a - \bar{\partial}_a \otimes \delta\bar{x}^a = (\partial_a - N^b{}_a\bar{\partial}_b) \otimes \mathrm{d}x^a - \bar{\partial}_a \otimes (\mathrm{d}\bar{x}^a + N^a{}_b\mathrm{d}x^b) \tag{38.5.20}$$

and analogously for $\tilde{\mathbb{P}}$. For their difference we then find

$$\mathbb{P} - \tilde{\mathbb{P}} = 2(\tilde{N}^a{}_b - N^a{}_b)\bar{\partial}_a \otimes \mathrm{d}x^b = (\bar{x}^c\bar{\partial}_b N^a{}_c - N^a{}_b)\bar{\partial}_a \otimes \mathrm{d}x^b = -\mathbf{T}\,. \tag{38.5.21}$$

This relation is useful to characterize a non-linear connection by its properties. We use it to show the following.

Theorem 38.5.4. *Let M be a manifold equipped with a non-linear connection on the tangent bundle, $\tilde{X} = \Theta\mathbf{c}$ the induced semispray, and $\tilde{\mathbb{P}} = -\mathcal{L}_{\tilde{X}}J$ the almost product structure of the non-linear connection induced by $\tilde{X}$. Then the induced non-linear connection agrees with the original one if and only if the latter is symmetric and homogeneous.*

Proof. Assume first that the connections agree, $\mathbb{P} = \tilde{\mathbb{P}}$. From theorem 38.3.2 follows that the induced non-linear connection defined by $\tilde{\mathbb{P}}$ is symmetric, and so it can agree with the original connection only if the latter is symmetric as well. Further, theorem 38.4.2 states that $\tilde{X}$ is horizontal with respect to $\mathbb{P}$. If $\mathbb{P} = \tilde{\mathbb{P}}$, then it is also horizontal with respect to $\tilde{\mathbb{P}}$. From theorem 38.5.2 then follows that the induced connection defined by $\tilde{\mathbb{P}}$, and hence the original connection defined by $\mathbb{P}$, is homogeneous.

Conversely, if the original connection is homogeneous and symmetric, then $\mathbb{T} = 0$ and $\mathbf{t} = 0$, whence $\mathbf{T} = 0$ and thus $\mathbb{P} = \tilde{\mathbb{P}}$ according to theorem 38.5.3.

Note that this proves that $\mathbf{T} = 0$ if and only if $\mathbb{T} = 0$ and $\mathbf{t} = 0$, as claimed in theorem 37.5.3. ∎

We can also see this from the coordinate expressions we derived above. Already the first equality in (38.5.19) shows that the induced connection is symmetric, $\bar{\partial}_{[a}\tilde{N}^c{}_{b]} = 0$, as shown in theorem 38.3.2. A necessary condition for the equality of both connections, $\tilde{N}^a{}_b = N^a{}_b$, is therefore that also the original connection is symmetric, $\bar{\partial}_{[a}N^c{}_{b]} = 0$. Imposing this condition, we have

$$\tilde{N}^a{}_b = \frac{1}{2}(N^a{}_b + \bar{x}^c\bar{\partial}_b N^a{}_c) . = \frac{1}{2}(N^a{}_b + \bar{x}^c\bar{\partial}_c N^a{}_b) , \tag{38.5.22}$$

and so both connections are equal if and only if in addition holds

$$N^a{}_b = \bar{x}^c\bar{\partial}_c N^a{}_b , \tag{38.5.23}$$

which means that the original (and hence also the induced) connection is homogeneous.

The previous theorem shows that a homogeneous and symmetric connection is fully determined by its induced semispray. This can be generalized to the case of an arbitrary non-linear connection as follows.

Theorem 38.5.5. *A non-linear connection is uniquely determined by its induced semispray* $\Theta\mathbf{c}$ *and strong torsion* $\mathbf{T}$.

Proof. Let $\mathbb{P}_1$ and $\mathbb{P}_2$ be the almost product structures of two non-linear connections, and assume that they have the same strong torsion $\mathbf{T}_1 = \mathbf{T}_2 = \mathbf{T}$ and the same induced semispray, which hence gives the same induced non-linear connection $\tilde{\mathbb{P}}_1 = \tilde{\mathbb{P}}_2 = \tilde{\mathbb{P}}$. From theorem 38.5.3 then follows

$$\mathbb{P}_1 = \tilde{\mathbb{P}} - \mathbf{T} = \mathbb{P}_2 . \tag{38.5.24}$$

∎

38.6 Semisprays and autoparallel curves

The previous sections have shown that there exists a close relationship between semisprays and non-linear connections. This relationship can be used in order to express concepts from one of these notions to the other. Here we apply this relationship to autoparallel curves, where we find the following useful relation.

Theorem 38.6.1. *Let* M *be a manifold equipped with a non-linear connection. A curve* $\gamma : \mathbb{R} \to M$ *is autoparallel if and only if its canonical lift is an integral curve of the induced semispray.*

Proof. Let Θ be the adjoint structure of a non-linear connection and $\Theta\mathbf{c}$ the induced semispray. We proceed in two steps.

1. According to theorem 38.1.3, every integral curve of a semispray is the canonical lift of a curve on M. This curve is horizontal, since $\Theta\mathbf{c}$ is horizontal. Following definition 37.7.1, a curve $\gamma : \mathbb{R} \to M$ is an autoparallel if and only if its canonical lift $\dot{\gamma} : \mathbb{R} \to TM$ is horizontal. Hence, every integral curve of $\Theta\mathbf{c}$ is the canonical lift of an autoparallel curve.

2. Let $\gamma : \mathbb{R} \to M$ be an autoparallel curve, and so its canonical lift $\dot{\gamma} : \mathbb{R} \to TM$ is horizontal, i.e., it is a horizontal lift of γ. ►...◄

■

Chapter 39

D-tensors and d-connections

39.1 Pullback formalism

[Ant03]

39.2 D-tensors

[MA94]

39.3 D-connections and N-linear connections

Following the definitions in section 39.2, we can express d-tensors as tensors on the tangent bundle of a manifold equipped with a non-linear connection, which respect the split of the double tangent bundle into horizontal and vertical parts. Given such tensor fields, one is naturally interested in taking their derivatives with respect to a linear connection, hence in this case a Koszul connection on the bundle $\varpi : TTM \to M$. However, such a connection will, in general, not preserve the split $TTM \cong HTM \otimes VTM$ into horizontal and vertical bundles. The class of connections which has this property may be considered *distinguished*, in the same sense as d-tensors. We define them as follows.

Definition 39.3.1 (D-connection)**.** Let M be a manifold with a non-linear connection. A Koszul connection ∇ on the double tangent bundle $\varpi : TTM \to TM$ is called a *d-connection* (or *distinguished connection*) if and only if $\nabla\mathbb{P} = 0$, where $\mathbb{P}$ is the almost product structure of the non-linear connection.

Here we have chosen to use the almost product structure in the definition. However, we could just as well have used any of the projectors, as the following statement shows.

Theorem 39.3.1. *A Koszul connection on the double tangent bundle $\varpi : TTM \to TM$ is a d-connection if and only if it preserves the horizontal projector, $\nabla\mathbf{h} = 0$ (and hence, equivalently, also the vertical projector, $\nabla\mathbf{v} = 0$).*

Proof. For any Koszul connection and vector field X one has

$$(\nabla\mathrm{id}_{TTM})X = \nabla(\mathrm{id}_{TTM}X) - \mathrm{id}_{TTM}\nabla X = \nabla X - \nabla X = 0\,. \tag{39.3.1}$$

Since the almost product structure, the horizontal and the vertical projector only differ up to the identity id_{TTM} and constant factors, any of them is preserved by the connection if and only if any other is. ■

With this theorem in place, it is now easy to show that a d-connection indeed preserves the horizontal and vertical distributions. We state this property as follows.

Theorem 39.3.2. *A d-connection ∇ preserves the horizontal and vertical parts of a vector field,*

$$\nabla_X Y = \mathbf{h}\nabla_X \mathbf{h}Y + \mathbf{v}\nabla_X \mathbf{v}Y \tag{39.3.2}$$

for all $X, Y \in \mathrm{Vect}(TM)$.

Proof. Since $\nabla\mathbf{h} = \nabla\mathbf{v} = 0$ for a d-connection, we have

$$\mathbf{h}\nabla_X \mathbf{h}Y + \mathbf{v}\nabla_X \mathbf{v}Y = \mathbf{h}^2\nabla_X Y + \mathbf{v}^2\nabla_X Y = \mathbf{h}\nabla_X Y + \mathbf{v}\nabla_X Y = \nabla_X Y\,. \tag{39.3.3}$$

■

As discussed in section 39.1, we can regard d-tensors also as tensors over a particular pullback bundles, and find isomorphisms between horizontal and vertical bundles. In order to also preserve these isomorphism, we need to impose an additional condition on the d-connection, which leads us to the following definition.

Definition 39.3.2 (N-linear connection)**.** A d-connection ∇ on a manifold M is called a N-linear connection if and only if $\nabla J = 0$, where J is the tangent structure.

One may now expect that also the other bundle isomorphisms we constructed are preserved by a N-linear connection. The following statement shows that this expectation is justified.

Theorem 39.3.3. *A d-connection ∇ is a N-linear connection if and only if $\nabla\Theta = 0$, or equivalently $\nabla\mathbb{F} = 0$.*

Proof. Since $\mathbb{F} = \Theta - J$, it is sufficient to consider only one of these objects, and so we will prove it for Θ. Given a N-linear connection ∇ and a vector field $X, Y \in \mathrm{Vect}(TM)$, we have

$$\begin{aligned}
(\nabla_X\Theta)Y &= \nabla_X\Theta Y - \Theta\nabla_X Y\\
&= \mathbf{h}\nabla_X\Theta Y - \Theta\nabla_X Y\\
&= \Theta J\nabla_X\Theta Y - \Theta\nabla_X Y\\
&= \Theta\nabla_X J\Theta Y - \Theta\nabla_X Y\\
&= \Theta\nabla_X \mathbf{v}Y - \Theta\nabla_X Y\\
&= \Theta\mathbf{v}\nabla_X Y - \Theta\nabla_X Y\\
&= \Theta\nabla_X Y - \Theta\nabla_X Y\\
&= 0\,.
\end{aligned} \tag{39.3.4}$$

Conversely, let ∇ be a d-connection such that $\nabla\Theta = 0$. Then we have

$$\begin{aligned}
(\nabla_X J)Y &= \nabla_X JY - J\nabla_X Y\\
&= \mathbf{v}\nabla_X JY - J\nabla_X Y\\
&= J\Theta\nabla_X JY - J\nabla_X Y\\
&= J\nabla_X\Theta JY - J\nabla_X Y\\
&= J\nabla_X \mathbf{h}Y - J\nabla_X Y\\
&= J\mathbf{h}\nabla_X Y - J\nabla_X Y\\
&= J\nabla_X Y - J\nabla_X Y\\
&= 0\,.
\end{aligned} \tag{39.3.5}$$

Since this holds for all vector fields X, Y, both conditions are equivalent. ■

For practical calculations, it is useful to study also the component expressions of a N-linear connection. Since any d-connection preserves the horizontal and vertical distributions, it is most convenient to work in the Berwald basis, and it follows immediately that the covariant derivatives of the basis vectors with respect to a d-connection can be expressed in the form

$$\nabla_{\delta_a}\delta_b = F^c{}_{ba}\delta_c\,, \qquad \nabla_{\delta_a}\bar{\partial}_b = \bar{F}^c{}_{ba}\bar{\partial}_c\,, \tag{39.3.6a}$$
$$\nabla_{\bar{\partial}_a}\delta_b = C^c{}_{ba}\delta_c\,, \qquad \nabla_{\bar{\partial}_a}\bar{\partial}_b = \bar{C}^c{}_{ba}\bar{\partial}_c\,, \tag{39.3.6b}$$

since the covariant derivative of a horizontal (vertical) vector field must again be horizontal (vertical). For a N-linear connection, in addition also the tangent structure must be preserved. This yields the further conditions

$$0 = J\nabla_{\delta_a}\delta_b - \nabla_{\delta_a}J\delta_b = F^c{}_{ba}J\delta_c - \nabla_{\delta_a}\bar{\partial}_b = F^c{}_{ba}\bar{\partial}_c - \bar{F}^c{}_{ba}\bar{\partial}_c\,, \tag{39.3.7a}$$
$$0 = J\nabla_{\bar{\partial}_a}\delta_b - \nabla_{\bar{\partial}_a}J\delta_b = C^c{}_{ba}J\delta_c - \nabla_{\bar{\partial}_a}\bar{\partial}_b = C^c{}_{ba}\bar{\partial}_c - \bar{C}^c{}_{ba}\bar{\partial}_c\,, \tag{39.3.7b}$$

so that for a N-linear connection we have $F^c{}_{ba} = \bar{F}^c{}_{ba}$ and $C^c{}_{ba} = \bar{C}^c{}_{ba}$. Hence, in a given coordinate system $(x^a, \bar{x}^a)$, a N-linear connection is fully determined by its coefficients $F^c{}_{ba}(x, \bar{x})$ and $C^c{}_{ba}(x, \bar{x})$, in addition to the coefficients $N^b{}_a(x, \bar{x})$, which determine the non-linear connection and hence the Berwald basis.

Given a d-tensor Q, its covariant derivative ∇Q with respect to a N-linear connection will still not be a d-tensor field, since ∇ introduces both horizontal and vertical components for the additional tensor factor. In order to obtain a d-tensor field, one must thus project onto either of the corresponding subspaces. This can be achieved by the following definition.

Definition 39.3.3 (Horizontal and vertical part of a N-linear connection)**.** Let ∇ be a N-linear connection on a manifold M. Its *horizontal and vertical part* ∇^h and ∇^v are defined such that

$$\nabla^h_X Y = \nabla_{\mathbf{h}X} Y\,, \quad \nabla^v_X Y = \nabla_{\mathbf{v}X} Y \tag{39.3.8}$$

for all vector fields $X, Y \in \mathrm{Vect}(TM)$.

Note that ∇^h and ∇^v by themselves are *not* Koszul connections, since

$$\nabla^h_X f = \mathbf{h}Xf \neq Xf\,, \quad \nabla^v_X f = \mathbf{v}Xf \neq Xf \tag{39.3.9}$$

for general vector fields X and functions f. Only their sum is a Koszul connection.

39.4 Torsion

As for any linear connection on a tangent bundle, one can define the torsion as detailed in section 30.4. Making the necessary replacement $M \to TM$ in definition 30.4.2, we arrive at the following definition.

Definition 39.4.1 (Torsion of a N-linear connection)**.** Let ∇ be a N-linear connection on a manifold M. Its *torsion* is the vector-valued two-form $T \in \Omega^2(TM, TTM)$ defined by

$$T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y] \tag{39.4.1}$$

for all vector fields $X, Y \in \mathrm{Vect}(TM)$.

It is instructive to decompose the torsion into its horizontal and vertical parts, i.e., first decompose the vector fields X and Y into their horizontal and vertical parts, and then decompose $T(X, Y)$. Using the fact that a d-connection preserves horizontal and vertical parts of vector fields according to theorem 39.3.2, and that the commutator of vertical vector fields is again vertical, we find the following expressions for the components:

1. The (h)h-torsion:
$$\mathbf{h}T(\mathbf{h}X, \mathbf{h}Y) = \nabla^h_X \mathbf{h}Y - \nabla^h_Y \mathbf{h}X - \mathbf{h}[\mathbf{h}X, \mathbf{h}Y] \,. \tag{39.4.2a}$$
2. The (v)h-torsion:
$$\mathbf{v}T(\mathbf{h}X, \mathbf{h}Y) = -\mathbf{v}[\mathbf{h}X, \mathbf{h}Y] = \mathbf{R}(X, Y) \,. \tag{39.4.2b}$$
3. The (h)hv-torsion:
$$\mathbf{h}T(\mathbf{h}X, \mathbf{v}Y) = -\nabla^v_Y \mathbf{h}X - \mathbf{h}[\mathbf{h}X, \mathbf{v}Y] \,. \tag{39.4.2c}$$
4. The (v)hv-torsion:
$$\mathbf{v}T(\mathbf{h}X, \mathbf{v}Y) = \nabla^h_X \mathbf{v}Y - \mathbf{v}[\mathbf{h}X, \mathbf{v}Y] \,. \tag{39.4.2d}$$
5. The (h)v-torsion:
$$\mathbf{h}T(\mathbf{v}X, \mathbf{v}Y) = 0 \,. \tag{39.4.2e}$$
6. The (v)v-torsion:
$$\mathbf{v}T(\mathbf{v}X, \mathbf{v}Y) = \nabla^v_X \mathbf{v}Y - \nabla^v_Y \mathbf{v}X - \mathbf{v}[\mathbf{v}X, \mathbf{v}Y] \,. \tag{39.4.2f}$$

In the case of a general affine connection discussed in section 30.4 we have called a connection *symmetric* if and only if its torsion vanishes. However, in the case of N-linear connections, this condition turns out to be too restrictive. Depending on the choice of the non-linear connection, there may not exist any N-linear connection with vanishing torsion. This can be seen from the decomposition shown above, since the component $\mathbf{v}T(\mathbf{h}X, \mathbf{h}Y) = \mathbf{R}(X, Y)$ is fully determined by the curvature of the non-linear connection. Hence, if the non-linear connection has non-vanishing curvature, any N-linear connection must have non-vanishing torsion. For a N-linear connection, one therefore uses a weaker condition to define a notion of being symmetric, which restricts only independent components of the connection. This is defined as follows.

Definition 39.4.2 (Symmetric N-linear connection)**.** A N-linear connection is called *symmetric* if and only if $\mathbf{h}T(\mathbf{h}X, \mathbf{h}Y) = \mathbf{v}T(\mathbf{v}X, \mathbf{v}Y) = 0$ for all $X, Y \in \mathrm{Vect}(TM)$.

Using the expression (39.3.6) for the N-linear connection coefficients, as well as the Berwald basis (37.1.1), one can calculate the components of the torsion in the Berwald basis. One finds that these are given by

$$\mathbf{h}T(\delta_a, \delta_b) = (F^c{}_{ba} - F^c{}_{ab})\delta_c \,, \tag{39.4.3a}$$
$$\mathbf{v}T(\delta_a, \delta_b) = (\delta_a N^c{}_b - \delta_b N^c{}_a)\bar{\partial}_c \,, \tag{39.4.3b}$$
$$\mathbf{h}T(\bar{\partial}_a, \delta_b) = C^c{}_{ba}\delta_c \,, \tag{39.4.3c}$$
$$\mathbf{v}T(\bar{\partial}_a, \delta_b) = (\bar{\partial}_a N^c{}_b - F^c{}_{ab})\bar{\partial}_c \,, \tag{39.4.3d}$$
$$\mathbf{v}T(\bar{\partial}_a, \bar{\partial}_b) = (C^c{}_{ba} - C^c{}_{ab})\bar{\partial}_c \,. \tag{39.4.3e}$$

We thus see that for a symmetric N-linear connection, as given in definition 39.4.2, the connection coefficients are symmetric in their lower indices, $F^a{}_{[bc]} = C^a{}_{[bc]} = 0$, in analogy to the symmetry of the connection coefficients for an affine connection shown in section 30.4. Also note that the coefficients $F^a{}_{bc}$ and $C^a{}_{bc}$ are fully determined by the non-linear connection coefficients $N^a{}_b$, as well as the mixed torsion components $T(\mathbf{h}X, \mathbf{v}Y)$.

39.5 Curvature

Since a N-connection is a connection, we can define its curvature, thus specializing the notions given in sections 26.10, 28.12 and in particular 30.3. From the latter we simply replace $M \to TM$ to arrive at the following definition.

Definition 39.5.1 (Curvature of a N-linear connection)**.** Let ∇ be a N-linear connection on a manifold M. Its *curvature* is the endomorphism-valued two-form $R \in \Omega^2(TM, \mathrm{End}(TTM))$ defined by

$$R(X,Y)Z = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z \tag{39.5.1}$$

for all vector fields $X, Y, Z \in \mathrm{Vect}(TM)$.

We can now in principle proceed as for the torsion shown in section 39.4 and decompose the curvature into its horizontal and vertical components. Since in the definition 39.5.1 three vector fields appear, and the result is again a vector field, decomposing each of them would yield 16 components, and 12 if we take into account the antisymmetry of the curvature in its two form arguments. However, it turns out that this task can be simplified significantly if we take into account the properties of N-linear connections. First, we use the fact that a d-connection preserves the horizontal and vertical to show the following.

Theorem 39.5.1. *The curvature R of a N-linear connection preserves the horizontal and vertical distributions,*

$$R(X,Y)Z = \mathbf{h}R(X,Y)\mathbf{h}Z + \mathbf{v}R(X,Y)\mathbf{v}Z \tag{39.5.2}$$

for all vector fields $X, Y, Z \in \mathrm{Vect}(TM)$.

Proof. By theorem 39.3.2, a d-connection preserves the horizontal and vertical distributions, such that

$$\begin{aligned} R(X,Y)Z &= \mathbf{h}\nabla_X\nabla_Y\mathbf{h}Z + \mathbf{v}\nabla_X\nabla_Y\mathbf{v}Z - \mathbf{h}\nabla_Y\nabla_X\mathbf{h}Z - \mathbf{v}\nabla_Y\nabla_X\mathbf{v}Z - \mathbf{h}\nabla_{[X,Y]}\mathbf{h}Z - \mathbf{v}\nabla_{[X,Y]}\mathbf{v}Z \\ &= \mathbf{h}R(X,Y)\mathbf{h}Z + \mathbf{v}R(X,Y)\mathbf{v}Z\,. \quad \blacksquare \end{aligned} \tag{39.5.3}$$

Further, we can use the fact that a N-linear connection acts identically on horizontal and vertical vectors, i.e., it commutes with the tangent structure. Hence, the same holds also for the curvature,

$$R(X,Y)JZ = JR(X,Y)Z\,, \tag{39.5.4}$$

and so it suffices to consider only horizontal (or only vertical) vector fields Z. Thus, we have to decompose only X and Y into horizontal and vertical parts, and consider each of them separately, leaving only three independent components of the curvature. Writing them as endomorphisms acting on d-vector fields, we find that their components are given as follows.

1. The horizontal part:

$$R(\delta_a, \delta_b) = \left[\delta_a F^d{}_{cb} - \delta_b F^d{}_{ca} + F^d{}_{ea}F^e{}_{cb} - F^d{}_{eb}F^e{}_{ca} + C^d{}_{ce}(\delta_a N^e{}_b - \delta_b N^e{}_a)\right](\delta_d \otimes \mathrm{d}x^c + \bar{\partial}_d \otimes \delta\bar{x}^c)\,. \tag{39.5.5a}$$

2. The mixed part:

$$R(\delta_a, \bar{\partial}_b) = \left(\delta_a C^d{}_{cb} - \bar{\partial}_b F^d{}_{ca} + F^d{}_{ea}C^e{}_{cb} - C^d{}_{eb}F^e{}_{ca} - C^d{}_{ce}\bar{\partial}_b N^e{}_a\right)(\delta_d \otimes \mathrm{d}x^c + \bar{\partial}_d \otimes \delta\bar{x}^c)\,. \tag{39.5.5b}$$

3. The vertical part:

$$R(\bar{\partial}_a, \bar{\partial}_b) = \left(\bar{\partial}_a C^d{}_{cb} - \bar{\partial}_b C^d{}_{ca} + C^d{}_{ea} C^e{}_{cb} - C^d{}_{eb} C^e{}_{ca}\right)\left(\delta_d \otimes \mathrm{d}x^c + \bar{\partial}_d \otimes \delta \bar{x}^c\right). \quad (39.5.5c)$$

39.6 Bianchi identities

39.7 Autoparallel curves

39.8 Berwald connection

We have seen in section 39.3 that we can specify a N-linear connection by its action on the Berwald basis, hence fixing the connection coefficients (39.3.6), and that these are fully determined from the mixed components of the torsion, as detailed in section 39.4. Since the torsion is tensorial, there exists a canonical choice, namely to demand that these torsion components vanish. This leads to a canonically defined N-linear connection, which is fully determined from the non-linear connection as follows.

Definition 39.8.1 (Berwald connection)**.** Let M be a manifold with a non-linear connection. The *Berwald connection* is the unique N-linear connection $\overset{\text{Be}}{\nabla}$ satisfying $\overset{\text{Be}}{T}(\mathbf{h}X, \mathbf{v}Y) = 0$ for all $X, Y \in \mathrm{Vect}(TM)$.

By comparison with the component expressions (39.4.3) for the torsion, one finds that the connection coefficients of the Berwald connection are given by

$$\overset{\text{Be}}{F}{}^a{}_{bc} = \bar{\partial}_b N^a{}_c\,, \quad \overset{\text{Be}}{C}{}^a{}_{bc} = 0\,. \tag{39.8.1}$$

One can also express the Berwald connection without resorting to tensor components. It turns out that the following useful formula holds.

Theorem 39.8.1. *The Berwald connection is given by*

$$\overset{\text{Be}}{\nabla}_X Y = \mathbf{v}[\mathbf{h}X, \mathbf{v}Y] + \mathbf{h}[\mathbf{v}X, \mathbf{h}Y] + J[\mathbf{v}X, \Theta Y] + \Theta[\mathbf{h}X, JY] \tag{39.8.2}$$

for all $X, Y \in \mathrm{Vect}(TM)$.

Proof. We first check that the defining relation (39.8.2) yields a linear connection on the tangent bundle. For this purpose we check the linearity

$$\begin{aligned} \overset{\text{Be}}{\nabla}_{fX} Y &= \mathbf{v}[f\mathbf{h}X, \mathbf{v}Y] + \mathbf{h}[f\mathbf{v}X, \mathbf{h}Y] + J[f\mathbf{v}X, \Theta Y] + \Theta[f\mathbf{h}X, JY] \\ &= f\mathbf{v}[\mathbf{h}X, \mathbf{v}Y] + f\mathbf{h}[\mathbf{v}X, \mathbf{h}Y] + fJ[\mathbf{v}X, \Theta Y] + f\Theta[\mathbf{h}X, JY] \\ &\quad - (\mathbf{v}Yf)\mathbf{v}\mathbf{h}X - (\mathbf{h}Yf)\mathbf{h}\mathbf{v}X - (\Theta Yf)J\mathbf{v}X - (JYf)\Theta\mathbf{h}X \\ &= f\overset{\text{Be}}{\nabla}_X Y\,, \end{aligned} \tag{39.8.3}$$

as well as the Leibniz rule

$$\begin{aligned} \overset{\text{Be}}{\nabla}_X (fY) &= \mathbf{v}[\mathbf{h}X, f\mathbf{v}Y] + \mathbf{h}[\mathbf{v}X, f\mathbf{h}Y] + J[\mathbf{v}X, f\Theta Y] + \Theta[\mathbf{h}X, fJY] \\ &= f\mathbf{v}[\mathbf{h}X, \mathbf{v}Y] + f\mathbf{h}[\mathbf{v}X, \mathbf{h}Y] + fJ[\mathbf{v}X, \Theta Y] + f\Theta[\mathbf{h}X, JY] \\ &\quad + (\mathbf{h}Xf)\mathbf{v}\mathbf{v}Y + (\mathbf{v}Xf)\mathbf{h}\mathbf{h}Y + (\mathbf{v}Xf)J\Theta Y + (\mathbf{h}Xf)\Theta JY \\ &= f\overset{\text{Be}}{\nabla}_X Y + (Xf)Y\,. \end{aligned} \tag{39.8.4}$$

Next, we check that it is a d-connection, by calculating the covariant derivative of the almost product structure:

$$\begin{aligned}\overset{\text{Be}}{\nabla}_X(\mathbb{P}Y) - \mathbb{P}\overset{\text{Be}}{\nabla}_X Y &= \mathbf{v}[\mathbf{h}X, -\mathbf{v}Y] + \mathbf{h}[\mathbf{v}X, \mathbf{h}Y] + J[\mathbf{v}X, -\Theta Y] + \Theta[\mathbf{h}X, JY] \\ &\quad + \mathbf{v}[\mathbf{h}X, \mathbf{v}Y] - \mathbf{h}[\mathbf{v}X, \mathbf{h}Y] + J[\mathbf{v}X, \Theta Y] - \Theta[\mathbf{h}X, JY] \\ &= 0\,.\end{aligned} \tag{39.8.5}$$

To check that it is a N-linear connection, one similarly calculates the covariant derivative of the tangent structure:

$$\begin{aligned}\overset{\text{Be}}{\nabla}_X(JY) - J\overset{\text{Be}}{\nabla}_X Y &= \mathbf{v}[\mathbf{h}X, \mathbf{v}JY] + \mathbf{h}[\mathbf{v}X, \mathbf{h}JY] + J[\mathbf{v}X, \Theta JY] + \Theta[\mathbf{h}X, JJY] \\ &\quad - J\mathbf{v}[\mathbf{h}X, \mathbf{v}Y] - J\mathbf{h}[\mathbf{v}X, \mathbf{h}Y] - JJ[\mathbf{v}X, \Theta Y] - J\Theta[\mathbf{h}X, JY] \\ &= \mathbf{v}[\mathbf{h}X, JY] + J[\mathbf{v}X, \mathbf{h}Y] - J[\mathbf{v}X, \mathbf{h}Y] - \mathbf{v}[\mathbf{h}X, JY] \\ &= 0\,.\end{aligned} \tag{39.8.6}$$

Finally, we calculate the torsion component

$$\begin{aligned}\overset{\text{Be}}{T}(\mathbf{h}X, \mathbf{v}Y) &= \overset{\text{Be}}{\nabla}_{\mathbf{h}X}(\mathbf{v}Y) - \overset{\text{Be}}{\nabla}_{\mathbf{v}Y}(\mathbf{h}X) - [\mathbf{h}X, \mathbf{v}Y] \\ &= \mathbf{v}[\mathbf{h}X, \mathbf{v}Y] - \mathbf{h}[\mathbf{v}Y, \mathbf{h}X] - [\mathbf{h}X, \mathbf{v}Y] \\ &= 0\,.\end{aligned} \tag{39.8.7}$$

From the torsion decomposition in section 39.4 follows that this uniquely determines the N-linear connection. ■

Of course, one can also use the formula (39.8.2) to obtain a coordinate expression for the Berwald connection. Writing the vector fields X, Y as $X = X^a\delta_a + \bar{X}^a\bar{\partial}_a$ and $Y = Y^a\delta_a + \bar{Y}^a\bar{\partial}_a$ in the Berwald basis, one finds the covariant derivative

$$\overset{\text{Be}}{\nabla}_X Y = X^a\left[\left(\delta_a Y^b + \bar{\partial}_c N^b{}_a Y^c\right)\delta_b + \left(\delta_a \bar{Y}^b + \bar{\partial}_c N^b{}_a \bar{Y}^c\right)\bar{\partial}_b\right] + \bar{X}^a\left(\bar{\partial}_a Y^b\delta_b + \bar{\partial}_a\bar{Y}^b\bar{\partial}_b\right)\,, \tag{39.8.8}$$

which follows from $[\delta_a, \bar{\partial}_b] = \bar{\partial}_b N^c{}_a\bar{\partial}_c$ and shows that its connection coefficients are given by the formula (39.8.1).

We now apply the formula introduced above to show the following.

Theorem 39.8.2. *The Berwald connection of a homogeneous non-linear connection satisfies*

$$\nabla_X\mathbf{c} = \mathbf{v}X \tag{39.8.9}$$

for all vector fields $X \in \mathrm{Vect}(TM)$.

Proof. Since the Liouville vector field is vertical, we have for a homogeneous non-linear connection

$$\begin{aligned}\nabla_X\mathbf{c} &= \mathbf{v}[\mathbf{h}X, \mathbf{c}] + J[\mathbf{v}X, \Theta\mathbf{c}] \\ &= \mathbf{v}[\mathbf{h}X, \mathbf{c}] + J[\mathbf{v}X, \Theta\mathbf{c}] \\ &= -\mathbb{P}\mathbf{v}X - \mathbb{T}X \\ &= \mathbf{v}X\,,\end{aligned} \tag{39.8.10}$$

using theorem 37.4.2 and the properties of induced semisprays and induced non-linear connections, where the latter in particular implies that the tension $\mathbb{T}$ vanishes. ■

39.9 Dynamical covariant derivative

From the Berwald connection and the induced semispray of the underlying non-linear connection one can construct a particular derivative operator, which is defined as follows.

Definition 39.9.1 (Dynamical covariant derivative)**.** Let M be a manifold with a non-linear connection. The *dynamical covariant derivative* of a d-tensor field is its Berwald derivative $\blacktriangledown = \overset{\text{Be}}{\nabla}_{\Theta\mathbf{c}}$ with respect to the induced semispray $\Theta\mathbf{c}$.

Given this operator, we may now study its action on various geometric objects. We start with a few tensors which we have previously encountered.

Theorem 39.9.1. *The tangent structure, almost product structure, almost complex structure and adjoint structure are covariantly constant with respect to the dynamical covariant derivative,*

$$\blacktriangledown J = \blacktriangledown \mathbb{P} = \blacktriangledown \mathbb{F} = \blacktriangledown \Theta = 0\,. \tag{39.9.1}$$

Proof. This immediately follows from the fact that the Berwald connection is a N-linear connection. ∎

Another interesting case is the action on vertical lifts of functions and vector fields.

Theorem 39.9.2. *For every function $f \in C^{\infty}(M, \mathbb{R})$ and vector field $X \in \mathrm{Vect}(M)$ the dynamical covariant derivative satisfies*

$$\blacktriangledown \overset{V}{f} = \overset{C}{f}\,, \quad \blacktriangledown \overset{V}{X} = \mathbf{v}\overset{C}{X}\,. \tag{39.9.2}$$

Proof. The first statement immediately follows from the fact that the covariant derivative of a function is simply the application of the induced semispray $\Theta\mathbf{c}$ and theorem 38.1.9. For the second statement we calculate

$$\blacktriangledown \overset{V}{X} = \mathbf{v}[\Theta\mathbf{c}, \overset{V}{X}] = \mathbf{v}[\Theta\mathbf{c}, J\overset{C}{X}] = \blacktriangleright \ldots \blacktriangleleft \tag{39.9.3}$$

∎

39.10 Affine bundle of d-connections

39.11 Pullback and Lie derivative of d-tensors

39.12 Pullback and Lie derivative of d-connections

Chapter 40

Finsler geometry

40.1 Finsler functions and length functionals

The basic idea behind Finsler geometry is to introduce a length measure for curves, or more precisely singular curve segments $\gamma : [0,1] \to M$ on a manifold M. This length measure is assumed to be as general as possible, with the restriction that it should not depend on the parametrization of the curve. We will see below that this can be achieved by introducing a function on the tangent bundle TM. However, this function will not be differentiable at the zero section on TM. To overcome with the technical issues arising from this fact, we define the following bundle.

Definition 40.1.1 (Slit tangent bundle)**.** Let M be a manifold. Its *slit tangent bundle* $\mathring{T}M = TM \setminus \{0\}$ is the fiber bundle $\mathring{\tau} : \mathring{T}M \to M$ obtained from the tangent bundle TM by removing the zero section,

$$\mathring{T}M = \bigcup_{x \in M} T_xM \setminus \{0_x\}\,. \tag{40.1.1}$$

Note that $\mathring{T}M$ is *not* a vector bundle, since its fibers are not vector spaces - they do not have a zero element. However, the total space $\mathring{T}M$ is still a smooth manifold, and $\mathring{\tau} : \mathring{T}M \to M$ is a fiber bundle with typical fiber $\mathbb{R}^n \setminus \{0\}$, where $n = \dim M$. Also we find that various structures we introduced on TM can be restricted to $\mathring{T}M$. This applies in particular to the dilatations introduced in definition 19.9.1, which yield an action of $(\mathbb{R}, +)$ on $\mathring{T}M$, and hence also the Liouville vector field from definition 19.9.2, which restricts to a vector field on $\mathring{T}M$. This allows us to define homogeneity also for objects defined on $\mathring{T}M$. We make use of this fact in the following definition.

Definition 40.1.2 (Finsler function)**.** Let M be a manifold. A non-negative function $F : TM \to \mathbb{R}$ is called a *Finsler function* if and only if it satisfies the following conditions:

1. F is positive on $\mathring{T}M$.
2. F is smooth on $\mathring{T}M$ and continuous on TM.
3. F is positively homogeneous of order 1.

4. The bilinear form $g_u : T_{\mathring{\tau}(u)}M \times T_{\mathring{\tau}(u)}M \to \mathbb{R}$ defined by

$$g_u(v, w) = \frac{1}{2} \frac{\partial^2}{\partial s \partial t} F^2(u + sv + tw)\Big|_{s=t=0} \tag{40.1.2}$$

is non-degenerate for all $u \in \mathring{T}M$ and has constant signature.

There are different alternative definitions, in particular concerning the bilinear form g_u. In the classical definition of Finsler geometry, one demands that g_u is positive definite [Run59, BCS91, MA94, She01b, SLK14, SS16]. Another definition does not make any restrictions on the signature [She01a]. Here we follow a definition which requires only that the signature is the same everywhere [BM07]. Another property of the Finsler function is usually not part of its definition, but may be included as well:

Definition 40.1.3 (Reversible Finsler function)**.** A Finsler function F on a manifold M is called *reversible* if $F(-v) = F(v)$ for all $v \in TM$.

With the help of the Finsler function, we can now define the length of a curve, as we aimed for.

Definition 40.1.4 (Finsler length)**.** Let M be a manifold with Finsler function F and $\gamma : [a, b] \to M$ a singular curve segment. The *Finsler length* of γ is the integral

$$\ell(\gamma) = \int_a^b (F \circ \dot{\gamma})(t)\, \mathrm{d}t\,. \tag{40.1.3}$$

In the previous definition, $\dot{\gamma} : [a, b] \to TM$ denotes the canonical lift of γ, following definition 7.3.2. By the virtue of the homogeneity of the Finsler function, we can now show the following.

Theorem 40.1.1. *The Finsler length of a singular curve segment is invariant under orientation-preserving reparametrization.*

Proof. Consider a smooth function $\varphi : [a, b] \to [\tilde{a}, \tilde{b}]$ with $\varphi(a) = \tilde{a}$, $\varphi(b) = \tilde{b}$ and $\varphi'(t) > 0$ for all $t \in [a, b]$, and define $\gamma = \tilde{\gamma} \circ \varphi : [a, b] \to M$. Then we have

$$\dot{\gamma}(t) = \varphi'(t)\dot{\tilde{\gamma}}(\varphi(t)) \tag{40.1.4}$$

for all $t \in [a, b]$. From the positive 1-homogeneity of F then follows

$$F(\dot{\gamma}(t)) = \varphi'(t)F(\dot{\tilde{\gamma}}(\varphi(t))) \tag{40.1.5}$$

for all $t \in [a, b]$. From the change-of-variable formula for integrals finally follows the length

$$\begin{aligned}\ell(\gamma) &= \int_a^b (F \circ \dot{\gamma})(t)\,\mathrm{d}t \\ &= \int_a^b F(\dot{\gamma}(t))\,\mathrm{d}t \\ &= \int_a^b F(\dot{\tilde{\gamma}}(\varphi(t)))\varphi'(t)\,\mathrm{d}t \\ &= \int_{\tilde{a}}^{\tilde{b}} F(\dot{\tilde{\gamma}}(\tilde{t}))\,\mathrm{d}\tilde{t} \\ &= \int_{\tilde{a}}^{\tilde{b}} (F \circ \dot{\tilde{\gamma}})(\tilde{t})\,\mathrm{d}\tilde{t} \\ &= \ell(\tilde{\gamma})\,. \quad \blacksquare\end{aligned} \tag{40.1.6}$$

40.2 Finsler metric

The bilinear form introduced in the definition 40.1.2 is very commonly used, and therefore carries its own name.

Definition 40.2.1 (Finsler metric)**.** The *Finsler metric* $g \in \Gamma(\mathring{\tau}^* \operatorname{Sym}^2 T^*M)$ is defined by

$$g_u(v, w) = \frac{1}{2}\frac{\partial^2}{\partial s \partial t} F^2(u + sv + tw)\bigg|_{s=t=0}\,. \tag{40.2.1}$$

40.3 Hilbert form

There is another possible interpretation for the Finsler length integral. Recall from section 25.1 that one-forms may be integrated over curve segments. One may wonder whether also the Finsler length can be expressed in this way as an integral of a one-form. This is indeed the case, however, the one-form will be defined on the tangent bundle instead of the manifold itself. We define it as follows.

Definition 40.3.1 (Hilbert one-form)**.** Let M be a manifold with Finsler function F. The *Hilbert one-form* is the horizontal one-form $\alpha \in \Omega^1(\mathring{T}M)$ defined by

$$\alpha = J^*(\mathrm{d}F)\,. \tag{40.3.1}$$

One easily checks the following property.

Theorem 40.3.1. *The Hilbert form is homogeneous of order 0.*

Proof. This follows from the fact that F is 1-homogeneous, while J^* is homogeneous of order -1. ■

It is also helpful to note the coordinate expression of the Hilbert form, using the induced coordinates on $\mathring{T}M$. From the differential of F, given by

$$\mathrm{d}F = \partial_a F \,\mathrm{d}x^a + \bar{\partial}_a F \,\mathrm{d}\bar{x}^a \tag{40.3.2}$$

and the coordinate expression (29.3.6) of the cotangent structure J^* one finds

$$\alpha = \bar{\partial}_a F \,\mathrm{d}x^a \,. \tag{40.3.3}$$

Using the Hilbert form, we can now come to the following statement.

Theorem 40.3.2. *The Finsler length of a singular curve segment $\gamma : [a, b] \to M$ is given by the pullback of the Hilbert form as*

$$\ell(\gamma) = \int_{[a,b]} \dot{\gamma}^* \alpha \,. \tag{40.3.4}$$

Proof. The one-form on $[a, b]$ under the integral can be written as

$$\dot{\gamma}^*\alpha = \dot{\gamma}^* J^* \mathrm{d}F = \langle \partial_t, \dot{\gamma}^* J^* \mathrm{d}F\rangle \,\mathrm{d}t = \langle (J \circ \dot{\gamma}_*)(\partial_t), \mathrm{d}F\rangle \,\mathrm{d}t \,, \tag{40.3.5}$$

by expanding in the canonical coordinate basis and using the definition of the pullback of a differential form. Note that the pushforward $\dot{\gamma}_*(\partial_t)$ of the coordinate basis vector at t is just the tangent vector $\ddot{\gamma}(t)$, and so we have

$$(J \circ \dot{\gamma}_*)(\partial_t) = (J \circ \ddot{\gamma})(t) = (\mathbf{c} \circ \dot{\gamma})(t) \,, \tag{40.3.6}$$

using theorem 29.2.8. Using the fact that F is 1-homogeneous by definition, and so

$$\mathbf{c} \lrcorner \mathrm{d}F = \mathbf{c}F = F \,, \tag{40.3.7}$$

we have

$$\dot{\gamma}^*\alpha = \langle (\mathbf{c} \circ \dot{\gamma})(t), \mathrm{d}F\rangle \,\mathrm{d}t = (F \circ \dot{\gamma})(t)\mathrm{d}t \,, \tag{40.3.8}$$

which leads to the length functional given in definition 40.1.4. ■

40.4 Cartan forms

There is another set of differential forms which are commonly used in Finsler geometry, which differ from the Hilbert form by their homogeneity and which we define here for later use when we discuss connections and geodesics. We start with the following definition.

Definition 40.4.1 (Cartan one-form). Let M be a manifold with Finsler function F. The *Cartan one-form* is the horizontal one-form $\theta \in \Omega^1(\mathring{T}M)$ defined by

$$\theta = \frac{1}{2} J^*(\mathrm{d}F^2) = F\alpha \,. \tag{40.4.1}$$

Before we discuss its properties, we continue with another definition.

Definition 40.4.2 (Cartan two-form). Let M be a manifold with Finsler function F. The *Cartan two-form* is the two-form $\omega \in \Omega^2(\mathring{T}M)$ defined by

$$\omega = \mathrm{d}\theta \,. \tag{40.4.2}$$

The following is now straightforward.

Theorem 40.4.1. *The Cartan one- and two-forms are homogeneous of order 1.*

Proof. From the homogeneity of F and α follows, using definition 40.4.1, that θ is 1-homogeneous. Using theorem 19.9.4, the same follows also for ω. ■

The Cartan two-form plays an important role in Finsler geometry. It equips the tangent bundle with the structure of a symplectic manifold, as we will show next.

Theorem 40.4.2. *The Cartan two-form is a symplectic form.*

Proof. Following definition 35.1.2, we must show that ω is closed and non-degenerate. The first property follows directly from the definition, since

$$\mathrm{d}\omega = \mathrm{d}\mathrm{d}\theta = 0\,. \tag{40.4.3}$$

To show that it is non-degenerate, consider two vector fields $X, Y \in \mathrm{Vect}(M)$, as well as their lifts $\overset{\mathrm{V}}{X} = J\overset{\mathrm{C}}{X}$ and $\overset{\mathrm{C}}{Y}$. From theorem 9.4.2 we have

$$\begin{aligned}
\overset{\mathrm{C}}{Y} \lrcorner \overset{\mathrm{V}}{X} \lrcorner \omega &= \overset{\mathrm{C}}{Y} \lrcorner \overset{\mathrm{V}}{X} \lrcorner \mathrm{d}\theta \\
&= \overset{\mathrm{V}}{X}(\overset{\mathrm{C}}{Y} \lrcorner \theta) - \overset{\mathrm{C}}{Y}(\overset{\mathrm{V}}{X} \lrcorner \theta) - [\overset{\mathrm{V}}{X}, \overset{\mathrm{C}}{Y}] \lrcorner \theta \\
&= \frac{1}{2}\left[\overset{\mathrm{V}}{X}(\overset{\mathrm{C}}{Y} \lrcorner J^*(\mathrm{d}F^2)) - \overset{\mathrm{C}}{Y}(\overset{\mathrm{V}}{X} \lrcorner J^*(\mathrm{d}F^2)) - [X,Y]^{\mathrm{V}} \lrcorner J^*(\mathrm{d}F^2)\right] \\
&= \frac{1}{2}\left[\overset{\mathrm{V}}{X}(J\overset{\mathrm{C}}{Y} \lrcorner \mathrm{d}F^2) - \overset{\mathrm{C}}{Y}(J\overset{\mathrm{V}}{X} \lrcorner \mathrm{d}F^2) - J[X,Y]^{\mathrm{V}} \lrcorner \mathrm{d}F^2\right] \\
&= \frac{1}{2}\overset{\mathrm{V}}{X}\overset{\mathrm{V}}{Y}F^2\,.
\end{aligned} \tag{40.4.4}$$

Following the definition 29.5.1 of the vertical lift, as well as the proof of theorem 29.5.5, we have for $u \in \mathring{T}M$ the relation

$$\frac{1}{2}\left(\overset{\mathrm{V}}{X}\overset{\mathrm{V}}{Y}\right)F^2(u) = \frac{1}{2}\frac{\partial^2}{\partial s\partial t}F^2(u + sX(\mathring{\tau}(u)) + tY(\mathring{\tau}(u)))\bigg|_{s=t=0} = g_u(X(\mathring{\tau}(u)), Y(\mathring{\tau}(u)))\,. \tag{40.4.5}$$

By definition, the Finsler metric is non-degenerate. It follows that for every vertical vector at u, which can be expressed as the vertical lift of a vector field on M evaluated at u, there exists a vector at $\mathring{\tau}(u)$ so that the expression given in the preceding equation is non-vanishing. Conversely, for every vector at $\mathring{\tau}(u)$, there exists a vertical vector at u such that the expression given above is again non-vanishing. Since the vertical and complete lifts span $T\mathring{T}M$, it follows that ω is non-degenerate. ■

Note in particular that the expression we obtained at the end of the proof is symmetric in X and Y. This is not a coincidence, and we will now prove this in a more general approach.

Theorem 40.4.3. *For any two vector fields X, Y, the Cartan two-form satisfies*

$$\omega(JX, Y) = \omega(JY, X)\,. \tag{40.4.6}$$

Proof. We use the fact that the Nijenhuis tensor of the tangent structure vanishes and calculate

$$\begin{aligned}
0 &= \frac{1}{2}N_J(X,Y) \lrcorner \mathrm{d}F^2 \\
&= \frac{1}{2}([JX,JY] + J^2[X,Y] - J[JX,Y] - J[X,JY]) \lrcorner \mathrm{d}F^2 \\
&= \frac{1}{2}\left\{([JX,JY] \lrcorner \mathrm{d}F^2 - [JX,Y] \lrcorner J^*\mathrm{d}F^2 - [X,JY] \lrcorner J^*\mathrm{d}F^2\right\} \\
&= \frac{1}{2}\left[(JX)((JY)F^2) - (JY)((JX)F^2)\right] - [JX,Y] \lrcorner \theta - [X,JY] \lrcorner \theta \\
&= (JX)(Y \lrcorner \theta) - (JY)(X \lrcorner \theta) - [JX,Y] \lrcorner \theta - [X,JY] \lrcorner \theta \\
&= \mathrm{d}\theta(JX,Y) - \mathrm{d}\theta(JY,X) + Y(JX \lrcorner \theta) - X(JY \lrcorner \theta) \\
&= \omega(JX,Y) - \omega(JY,X)\,,
\end{aligned} \tag{40.4.7}$$

where we used the fact that θ vanishes on the vertical vector fields JX and JY. ■

Also for the Cartan one- and two-forms one may easily derive coordinate expressions. For the Cartan one-form one easily finds, in analogy to the Hilbert form (40.3.3), the expression

$$\theta = \frac{1}{2}\bar{\partial}_a F^2\,\mathrm{d}x^a = F\bar{\partial}_a F\,\mathrm{d}x^a = g_{ab}\bar{x}^a \mathrm{d}x^b\,. \tag{40.4.8}$$

For the Cartan two-form, it is now straightforward to take the exterior derivative and obtain

$$\omega = \mathrm{d}\theta = \frac{1}{2}\partial_a\bar{\partial}_b F^2\,\mathrm{d}x^a \wedge \mathrm{d}x^b + \frac{1}{2}\bar{\partial}_a\bar{\partial}_b F^2\,\mathrm{d}\bar{x}^a \wedge \mathrm{d}x^b\,. \tag{40.4.9}$$

40.5 Geodesic spray

Having introduced the Cartan one- and two-forms, it is not straightforward to arrive at another important notion in Finsler geometry, which we define as follows.

Definition 40.5.1 (Geodesic spray)**.** Let M be a manifold with Finsler function F. The *geodesic spray* $\mathbf{S}$ is the unique vector field on $\mathring{T}M$ which satisfies

$$\iota_{\mathbf{S}}\omega = -\frac{1}{2}\mathrm{d}F^2\,. \tag{40.5.1}$$

The reason for the first part of its name will become clear in the following section 40.6, where we discuss Finsler geodesics. The second part of the name indicates that $\mathbf{S}$ is a spray, which we will check next.

Theorem 40.5.1. *The geodesic spray* $\mathbf{S}$ *is a spray.*

Proof. In the following we will make use of the fact that the Cartan two-form is non-degenerate. Let $X \in \mathrm{Vect}(\mathring{T}M)$ be a vector field. Using the relationship 40.4.3 between the Cartan two-form and the tangent structure, we have

$$\omega(J\mathbf{S},X) = -\omega(\mathbf{S},JX) = \frac{1}{2}JX \lrcorner \mathrm{d}F^2 = X \lrcorner \theta\,. \tag{40.5.2}$$

We can compare this to

$$\begin{aligned}\omega(\mathbf{c}, X) &= \mathrm{d}\theta(\mathbf{c}, X)\\ &= \mathbf{c}(X \lrcorner \theta) - X(\mathbf{c} \lrcorner \theta) - [\mathbf{c}, X] \lrcorner \theta\\ &= \mathcal{L}_{\mathbf{c}}(X \lrcorner \theta) - [\mathbf{c}, X] \lrcorner \theta\\ &= X \lrcorner \mathcal{L}_{\mathbf{c}}\theta\\ &= X \lrcorner \theta\,,\end{aligned} \tag{40.5.3}$$

where we use the fact that $\mathbf{c}$ is vertical, so that $\mathbf{c} \lrcorner \theta$ vanishes, the commutator (16.5.8) of the Lie derivative with the interior product and the 1-homogeneity of θ. We thus see that

$$\omega(J\mathbf{S}, X) = X \lrcorner \theta = \omega(\mathbf{c}, X) \tag{40.5.4}$$

for arbitrary vector fields X, Since ω is non-degenerate, it thus follows that $J\mathbf{S} = \mathbf{c}$. Further, by theorem 40.4.1, the Cartan two-form is 1-homogeneous, while F^2 is 2-homogeneous, since the Finsler function is 1-homogeneous by definition. Since the Cartan two-form is furthermore non-degenerate, it preserves the homogeneity (i.e., its interior product with a vector field will not drop any components of the vector field, and send any homogeneous component to another homogeneous component of homogeneity increased by 1). Hence, the geodesic spray must be 1-homogeneous, and thus a spray. ■

The geodesic spray has a number of useful properties, which we will study now and in the following sections. The following property is immediately evident.

Theorem 40.5.2. *The Finsler function is constant along the geodesic spray,* $\mathbf{S}F = 0$.

Proof. Using the fact that the Finsler function is non-vanishing, we can write

$$\mathbf{S}F = \iota_{\mathbf{S}}\mathrm{d}F = \frac{1}{2F}\iota_{\mathbf{S}}\mathrm{d}F^2 = -\frac{1}{F}\iota_{\mathbf{S}}\iota_{\mathbf{S}}\omega = 0\,. \tag{40.5.5}$$

■

These properties can also easily be seen from a coordinate derivation. First, we expand

$$\mathrm{d}F^2 = \partial_a F^2 \mathrm{d}x^a + \bar{\partial}_a F^2 \mathrm{d}\bar{x}^a\,. \tag{40.5.6}$$

Denoting $\mathbf{S} = S^a\partial_a + \bar{S}^a\bar{\partial}_a$, we further have

$$\begin{aligned}-2\iota_{\mathbf{S}}\omega &= S^b(\partial_a\bar{\partial}_b F^2 - \partial_b\bar{\partial}_a F^2)\mathrm{d}x^a + \bar{\partial}_a\bar{\partial}_b F^2(S^b\mathrm{d}\bar{x}^a - \bar{S}^a\mathrm{d}x^b)\\ &= \left[S^b(\partial_a\bar{\partial}_b F^2 - \partial_b\bar{\partial}_a F^2) - 2\bar{S}^b g_{ab}\right]\mathrm{d}x^a + 2S^b g_{ab}\mathrm{d}\bar{x}^a\,,\end{aligned} \tag{40.5.7}$$

using the coordinate expression for the Finsler metric to arrive at the second line. We thus obtain the two equations

$$\partial_a F^2 = S^b(\partial_a\bar{\partial}_b F^2 - \partial_b\bar{\partial}_a F^2) - 2\bar{S}^b g_{ab}\,, \tag{40.5.8a}$$

$$\bar{\partial}_a F^2 = 2S^b g_{ab}\,. \tag{40.5.8b}$$

Using the properties of the Finsler metric derived in section 40.2, we can thus solve the second equation, from which we derive

$$S^a = \frac{1}{2}g^{ab}\bar{\partial}_b F^2 = \bar{x}^a\,. \tag{40.5.9}$$

We can then insert this result into the first equation, which we solve to obtain

$$\begin{aligned}\bar{S}^a &= \frac{1}{2}g^{ab}\left[\bar{x}^c(\partial_b\bar{\partial}_c F^2 - \partial_c\bar{\partial}_b F^2) - \partial_b F^2\right]\\ &= \frac{1}{2}g^{ab}(\partial_b F^2 - \bar{x}^c\partial_c\bar{\partial}_b F^2)\,,\end{aligned} \tag{40.5.10}$$

using the homogeneity of the Finsler function to obtain the second line.

From the obtained result we now see the aforementioned properties of the geodesic spray. The expression for S^a shown in equation (40.5.9) is the characteristic property of a semispray. In order see that it is a spray, we use the expression for $\bar{S}^a$ from equation (40.5.10) to obtain

$$\begin{aligned}\bar{x}^b\bar{\partial}_b\bar{S}^a &= \frac{1}{2}\bar{x}^d\bar{\partial}_d\left[g^{ab}(\partial_bF^2 - \bar{x}^c\partial_c\bar{\partial}_bF^2)\right]\\ &= \frac{1}{2}g^{ab}\left(\bar{x}^d\bar{\partial}_d\partial_bF^2 - \bar{x}^d\bar{\partial}_d\bar{x}^c\partial_c\bar{\partial}_bF^2 - \bar{x}^c\bar{x}^d\bar{\partial}_d\partial_c\bar{\partial}_bF^2\right)\\ &= \frac{1}{2}g^{ab}\left(2\partial_bF^2 - \bar{x}^c\partial_c\bar{\partial}_bF^2 - \bar{x}^c\partial_c\bar{\partial}_bF^2\right)\\ &= g^{ab}(\partial_bF^2 - \bar{x}^c\partial_c\bar{\partial}_bF^2)\\ &= 2\bar{S}^a\,,\end{aligned} \tag{40.5.11}$$

using the homogeneity of the Finsler metric and the Finsler function, and so one obtains desired homogeneity property, which shows that $\mathbf{S}$ is a spray. Finally, for the Finsler function one has

$$\begin{aligned}\mathbf{S}F^2 &= \bar{x}^a\partial_aF^2 + \frac{1}{2}g^{ab}(\partial_bF^2 - \bar{x}^c\partial_c\bar{\partial}_bF^2)\bar{\partial}_aF^2\\ &= \bar{x}^a\partial_aF^2 + \bar{x}^b(\partial_bF^2 - \bar{x}^c\partial_c\bar{\partial}_bF^2)\\ &= 0\,,\end{aligned} \tag{40.5.12}$$

once again using its homogeneity and the properties of the Finsler metric.

40.6 Finsler geodesics

40.7 Induced non-linear connection

We have seen in the previous section that the Finsler function defines a distinguished spray, which we called the geodesic spray. Following our discussion in section 38.3, a spray induces a unique homogeneous and torsion-free non-linear connection. This holds, of course, also for the geodesic spray. From theorem 38.3.1 we derive the following definition.

Definition 40.7.1 (Induced non-linear connection). Let M be a manifold with Finsler function F. The *induced non-linear connection* on M is defined by its almost product structure $\mathbb{P}$ as

$$\mathbb{P} = -\mathcal{L}_{\mathbf{S}}J\,, \tag{40.7.1}$$

where $\mathbf{S}$ is the geodesic spray.

Using the relation (38.3.6) derived in section 38.3, as well as the expression (40.5.10) for the geodesic spray, one finds that the coefficients of the non-linear connection are given by

$$N^a{}_b = -\frac{1}{2}\bar{\partial}_b\bar{S}^a = \frac{1}{4}\bar{\partial}_b\left[g^{ad}(\bar{x}^c\partial_c\bar{\partial}_dF^2 - \partial_dF^2)\right]\,. \tag{40.7.2}$$

A few properties of the induced non-linear connection follow immediately from its definition in terms of a spray, and have been proven already in chapter 38. We briefly summarize them here:

1. The induced non-linear connection is symmetric: $\mathbf{t} = 0$ (theorem 38.3.2).
2. The induced non-linear connection is homogeneous: $\mathbb{T} = 0$ (theorem 38.5.2).

3. The geodesic spray is the unique horizontal semispray (theorem 38.4.2), and is induced by the connection: $\mathbf{S} = \Theta\mathbf{c}$ (theorem 38.5.2).

Besides these properties, the non-linear connection also has a close relationship with the Finsler function from which it is induced. We find the following property.

Theorem 40.7.1. *The Finsler function is constant along any horizontal vector field.*

Proof. We can write any horizontal vector field in the form $\mathbf{h}X$ for an arbitrary vector field X. Using the definition of the induced non-linear connection, we then have

$$\begin{aligned}
2\mathbf{h}XF^2 &= (X + \mathbb{P}X)F^2 \\
&= [X - (\mathcal{L}_{\mathbf{S}}J)X]F^2 \\
&= (X + J[\mathbf{S}, X] - [\mathbf{S}, JX])F^2 \\
&= XF^2 + J[\mathbf{S}, X] \lrcorner \mathrm{d}F^2 - \mathbf{S}((JX)F^2) + (JX)(\mathbf{S}F^2) \\
&= XF^2 + [\mathbf{S}, X] \lrcorner J^*\mathrm{d}F^2 - \mathbf{S}(X \lrcorner J^*\mathrm{d}F^2) \\
&= XF^2 + 2[\mathbf{S}, X] \lrcorner \theta - 2\mathbf{S}(X \lrcorner \theta) \\
&= XF^2 + 2\mathbf{S}(X \lrcorner \theta) - 2X(\mathbf{S} \lrcorner \theta) - 2X \lrcorner \mathbf{S} \lrcorner \mathrm{d}\theta - 2\mathbf{S}(X \lrcorner \theta) \\
&= XF^2 - X(\mathbf{S} \lrcorner J^*\mathrm{d}F^2) - 2X \lrcorner \mathbf{S} \lrcorner \omega \\
&= XF^2 - X(J\mathbf{S} \lrcorner \mathrm{d}F^2) + X \lrcorner \mathrm{d}F^2 \\
&= XF^2 - X(\mathbf{c} \lrcorner \mathrm{d}F^2) + XF^2 \\
&= 2XF^2 - X(\mathbf{c}F^2) \\
&= 2XF^2 - 2XF^2 \\
&= 0\,,
\end{aligned} \tag{40.7.3}$$

where we used the numerous definitions and theorems given previously in this chapter. ■

This can also be seen from the coordinate expression (40.7.2) of the non-linear connection. Using the fact that the horizontal vector fields are spanned by the Berwald basis elements δ_a, we have

$$\begin{aligned}
\delta_a F &= (\partial_a - N^b{}_a\bar{\partial}_b)F \\
&= \partial_a F - \frac{1}{4}\bar{\partial}_a\left[g^{bc}(\bar{x}^d\partial_d\bar{\partial}_c F^2 - \partial_c F^2)\right]\bar{\partial}_b F \\
&= \partial_a F - \frac{1}{4F}\left[g^{bc}\bar{\partial}_a(\bar{x}^d\partial_d\bar{\partial}_c F^2 - \partial_c F^2) - g^{bf}g^{gc}\bar{\partial}_a g_{fg}(\bar{x}^d\partial_d\bar{\partial}_c F^2 - \partial_c F^2)\right]g_{be}\bar{x}^e \\
&= \partial_a F - \frac{1}{4F}\left[\bar{x}^c\bar{\partial}_a(\bar{x}^d\partial_d\bar{\partial}_c F^2 - \partial_c F^2) - \frac{1}{2}\bar{x}^f g^{gc}\bar{\partial}_a\bar{\partial}_f\bar{\partial}_g F^2(\bar{x}^d\partial_d\bar{\partial}_c F^2 - \partial_c F^2)\right] \\
&= \partial_a F - \frac{1}{4F}\left(\bar{x}^c\partial_a\bar{\partial}_c F^2 + \bar{x}^c\bar{x}^d\bar{\partial}_a\partial_d\bar{\partial}_c F^2 - \bar{x}^c\bar{\partial}_a\partial_c F^2\right) \\
&= \partial_a F - \frac{1}{4F}\left(2\partial_a F^2 + \bar{x}^d\bar{\partial}_a\partial_d F^2 - \bar{x}^c\bar{\partial}_a\partial_c F^2\right) \\
&= \partial_a F - \frac{1}{4F}4F\partial_a F \\
&= 0\,.
\end{aligned} \tag{40.7.4}$$

Given a non-linear connection, it is most convenient to express tensor fields on the (slit) tangent bundle in terms of the Berwald basis introduced in definition 37.1.3. For the geodesic spray, it follows already from the fact that it is induced by the non-linear connection and the relation (38.4.2) that it is given by

$$\mathbf{S} = \bar{x}^a\delta_a\,. \tag{40.7.5}$$

For the Cartan two-form, we have

$$
\begin{aligned}
\omega &= \frac{1}{2}\partial_a\bar{\partial}_b F^2\,\mathrm{d}x^a \wedge \mathrm{d}x^b + g_{ab}\mathrm{d}\bar{x}^a \wedge \mathrm{d}x^b \\
&= \frac{1}{2}\partial_a\bar{\partial}_b F^2\,\mathrm{d}x^a \wedge \mathrm{d}x^b + g_{ab}(\delta\bar{x}^a - N^a{}_c\mathrm{d}x^c) \wedge \mathrm{d}x^b \\
&= \left(\frac{1}{2}\partial_a\bar{\partial}_b F^2 - g_{bc}N^c{}_a\right)\mathrm{d}x^a \wedge \mathrm{d}x^b + g_{ab}\delta\bar{x}^a \wedge \mathrm{d}x^b \\
&= \frac{1}{2}\left(\partial_a\bar{\partial}_b F^2 - N^c{}_a\bar{\partial}_c\bar{\partial}_b F^2\right)\mathrm{d}x^a \wedge \mathrm{d}x^b + g_{ab}\delta\bar{x}^a \wedge \mathrm{d}x^b \\
&= \frac{1}{2}\delta_a\bar{\partial}_b F^2\mathrm{d}x^a \wedge \mathrm{d}x^b + g_{ab}\delta\bar{x}^a \wedge \mathrm{d}x^b\,.
\end{aligned} \tag{40.7.6}
$$

For the first term, note that only the part antisymmetric in a and b contributes, and we can further exchange the derivatives as

$$
\delta_{[a}\bar{\partial}_{b]}F^2 = \bar{\partial}_{[b}\delta_{a]}F^2 + \bar{\partial}_{[b}N^c{}_{a]}\bar{\partial}_c F^2 = 0\,, \tag{40.7.7}
$$

where the first term vanishes as the consequence of theorem 40.7.1, while the second term is the weak torsion (37.5.5), and thus also vanishes, since the non-linear connection is symmetric. We are thus left with

$$
\omega = g_{ab}\delta\bar{x}^a \wedge \mathrm{d}x^b\,. \tag{40.7.8}
$$

Note that from this expression also our result from theorem 40.4.2 becomes evident, where we showed that the Cartan two-form is non-degenerate.

40.8 Induced d-tensors

Having constructed a non-linear connection from the Finsler geometry, one also has a notion of d-tensors and the Berwald linear derivative. Some of the former we have already encountered, such as the horizontal geodesic spray, the vertical Liouville vector field and the horizontal Cartan 1-form. We will now construct further helpful d-tensors which are obtained by taking derivatives of the Finsler function. The first one we study is essentially already familiar.

Definition 40.8.1 (Finsler metric)**.** Let M be a manifold with Finsler function F. The *Finsler metric* is the d-tensor defined by

$$
\mathbf{g} = \frac{1}{2}\overset{\mathrm{Be}}{\nabla}{}^v\overset{\mathrm{Be}}{\nabla}{}^v F^2\,. \tag{40.8.1}
$$

Theorem 40.8.1. *The dynamical covariant derivative of the Finsler metric vanishes.*

Proof. ►...◄ ■

Definition 40.8.2 (Cartan tensor)**.** Let M be a manifold with Finsler function F. The *Cartan tensor* is the d-tensor defined by

$$
\mathbf{C} = \frac{1}{2}\overset{\mathrm{Be}}{\nabla}{}^v\mathbf{g} = \frac{1}{4}\overset{\mathrm{Be}}{\nabla}{}^v\overset{\mathrm{Be}}{\nabla}{}^v\overset{\mathrm{Be}}{\nabla}{}^v F^2\,. \tag{40.8.2}
$$

Definition 40.8.3 (Landsberg tensor)**.** Let M be a manifold with Finsler function F. The *Landsberg tensor* is the d-tensor defined by

$$\mathbf{P} = \frac{1}{2}\overset{\mathrm{Be}}{\nabla}{}^{h}\mathbf{g} = \frac{1}{4}\overset{\mathrm{Be}}{\nabla}{}^{h}\overset{\mathrm{Be}}{\nabla}{}^{v}\overset{\mathrm{Be}}{\nabla}{}^{v}F^{2}\,. \tag{40.8.3}$$

Theorem 40.8.2. *The Cartan and Landsberg tensors are related by*

$$\mathbf{P} = -J^{*}\blacktriangledown\mathbf{C}\,. \tag{40.8.4}$$

Proof. ▶. . . ◀ ■

40.9 Sasaki metric

Definition 40.9.1 (Sasaki metric)**.** Let M be a manifold with Finsler function F. The *Sasaki metric* G is the metric on the tangent bundle TM which is defined such that for vector fields $X, Y \in \mathrm{Vect}(\mathring{T}M)$ holds

$$G(X, Y) = \omega\left(X, JY - \frac{\Theta Y}{F^{2}}\right), \tag{40.9.1}$$

where ω is the Cartan two-form.

40.10 Volume form

40.11 Induced linear connections

We have seen in section 40.7 that a Finsler geometry uniquely defines a non-linear, homogeneous connection on the tangent bundle. Following our treatment in section 39.3.1, we thus have the possibility to define the notion of a d-connection on Finsler geometries. In this section we will show that we can derive a number of particular d-connections from the Finsler function, each defined such that it satisfies a number of particular properties. We have, in fact, already encountered an example for such a d-connection, namely the Berwald connection in section 39.8, which is defined by any non-linear connection, and hence in particular also by the non-linear connection we derived from the Finsler geometry in section 40.7. We can, however, also bypass constructing the non-linear connection first, and define the Berwald connection directly from the Finsler function (alongside the non-linear connection, but without using the formula (40.7.1)). This construction is given by the following statement.

Theorem 40.11.1. *The Berwald connection on a Finsler geometry is the unique N-linear connection which satisfies the following:*

1. *The Liouville vector field is horizontally constant:* $\overset{Be}{\nabla}{}^{h}\mathbf{c} = 0$.
2. *The Finsler function is horizontally constant:* $\overset{Be}{\nabla}{}^{h}F^{2} = 0$.
3. *The torsion satisfies* $\mathbf{h}\overset{Be}{T}(\mathbf{h}X, \mathbf{h}Y) = \overset{Be}{T}(\mathbf{h}X, \mathbf{v}Y) = 0$ *for all* $X, Y \in \mathrm{Vect}(\mathring{T}M)$.

Proof. ▶. . . ◀ ■

Given a Finsler function, one may define a number of other connections besides the Berwald connection, by imposing different sets of defining conditions. One of the most natural, which is reminiscent of the Levi-Civita connection discussed in section 31.10.1, is to demand that the Finsler metric $\mathbf{g}$ is covariantly constant and that it is symmetric, in the sense of definition 39.4.2. This leads us to the following definition.

Definition 40.11.1 (Cartan linear connection)**.** The Cartan linear connection on a Finsler geometry is the unique N-linear connection which satisfies the following:

1. The Liouville vector field is horizontally constant: $\overset{\text{Ca}}{\nabla}{}^h\mathbf{c} = 0$.
2. The Finsler metric is covariantly constant: $\overset{\text{Ca}}{\nabla}\mathbf{g} = 0$.
3. The connection is symmetric: $\mathbf{h}\overset{\text{Ca}}{T}(\mathbf{h}X, \mathbf{h}Y) = \mathbf{v}\overset{\text{Ca}}{T}(\mathbf{v}X, \mathbf{v}Y) = 0$ for all $X, Y \in \operatorname{Vect}(\mathring{T}M)$.

▶Derive coordinate expression.◀

$$\overset{\text{Ca}}{F}{}^a{}_{bc} = \frac{1}{2}g^{ad}(\delta_b g_{dc} + \delta_c g_{bd} - \delta_d g_{bc})\,, \quad \overset{\text{Ca}}{C}{}^a{}_{bc} = \frac{1}{2}g^{ad}\bar{\partial}_d g_{bc}\,. \tag{40.11.1}$$

One may also combine the conditions which are imposed for the Berwald and Cartan connections. One possibility is to demand that the Finsler metric is only *horizontally* constant. This leads to the following definition.

Definition 40.11.2 (Chern-Rund linear connection)**.** The Chern-Rund linear connection on a Finsler geometry is the unique N-linear connection which satisfies the following:

1. The Liouville vector field is horizontally constant: $\overset{\text{CR}}{\nabla}{}^h\mathbf{c} = 0$.
2. The Finsler metric is horizontally constant: $\overset{\text{CR}}{\nabla}{}^h\mathbf{g} = 0$.
3. The torsion satisfies $\mathbf{h}\overset{\text{CR}}{T}(\mathbf{h}X, Y) = 0$ for all $X, Y \in \operatorname{Vect}(\mathring{T}M)$.

▶Derive coordinate expression.◀

$$\overset{\text{CR}}{F}{}^a{}_{bc} = \frac{1}{2}g^{ad}(\delta_b g_{dc} + \delta_c g_{bd} - \delta_d g_{bc})\,, \quad \overset{\text{CR}}{C}{}^a{}_{bc} = 0\,. \tag{40.11.2}$$

Another possibility to define a N-linear connection is now straightforward, namely to demand that the Finsler metric is only *vertically* constant. This leads to the following definition.

Definition 40.11.3 (Hashiguchi linear connection)**.** The Hashiguchi linear connection on a Finsler geometry is the unique N-linear connection which satisfies the following:

1. The Liouville vector field is horizontally constant: $\overset{\text{Ha}}{\nabla}{}^h\mathbf{c} = 0$.

2. The Finsler function is horizontally constant: $\overset{\mathrm{Ha}}{\nabla}{}^h F^2 = 0$.

3. The Finsler metric is vertically constant: $\overset{\mathrm{Ha}}{\nabla}{}^v \mathbf{g} = 0$.

4. The torsion satisfies $\mathbf{h}\overset{\mathrm{Ha}}{T}(\mathbf{h}X, \mathbf{h}Y) = \mathbf{v}\overset{\mathrm{Ha}}{T}(X, \mathbf{v}Y) = 0$ for all $X, Y \in \mathrm{Vect}(\mathring{T}M)$.

►Derive coordinate expression.◄

$$\overset{\mathrm{Ha}}{F}{}^a{}_{bc} = \bar{\partial}_b N^a{}_c \,, \quad \overset{\mathrm{Ha}}{C}{}^a{}_{bc} = \frac{1}{2} g^{ad} \bar{\partial}_d g_{bc} \,. \tag{40.11.3}$$

We can thus summarize the connection coefficients of the four N-linear connections as follows:

$$\overset{\mathrm{Be}}{F}{}^a{}_{bc} = \overset{\mathrm{Ha}}{F}{}^a{}_{bc} = \bar{\partial}_b N^a{}_c \,, \qquad \overset{\mathrm{Ca}}{F}{}^a{}_{bc} = \overset{\mathrm{CR}}{F}{}^a{}_{bc} = \frac{1}{2} g^{ad} (\delta_b g_{dc} + \delta_c g_{bd} - \delta_d g_{bc}) \,, \tag{40.11.4a}$$

$$\overset{\mathrm{Be}}{C}{}^a{}_{bc} = \overset{\mathrm{CR}}{C}{}^a{}_{bc} = 0 \,, \qquad \overset{\mathrm{Ca}}{C}{}^a{}_{bc} = \overset{\mathrm{Ha}}{C}{}^a{}_{bc} = \frac{1}{2} g^{ad} \bar{\partial}_d g_{bc} \,. \tag{40.11.4b}$$

40.12 Unit tangent bundle

Chapter 41

Klein geometries and homogeneous spaces

In the following we will study a class of spaces, which can be viewed from two different perspectives, and are then denoted either as *Klein geometries* (after Felix Klein, who studied these spaces as part of his Erlangen programme) or as *homogeneous spaces*. This chapter mostly follows the treatment in [Sha97, ch. 4].

41.1 Klein geometries

In this section, we start with the first perspective on Klein geometries. For this purpose, we start with the definition of the central object we discuss in this section.

Definition 41.1.1 (Klein geometry)**.** A *Klein geometry* is a pair (G, H), where G is a Lie group, called the *principal group* of the Klein geometry, and H is a closed subgroup of G, such that the (left) coset space G/H, called the *space* of the Klein geometry, is connected.

By its definition, a Klein geometry gives rise to a number of interesting and useful geometric structures. First, recall from the definition 15.4.1 of a coset space G/H that its elements are the cosets (equivalence classes) gH for $g \in G$, or in other words the orbits of the right action $\cdot : G \times H \to G, (g, h) \mapsto gh$ of H on G by right multiplication. For the definition of a Klein geometry, we have demanded that H is a closed subgroup of G - not without reason, as the following statement shows.

Theorem 41.1.1. *The space G/H of a Klein geometry (G, H) is a smooth manifold of dimension $\dim G/H = \dim G - \dim H$.*

Proof. The proof makes use of the exponential map from definition 15.7.2, and is only briefly sketched here. Let $m = \dim H$ and $n = \dim G$. One first chooses a basis $(a_1, \ldots, a_n)$ of $\mathfrak{g}$, such that $(a_1, \ldots, a_m)$ is a basis of $\mathfrak{h}$. To construct a chart of G around the unit element $e \in G$, one uses the fact that there exists an open set $U \subset G$ with $e \in U$ such that $\exp|_{\exp^{-1}(U)}$ is a diffeomorphism onto U. Similarly, there exists an open set $V \subset H$ with $e \in V$ such that $\exp|_{\exp^{-1}(V)}$ is a diffeomorphism onto V. To obtain a chart around eH on G/H, ▶. . . ◀ ■

Further, a Klein geometry defines a number of maps relating the manifolds G, H and G/H. First, we note the existence of the following group action.

Theorem 41.1.2. *The principal group G of a Klein geometry (G, H) acts transitively on its space G/H from the left via the group action $\rho : (g, \tilde{g}H) \mapsto (g\tilde{g})H$.*

Proof. We have already shown that G/H is a smooth manifold. To see that the prescription ρ given above constitutes a (left) group action, one uses the associativity of group multiplication to find

$$\rho(g_1.\rho(g_2, g_3H)) = \rho(g_1, (g_2g_3)H) = (g_1g_2g_3)H = \rho(g_1g_2, g_3H) \tag{41.1.1}$$

for all $g_1, g_2, g_3 \in G$. Smoothness of ρ follows from the fact that multiplication in a Lie group is smooth. Finally, to see that ρ is transitive, let $g_1, g_2 \in G$. To check that there exists $g \in G$ such that $\rho(g, g_1H) = g_2H$, simply choose $g = g_2g_1^{-1}$. ■

Another important action in the context of Klein geometries is that of the group H on G. It allows us to define a principal fiber bundle as follows:

Theorem 41.1.3. *The canonical projection $\pi : G \to G/H, g \mapsto gH$ from a group G onto the coset space G/H defined by a closed subgroup $H \subset G$ defines a principal H-bundle, where the right action $\cdot : G\times H \to G$ of H on G is given by multiplication from the right, $(g, h) \mapsto g\cdot h = gh$.*

Proof. Again, we only sketch the proof. We have already show in theorem 41.1.1 that G/H is a manifold. To further see that $\pi : G \to G/H$ is a fiber bundle, one first constructs a local trivialization around the coset $eH \in G/H$ of the unit element $e \in G$, again by using the exponential map from definition 15.7.2. From this one obtains for all $g \in G$ a trivialization around $gH \in G/H$ by left multiplication with g. This defines the fiber bundle geometry.

To complete the proof, we explicitly show that the fiber bundle we constructed carries a free, transitive, fiber preserving right action of H, given by $g \cdot h = gh \in G$ for $g \in G$ and $h \in H$. To show this, note first that for $g, g' \in G$ we have $gH = \pi(g) = \pi(g') = g'H$ if and only if

$$\{gh, h \in H\} = \{g'h, h \in H\}, \tag{41.1.2}$$

and hence if and only if there exists $h \in H$ such that $g' = gh$. This has two implications. First, it follows that $\pi(gh) = \pi(g)$, so that the action preserves the fibers. Second, if g and g' belong to the same fiber, $\pi(g) = \pi(g')$, there exists $h \in H$ such that $g' = gh$, and so the action is transitive on the fibers. Also h is unique, since if we would have any other $h' \in H$ which also satisfies $g' = gh'$, we could multiply from the left by g^{-1} and obtain

$$h = g^{-1}g' = h'. \tag{41.1.3}$$

It thus follows that the group action is also free on the fibers. Hence, it defines a principal H-bundle. ■

41.2 Geometric orientation

Depending on the properties of the Lie groups G and H, we can distinguish different types of Klein geometries.

Definition 41.2.1 (Geometrically oriented Klein geometry). A Klein geometry (G, H) is called *geometrically oriented* if its principal group G is connected.

It must be noted that the space G/H of a geometrically oriented Klein geometry (G, H) is in general *not* orientable in the sense of definition 23.5.5! An example is given by projective spaces. In fact, given a Klein geometry which is not geometrically oriented, we can still obtain a geometrically oriented one by the following construction.

Definition 41.2.2 (Associated geometrically oriented Klein geometry)**.** Let (G, H) be a Klein geometry, $G_0 \subset G$ the connected component of the identity in G and $H_0 = G_0 \cap H$. Then (G_0, H_0) is called the *associated geometrically oriented Klein geometry* of (G, H).

Naturally the question arises how the Klein geometries (G, H) and (G_0, H_0) are related to each other. One finds that the following relations hold in the case that the coset space G/H is connected.

Theorem 41.2.1. *Let (G, H) be a Klein geometry such that G/H is connected and (G_0, H_0) its associated geometrically oriented Klein geometry. Then $G = G_0 \cdot H$ and $G/H \cong G_0/H_0$.*

Proof. To show the first part of the proposition, note that for every $g \in G$ there exists a path in G/H connecting gH and eH, since G/H is connected, i.e., a curve $\gamma : [0, 1] \to G/H$ with $\gamma(0) = eH$ and $\gamma(1) = gH$. Since $\pi : G \to G/H$ is a fiber bundle, there exists a curve $\Gamma : [0, 1] \to G$ such that $\gamma = \pi \circ \Gamma$, $\Gamma(1) = g$ and $\Gamma(0) = h \in \pi^{-1}(eH) = H$. From this another curve $\Gamma' : [0, 1] \to G, t \mapsto \Gamma(t)h^{-1}$ is obtained, which satisfies $\Gamma'(0) = hh^{-1} = e$ and $\Gamma'(1) = gh^{-1}$. Since this path connects e and gh^{-1}, it follows that $gh^{-1} \in G_0$ lies in the identity component G_0 of G. Hence, $g = gh^{-1}h$ can be written as a product of $gh^{-1} \in G_0$ and $h \in H$.

To show the second part, consider a map ▶. . . ◀ ■

We will keep the discussion of geometrically oriented Klein geometries for later, when we consider Cartan geometries modeled on these Klein geometries, where the notion will become relevant.

41.3 Kernel and effective Klein geometries

Another important characterization of Klein geometries arises from the question whether the action of G on G/H is effective. To study this question, it is useful to first define the following subgroup.

Definition 41.3.1 (Kernel of a Klein geometry)**.** Let (G, H) be a Klein geometry. Its *kernel* is the largest subgroup $K \subset H$ that is normal in G.

Recall from group theory that a subgroup $K \subset G$ is normal subgroup (written as $K \triangleleft G$) if and only if $gkg^{-1} \in K$ for all $k \in K$ and $g \in G$. While this definition of the kernel appears rather abstract, it still has a very direct geometric interpretation for the Klein geometry. It turns out that it can alternatively be characterized by the following property.

Theorem 41.3.1. *The kernel K of a Klein geometry (G, H) is the subgroup of G defined by*

$$K = \{k \in G : \rho(k, x) = x \forall x \in G/H\}\,, \tag{41.3.1}$$

where $\rho : G \times G/H \to G/H$ is the canonical left action, i.e., K is the subgroup containing those elements k which act trivially on G/H.

Proof. From the fact that ρ is a group action follows that

$$\rho(kk', x) = \rho(k, \rho(k', x)) = \rho(k, x) = x\,, \tag{41.3.2}$$

as well as

$$x = \rho(k^{-1}k, x) = \rho(k^{-1}, \rho(k, x)) = \rho(k^{-1}, x) \tag{41.3.3}$$

for all $k, k' \in K$ and $x \in G/H$, and so K is a subgroup of G. Further, setting in particular $x = eH$ leads to

$$kH = \rho(k, eH) = eH\,, \tag{41.3.4}$$

and so $k \in H$. Hence, $K \subset H$ is a subgroup of H. To show that it is normal in G, let $g \in G$. Then we have

$$\rho(gkg^{-1}, x) = \rho(g, \rho(k, \rho(g^{-1}, x))) = \rho(g, \rho(g^{-1}, x)) = x\,, \tag{41.3.5}$$

so that also $gkg^{-1} \in K$, showing that K is normal in G. Finally, we need to show that K is maximal. For this purpose, let $K' \in H$ any subgroup of H which is normal in G. Then for any $k \in K'$ and $x \in G/H$, we choose a representative $g \in G$ such that $x = gH$, and find

$$\rho(k, gH) = \rho(kg, eH) = \rho(gk'g^{-1}g, eH) = \rho(gk', eH) = \rho(g, eH) = gH\,, \tag{41.3.6}$$

where we used the fact that $k' \in K'$ since K' was assumed normal in G, as well as $\rho(k', eH) = eH$ since we assumed $K' \subset H$. Hence, K' stabilizes every $x \in G/H$, and so $K' \subset K$, proving that K is maximal. ■

From this property we can now define a particular type of Klein geometries.

Definition 41.3.2 (Effective Klein geometry)**.** A Klein geometry is called *effective* if its kernel is trivial, i.e., it contains only the identity element.

The name already suggests that there exists a relation between effective Klein geometries and effective Lie group actions, following their definition 15.3.2. The following turns out to be true.

Theorem 41.3.2. *A Klein geometry (G, H) is effective if and only if the action of G on G/H is effective.*

Proof. Recall from definition 15.3.2 that a Lie group action $\rho : G \times G/H \to G/H$ is effective if and only if for every distinct $g, g' \in G$ there exists $x \in G/H$ such that $\rho(g, x) \neq \rho(g', x)$. In particular, we may choose $g' = e$, from which follows that ρ is effective if and only if for every $g \in G \setminus \{e\}$ there exists $x \in G/H$ such that

$$\rho(g, x) \neq \rho(e, x) = x\,, \tag{41.3.7}$$

i.e., if and only if $g \notin K$, where K is the kernel of (G, H). Hence, ρ is effective if and only if $K = \{e\}$, and so (G/H) is effective. ■

Definition 41.3.3 (Locally effective Klein geometry)**.** A Klein geometry is called *locally effective* if its kernel is discrete, i.e., it consists of isolated points of G.

Theorem 41.3.3. *Let (G, H) be a Klein geometry with kernel K and $N \subset K$ a closed subgroup which is normal in G. Then $(G/N, H/N)$ is a Klein geometry whose space $(G/N)/(H/N)$ is canonically diffeomorphic to G/H.*

Proof. ▶. . . ◀ ■

Theorem 41.3.4. *The Klein geometry $(G/N, H/N)$ from theorem 41.3.3 is effective if and only if $N = K$.*

Proof. ▶...◀ ■

Definition 41.3.4 (Associated effective Klein geometry)**.** Let (G, H) be a Klein geometry with kernel K. Its *associated effective Klein geometry* is the Klein geometry $(G/K, H/K)$.

One may wonder why one considers ineffective Klein geometries at all, if the coset space is also described by their (unique) associated effective Klein geometry, instead of demanding effectiveness from the beginning. One reason is the fact that ineffective Klein geometries naturally in physics. One example is the description of spin, and the usage of spin groups, which we discuss in chapter 45.

41.4 Homogeneous spaces

Another approach towards essentially the same type of geometric structure starts not from the closed subgroup H of G, but from the space on which G acts. We define this space, equipped with a suitable action, as follows.

Definition 41.4.1 (Homogeneous space)**.** A *homogeneous space* is a manifold M together with a transitive left action $\rho : G \times M$ of a Lie group G.

Theorem 41.4.1. *Let (G, H) be a Klein geometry. Then G/H, together with the canonical action of G on G/H, is a homogeneous space.*

Proof. This follows directly from the definition 41.4.1 of a homogeneous space and the statements 41.1.1 that G/H is a manifold, and 41.1.2 that G acts transitively on G/H. ■

Theorem 41.4.2. *Let M be a connected homogeneous space with left action ρ of a Lie group g. Further, let $x \in M$ and $H_x = \{h \in G, \rho(h, x) = x\}$ the stabilizer of x. Then (G, H_x) is a Klein geometry and there exists an equivariant diffeomorphism $\varphi : G/H_x \to M$ such that $\varphi(eH) = x$.*

Proof. First, we must check that H_x is closed. This follows from the fact that $H_x = \rho_x^{-1}(x)$ is the preimage of the closed set $\{x\}$ under the (by definition continuous) map

$$\begin{array}{cccc} \rho_x & : & G & \to & M \\ & & g & \mapsto & \rho(g, x) \end{array} . \tag{41.4.1}$$

Next, consider the map

$$\begin{array}{cccc} \varphi & : & G/H_x & \to & M \\ & & gH_x & \mapsto & \rho(g, x) \end{array} . \tag{41.4.2}$$

First, we have to check that this is well-defined, i.e., that $\varphi(gH_x) = \varphi(ghH_x)$ for all $g \in G$ and $h \in H_x$. This follows from the fact that ρ is a left action and H_x is the stabilizer of x, so that

$$\varphi(ghH_x) = \rho(gh, x) = \rho(g, \rho(h, x)) = \rho(g, x) = \varphi(gH_x) . \tag{41.4.3}$$

Next, we check that it is injective. For this purpose, consider $g, g' \in G$ with $g^{-1}g' \notin H_x$.

▶...◀

We then check surjectivity. Let $y \in M$. From the assumption that ρ is a transitive group action follows that there exists $g \in G$ such that $y = \rho(g, x) = \varphi(gH_x)$, so that φ is indeed surjective, and hence bijective.

Finally, we must still check that φ is a diffeomorphism, i.e., that both φ and its inverse are smooth. ►. . . ◄ ■

There are many examples of homogeneous spaces and Klein geometries. One of the nicest examples is the following.

Example 41.4.1. Let $G = \mathrm{SO}(n+1)$. One can see $\mathrm{SO}(n+1)$ as the group of rotations in $n+1$ dimensions, i.e., linear maps $g : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ which map the unit sphere $S^n \subset \mathbb{R}^{n+1}$ to itself. This means that $\mathrm{SO}(n+1)$ acts (from the left) on the sphere S^n. One easily checks that this action is transitive, so that this turns the sphere into a homogeneous space. Let H be the stabilizer of the north pole, i.e., the set of those linear maps $h \in \mathrm{SO}(n+1)$ which map the north pole $o \in S^n$ to itself. The matrix representation of such an element h has a block form

$$\begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}, \tag{41.4.4}$$

where $A \in \mathrm{SO}(n) \cong H$. To relate S^n to the coset space G/H, recall that G acts transitively, and so for every $x \in S^n$ there exists a $g \in G$ such that $x = g(o)$. This element is unique up to right multiplication by $h \in H$, since the stabilizer of x is given by gHh^{-1}. A coset gH can thus uniquely be identified by the image $g(o) \in S^n$. This identification defines a diffeomorphism, so that $(\mathrm{SO}(n+1), \mathrm{SO}(n))$ is a Klein geometry with space $\mathrm{SO}(n+1)/\mathrm{SO(n)} \cong S^n$.

Many other examples can be constructed in a similar fashion. Here we list a few of the most relevant examples encountered in physics.

Example 41.4.2. The following manifolds are obtained as homogeneous spaces:

- The sphere $S^n \cong \mathrm{SO}(n+1)/\mathrm{SO}(n)$, as detailed in example 41.4.1.
- Euclidean space $E^n \cong \mathbb{R}^n \cong \mathrm{ISO}(n)/\mathrm{SO}(n)$, with the Euclidean group $\mathrm{ISO}(n) = \mathbb{R}^n \rtimes \mathrm{SO}(n)$ acting by translations and rigid rotations.
- Hyperbolic space
$$H^n = \{(x_0, \vec{x}) \in \mathbb{R}^n, x_0^2 - \|\vec{x}\|^2 = 1\} \cong \mathrm{SO}(1,n)/\mathrm{SO}(n)\,, \tag{41.4.5}$$
with the symmetry group $\mathrm{SO}(1,n)$.
- Affine space $A^n \cong \mathbb{R}^n \cong \mathrm{IGL}(n)/\mathrm{GL}(n)$, with the affine group $\mathrm{IGL}(n) = \mathbb{R}^n \rtimes \mathrm{GL}(n)$ acting by affine transformations, i.e., translations and linear transformations.
- Minkowski space of n dimensions $\mathrm{ISO}(1,n-1)/\mathrm{SO}(1,n-1)$, with the action of the Poincaré group $\mathrm{ISO}(1,n-1)$.
- De Sitter space of n dimensions $\mathrm{SO}(1,n)/\mathrm{SO}(1,n-1)$, with the action of the de Sitter group $\mathrm{SO}(1,n)$.
- Anti de Sitter space of n dimensions $\mathrm{SO}(2,n-1)/\mathrm{SO}(1,n-1)$, with the action of the anti de Sitter group $\mathrm{SO}(2,n-1)$.

41.5 Tangent bundle

Theorem 41.5.1. *For every Klein geometry (G,H), there exists a canonical isomorphism between the vector bundles $T(G/H)$ and $G \times_{\mathrm{Ad}} \mathfrak{g}/\mathfrak{h}$, where G is understood as a principal H-*

bundle over G/H and Ad *denotes the adjoint representation of H on $\mathfrak{g}/\mathfrak{h}$.*

Proof. ▶. . . ◀ ■

41.6 Mutation

Definition 41.6.1 (Geometrical isomorphism of Klein geometries)**.** Let (G_1, H_1) and (G_2, H_2) be Klein geometries. A *geometrical isomorphism* relating (G_1, H_1) and (G_2, H_2) is a Lie group isomorphism $\varphi : G_1 \to G_2$ such that $\varphi(H_1) = H_2$. If a geometrical isomorphism exists, (G_1, H_1) and (G_2, H_2) are called *geometrically isomorphic*.

Definition 41.6.2 (Mutation of Klein geometries)**.** Let (G_1, H_1) and (G_2, H_2) be Klein geometries. A *mutation* relating (G_1, H_1) and (G_2, H_2) is a pair (φ, λ), where $\varphi : H_1 \to H_2$ is a Lie group isomorphism and $\lambda : \mathfrak{g}_1 \to \mathfrak{g}_2$ is a vector space isomorphism such that $\lambda(a) = \varphi_{*e}(h)$ and

$$\lambda(\mathrm{Ad}(h)(a)) = \mathrm{Ad}(\varphi(h))(\lambda(a)) \tag{41.6.1}$$

for all $h \in H_1$ and $a \in \mathfrak{g}_1$. If a mutation exists, (G_1, H_1) and (G_2, H_2) are called *mutants*.

Chapter 42

Cartan geometry

42.1 Cartan connection

Using the notion of a Klein geometry, or equivalently a homogeneous space, introduced in the previous section, we can now come to the definition of a Cartan geometry. This is achieved as follows.

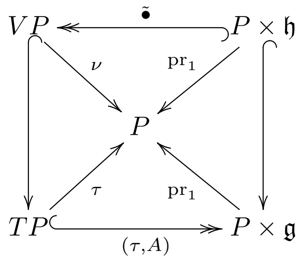
Definition 42.1.1 (Cartan geometry)**.** A *Cartan geometry* modeled on a Klein geometry (G, H) is a principal H-bundle $\pi : P \to M$ together with a $\mathfrak{g}$-valued one-form $A \in \Omega^1(P, \mathfrak{g})$ on P (the *Cartan connection*), such that

1. For each $p \in P$, $A_p : T_pP \to \mathfrak{g}$ is a linear isomorphism.
2. A is H-equivariant: $(R_h)^*A = \mathrm{Ad}(h^{-1}) \circ A \; \forall h \in H$.
3. $A(\tilde{a}) = a$ for all $a \in \mathfrak{h}$, where $\tilde{a}$ denotes the fundamental vector field of a.

A few remarks are in order.

1. From the condition that there exists an isomorphism between tangent spaces T_pP and $\mathfrak{g} \cong T_eG$, where $e \in G$ is the unit element, follows $\dim P = \dim G$. Further, since the fibers of a principal H-bundle have dimension $\dim H$, one has $\dim M = \dim G - \dim H = \dim G/H$.
2. By fixing a basis of $\mathfrak{g}$, a Cartan connection induces a basis of every tangent space T_pP, i.e., a section of the frame bundle. It thus follows that P is necessarily parallelizable. Note, however, that M need not be parallelizable.
3. A Cartan connection shares several properties with a principal connection as given in definition 27.1.2, in particular its equivariance and its action on the fundamental vector fields generated by the action of the structure group H on P. However, it is different, since it takes values in a different Lie algebra $\mathfrak{g}$ instead of $\mathfrak{h}$, and it has non-vanishing kernel, whereas a principal connection has the horizontal distribution as its kernel.
4. Since the fundamental vector fields span the vertical tangent space VP, it follows that a vector field $X \in \mathrm{Vect}(P)$ is vertical if and only if $\iota_X A$ takes values in $\mathfrak{h}$.

The fact that the Cartan connection defines a linear bijection between every tangent space T_pP and the Lie algebra $\mathfrak{g}$ can also be stated differently. For this purpose, consider the following commutative diagram:



(42.1.1)

The corners of this diagram show four vector bundles over P: the vertical tangent bundle $\nu : VP \to P$, the tangent bundle $\tau : TP \to P$ and the two trivial bundles $P \times \mathfrak{g}$ and $P \times \mathfrak{h}$, together with their respective projections onto the first factor P. Further, one has the canonical inclusions, since $VP \subset TP$ and $P \times \mathfrak{h} \subset P \times \mathfrak{g}$. The upper arrow represents the fundamental vector fields obtained from the right action on P, i.e., the map $(p, a) \mapsto \tilde{a}(p)$, where $\tilde{a} \in \mathrm{Vect}(P)$ is the fundamental vector field. This map is a vector bundle isomorphism covering the identity on P, since it establishes a linear isomorphism between $\mathfrak{h}$ and V_pP for all $p \in P$. The latter also holds for the map represented by the lower arrow, due to the first condition on a Cartan connection. Since TP and $P \times \mathfrak{g}$ are isomorphic, it follows that TP is trivial, so that P is parallelizable. Finally, the last condition in the definition 42.1.1 demands that the outer square of this diagram commutes.

The picture above now also allows to consider a dual picture, by taking the inverse of the lower arrow, to introduce a map $(\tau, A)^{-1} : P \times \mathfrak{g} \to TP$. Similarly to the fundamental vector fields represented by the upper arrow, we can fix $a \in \mathfrak{g}$, and obtain a map $\underline{A}(a) : P \to TP$, satisfying $\tau \circ \underline{A}(a) = \mathrm{id}_P$; this is a vector field on TP. This yields the following definition.

Definition 42.1.2 (Associated vector fields)**.** Let $(\pi : P \to M, A)$ be a Cartan geometry modeled on the Klein geometry (G, H). For each $a \in \mathfrak{g}$ the *associated vector field* $\underline{A}(a) \in \mathrm{Vect}(P)$ is the unique vector field such that $A(\underline{A}(a)) = a$.

Since the associated vector fields uniquely define the Cartan connection and vice versa, one can equivalently define a Cartan geometry in terms of the associated vector fields instead of the Cartan connection. In this case also the associated vector fields must satisfy an equivalent set of conditions, which takes the following form.

Theorem 42.1.1. *The associated vector fields $\underline{A}$ of a Cartan connection satisfy the following conditions, which are equivalent to the conditions on the Cartan connection posed in definition 42.1.1:*

1. *For each $p \in P$, $\underline{A}_p : \mathfrak{g} \to T_pP$ is a linear isomorphism.*
2. *$\underline{A}$ is H-equivariant: $R_{h*} \circ \underline{A} = \underline{A} \circ \mathrm{Ad}(h^{-1})\ \forall h \in H$.*
3. *$\underline{A}$ restricts to the fundamental vector fields on $\mathfrak{h}$: $\underline{A}(a) = \tilde{a}\ \forall a \in \mathfrak{h}$.*

Proof. These properties follow immediately from the definition 42.1.2 of the associated vector fields and the corresponding properties in definition 42.1.1 of a Cartan connection:

1. A Cartan connection defines for every $p \in P$ a linear isomorphism $A_p : T_pP \to \mathfrak{g}$. The associated vector fields are defined via its inverse, $\underline{A}_p = A_p^{-1}$, which is again a linear isomorphism.
2. Let $p \in P$, $h \in H$ and $v \in T_pP$. From the equivariance of the Cartan connection A follows

$$A(R_{h*}(v)) = ((R_h)^*(A))(v) = \mathrm{Ad}(h^{-1})(A(v)) . \tag{42.1.2}$$

Denoting $a = A(v) \in \mathfrak{g}$, and hence $v = \underline{A}_p(a)$, it then follows that

$$R_{h*}(\underline{A}_p(a)) = \underline{A}_{R_h(p)}(\mathrm{Ad}(h^{-1})(a)) \,, \tag{42.1.3}$$

which is the invariance condition for the associated vector fields.

3. The last condition follows immediately from the construction of $\underline{A}$ as the inverse of A. ■

The equivariance establishes a relation for the Cartan connection at different points belonging to the same fiber of P. It states that these are related by the group action in terms of right translations. Since these are generated by the fundamental vector fields, one also finds the following helpful statement.

Theorem 42.1.2. *The Lie derivative of the Cartan connection A and associated vector fields $\underline{A}$ with respect to the fundamental vector fields $\tilde{a}$ satisfies*

$$\mathcal{L}_{\tilde{a}}A = -[a, A] \,, \quad \mathcal{L}_{\tilde{a}}\underline{A}(b) = \underline{A}([a, b]) \tag{42.1.4}$$

for all $a \in \mathfrak{h}$ and $b \in \mathfrak{g}$.

Proof. ▶. . . ◀ ■

So far we have focused on the total space P of the principal bundle. We now turn our attention to the base manifold M, whose geometry we aim to describe using the Cartan connection. The first step towards this aim is the following important property.

Theorem 42.1.3. *For every Cartan geometry $(\pi : P \to M, A)$, there exists a canonical isomorphism between the vector bundles TM and $P \times_{\mathrm{Ad}} \mathfrak{g}/\mathfrak{h}$, where Ad denotes the adjoint representation of H on $\mathfrak{g}/\mathfrak{h}$.*

Proof. Let $p \in P$ and $w \in T_pP$, and denote $x = \pi(p) \in M$ and $v = \pi_*(w)$. The Cartan connection defines a unique element $a = A_p(w) \in \mathfrak{g}$, and hence a unique equivalence class $a + \mathfrak{h} \in \mathfrak{g}/\mathfrak{h}$. This further defines a unique element $[p, a + \mathfrak{h}] \in P \times_{\mathrm{Ad}} \mathfrak{g}/\mathfrak{h}$. We now study how this element depends on the choice of p and w. This will be done in two steps.

1. Let $w' \in T_pP$. Note that $\pi_*(w') = v = \pi_*(w)$ if and only if $w - w' \in V_pP$. Further, denote $a' = A_p(w')$, and note that $a' + \mathfrak{h} = a + \mathfrak{h}$, i.e., $a - a' \in \mathfrak{h}$, if and only if $w - w' \in V_pP$. Since the map $[p, \bullet] : \mathfrak{g}/\mathfrak{h} \to P_{\pi(p)} \times_{\mathrm{Ad}} \mathfrak{g}/\mathfrak{h}$ is a diffeomorphism, it follows that also $[p, a + \mathfrak{h}] = [p, a' + \mathfrak{h}]$ if and only if $w - w' \in V_pP$.

2. Let $\tilde{p} \in P$. Then $\pi(\tilde{p}) = x = \pi(p)$ if and only if there exists $h \in H$ such that $\tilde{p} = R_hp$. Further, $P_{\pi(\tilde{p})} \times_{\mathrm{Ad}} \mathfrak{g}/\mathfrak{h} = P_{\pi(p)} \times_{\mathrm{Ad}} \mathfrak{g}/\mathfrak{h}$ if and only if $\tilde{p} = R_hp$ for $h \in H$. Now denote $\tilde{w} = R_{h*}w$. Then we have

$$\pi_*(\tilde{w}) = \pi_*(R_{h*}w) = (\pi \circ R_h)_*(w) = \pi_*(w) = v \,. \tag{42.1.5}$$

It further follows from the equivariance of the Cartan connection that

$$\tilde{a} = A_{\tilde{p}}(\tilde{w}) = A_{p \cdot h}(R_{h*}w) = ((R_h)^*A)_p(w) = \mathrm{Ad}(h^{-1})(A_p(w)) = \mathrm{Ad}(h^{-1})(a) \,. \tag{42.1.6}$$

Hence, we have

$$[\tilde{p}, \tilde{a} + \mathfrak{h}] = [p \cdot h, \mathrm{Ad}(h^{-1})(a) + \mathfrak{h}] = [p, a + \mathfrak{h}] \,. \tag{42.1.7}$$

In summary, we find that for $\tilde{p} \in P$ and $w' \in T_{\tilde{p}}P$ we have that $[\tilde{p}, A_{\tilde{p}}(w') + \mathfrak{h}] = [p, A_p(w) + \mathfrak{h}]$ if and only if $\tilde{p} = R_p$ and $w' - R_{h*}w \in V_{\tilde{p}}P$, i.e., if and only if $\pi(p) = \pi(\tilde{p})$ and $\pi_*(w) = \pi_*(w')$. We thus have a one-to-one correspondence between elements of $P \times_{\mathrm{Ad}} \mathfrak{g}/\mathfrak{h}$ and vectors in TM, and this one-to-one correspondence preserves the base point $x \in M$. Moreover, it follows from the linearity of π_* and A_p that it also preserves the vector space structure on the fibers of these bundles, and hence constitutes a vector bundle isomorphism. ■

Example 42.1.1. Let M be a two-dimensional manifold, embedded in $\mathbb{R}^3$, and consider a sphere S^2, likewise embedded in $\mathbb{R}^3$. ▶. . . ◀

42.2 Curvature

Definition 42.2.1 (Curvature of a Cartan connection)**.** Let $(\pi : P \to M, A)$ be a Cartan geometry modeled on the Klein geometry (G, H). The *curvature* of the Cartan connection is the $\mathfrak{g}$-valued two-form

$$F = \mathrm{d}A + \frac{1}{2}[A \wedge A] \in \Omega^2(P, \mathfrak{g})\,. \tag{42.2.1}$$

Theorem 42.2.1. *The curvature F and associated vector fields $\underline{A}$ of a Cartan geometry $(\pi : P \to M, A)$ satisfy the relations*

$$\underline{A}([a, b]) - [\underline{A}(a), \underline{A}(b)] = \underline{A}(F(\underline{A}(a), \underline{A}(b))) \tag{42.2.2}$$

for all $a, b \in \mathfrak{g}$.

Proof. ▶. . . ◀ ■

Theorem 42.2.2. *The curvature F of a Cartan geometry $(\pi : P \to M, A)$ is horizontal, i.e., $\iota_X F = 0$ for any vertical vector field $X \in \Gamma(VP)$ on P.*

Proof. ▶. . . ◀ ■

Theorem 42.2.3 (Bianchi identity)**.** *The curvature F of a Cartan connection A satisfies $\mathrm{d}F = [F \wedge A]$.*

Proof. ▶. . . ◀ ■

42.3 First-order Cartan geometries

Definition 42.3.1 (First-order Cartan geometry)**.** A Cartan geometry with model Klein geometry (G, H) is called first-order Cartan geometry if the quotient representation of the adjoint representations H on $\mathfrak{g}/\mathfrak{h}$ is faithful. Otherwise, it is called higher-order Cartan geometry.

Definition 42.3.2 (Admissible frame)**.** Let $(\pi : P \to M, A)$ be a first-order Cartan geometry with model (G, H). An *admissible frame* over $x \in M$ is a vector space isomorphism $f : \mathfrak{g}/\mathfrak{h} \to T_xM$ such that $f = \varphi^{-1} \circ [p, \bullet]$ for some $p \in P$, where $\varphi : TM \to P \times_{\mathrm{Ad}} \mathfrak{g}/\mathfrak{h}$ denotes the bundle isomorphism from theorem 42.1.3.

Theorem 42.3.1. *The set Q of admissible frames f of a first-order Cartan geometry $(\pi : P \to M, A)$ forms a principal H-bundle over M with right action $f \cdot h = \varphi^{-1} \circ [p \cdot h, \bullet]$ for $f = \varphi^{-1} \circ [p, \bullet]$, which is canonically isomorphic to P.*

Proof. ▶. . . ◀ ■

42.4 Reductive Cartan geometries

Definition 42.4.1 (Reductive Cartan geometry). A Cartan geometry with model Klein geometry (G, H) is called *reductive* if the Lie algebra $\mathfrak{g}$ allows a decomposition of the form $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{z}$ into subrepresentations of the adjoint representation of H.

42.5 Cartan development

Definition 42.5.1 (Cartan development). Let $(\pi : P \to M, A)$ be a Cartan geometry modeled on (G, H) and $\phi : G \to G/H$ the canonical projection onto the coset space. The *Cartan development* of a curve $\gamma \in C^\infty(\mathbb{R}, P)$ through $z \in G/H$ is the unique curve $\chi \in C^\infty(\mathbb{R}, G/H)$ such that $\chi(0) = z$ and $\chi = \phi \circ \psi$, where $\psi \in C^\infty(\mathbb{R}, G)$ satisfies

$$\dot{\psi}(t) = L_{\psi(t)*} A(\dot{\gamma}(t)) \tag{42.5.1}$$

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

Theorem 42.5.1. *Let $\gamma_1, \gamma_2 \in C^\infty(\mathbb{R}, P)$ be two curves with $\pi \circ \gamma_1 = \pi \circ \gamma_2$. Then their Cartan developments $\chi_1, \chi_2 \in C^\infty(\mathbb{R}, G/H)$ through a common point $z \in G/H$ are the same.*

Proof. ▶. . . ◀ ■

Chapter 43

Complex geometry

43.1 Almost complex structures

So far we have considered manifolds as objects which locally inherit the geometric, and in particular differential, structure of real, Euclidean space. For several applications in physics it is helpful to consider a different approach, which allows introducing additional structure from the complex numbers. As is well known, the complex numbers $\mathbb{C}$ can be obtained by extending the real numbers $\mathbb{R}$ with an imaginary unit element i satisfying $i^2 = -1$, which acts by multiplication on complex numbers as $z \mapsto iz$. We have already encountered an object with similar properties in defining 37.2.3 of an almost complex structure in the context on non-linear connections on the tangent bundle. We can define a generalized version of this object as follows.

Definition 43.1.1 (Almost complex structure)**.** An *almost complex structure* on a manifold M is a rank $(1,1)$ tensor field J which satisfies $J \circ J = -\mathrm{id}_{TM}$ when regarded as a vector bundle endomorphism $J : TM \to TM$ covering the identity on M.

For every $p \in M$, the almost complex structure equips the tangent space T_pM with an automorphism $\xi \mapsto J\xi$ which squares to the identity, and is therefore indeed reminiscent of the multiplication by i on a complex vector space. Recall that, forgetting about this operation, a complex vector space may be regarded as a real vector space, which is of even dimension, and allows for a canonical orientation. We find that an almost complex structure induces similar properties.

Theorem 43.1.1. *A manifold equipped with an almost complex structure is even-dimensional and orientable.*

Proof. From the product rule 23.4.1 for the determinant of tensor densities follows

$$(\det J)^2 = \det(J \circ J) = \det(-\delta) = (-1)^{\dim M} . \tag{43.1.1}$$

Since $\det J$ is real, the right hand side must be non-negative, and so $\dim M$ must be even. Orientability will be proven alongside theorem 43.4.2. ∎

For the almost complex structure in the context of non-linear connections we have found a simple coordinate expression (37.2.19). There is no such canonical expression for generic almost complex structures; however, if we add another condition, we will find such an expression in the next section. To illustrate the notion, we give a simple example.

Example 43.1.1 (Almost complex structure on the sphere). Let

$$M = \{\vec{x} \in \mathbb{R}^3, \|\vec{x}\| = 1\} \cong S^2 \tag{43.1.2}$$

be the unit sphere, and note that its tangent bundle is described by the isomorphism

$$TM \cong \{(\vec{x}, \vec{u}) \in \mathbb{R}^3 \times \mathbb{R}^3, \|\vec{x}\| = 1, \vec{x} \cdot \vec{u} = 0\}. \tag{43.1.3}$$

Define

$$J : \begin{array}{ccc} TM & \to & TM \\ (\vec{x}, \vec{u}) & \mapsto & (\vec{x}, \vec{x} \times \vec{u}) \end{array}. \tag{43.1.4}$$

First, we check that its image indeed lies in TM. For this purpose, we write

$$(\vec{x}', \vec{u}') = J(\vec{x}, \vec{u}) = (\vec{x}, \vec{x} \times \vec{u}). \tag{43.1.5}$$

Now we have

$$\|\vec{x}'\| = \|\vec{x}\| = 1 \tag{43.1.6}$$

and

$$\vec{x}' \cdot \vec{u}' = \vec{x} \cdot (\vec{x} \times \vec{u}) = 0. \tag{43.1.7}$$

Further, the vector product is smooth and linear in the fiber component $\vec{u}$, and J covers the identity since $\vec{x}' = \vec{x}$. Hence, it is a vector bundle endomorphism covering the identity. We then calculate

$$\begin{aligned} J(J(\vec{x}, \vec{u})) &= J(\vec{x}, \vec{x} \times \vec{u}) \\ &= (\vec{x}, \vec{x} \times (\vec{x} \times \vec{u})) \\ &= (\vec{x}, (\vec{x} \cdot \vec{u})\vec{x} - (\vec{x} \cdot \vec{x})\vec{u}) \\ &= (\vec{x}, -\vec{u}) \\ &= -(\vec{x}, \vec{u}), \end{aligned} \tag{43.1.8}$$

where in the last line it is important to keep in mind that the vector space operations, and this taking the negative element, only act on the fiber part $\vec{u}$, but leave the base point $\vec{x}$ unchanged. Hence, J is an almost complex structure on M.

43.2 Complex vector bundles

We have seen that an almost complex structure equips each tangent space with a (real) linear function whose double application sends each tangent vector to its negative, similar to the multiplication by the imaginary unit i in a complex vector space. In the following, it will thus be useful to consider vector bundles whose fibers are not real, but complex vector spaces. Recall from definition 3.1.1 that we demanded the local trivializations of a vector bundle to restrict to linear functions on every fiber, in order for the linear and differentiable structures on the vector bundle to be compatible with each other, so that, e.g., the sections of a vector bundle form again a vector space. We demand a similar property also for complex vector bundles, but in this case taking the fibers to be *complex* vector spaces, and we demand that the *complex* linear structure is compatible with the differentiable one. To achieve this, the local trivializations must restrict to *complex* linear functions on every fiber. We make this precise as follows.

Definition 43.2.1 (Complex vector bundle). A *complex vector bundle* of rank $k \in \mathbb{N}$ is a fiber bundle $(E, B, \pi, \mathbb{C}^k)$ such that for all $p \in B$ the fiber $E_p = \pi^{-1}(p)$ is a complex vector space of complex dimension k and such that the restrictions of the local trivializations

$\phi : \pi^{-1}(U) \to U \times \mathbb{C}^k$ to a fiber E_p for $p \in U$ are complex vector space isomorphisms from E_p to $\{p\} \times \mathbb{C}^k$.

We will encounter various examples of complex vector bundles in the following sections. As for real vector bundles, we also need a notion of morphisms between complex vector bundles which preserve both the bundle structure and the complex vector space structure on each fiber. It is thus straightforward to define them as follows, in full analogy to the real case.

Definition 43.2.2 (Complex vector bundle morphism)**.** Let $(E_1, B_1, \pi_1, \mathbb{C}^{k_1})$ and $(E_2, B_2, \pi_2, \mathbb{C}^{k_2})$ be complex vector bundles. A *complex vector bundle morphism* (or *complex vector bundle homomorphism*) is a bundle morphism $\theta : E_1 \to E_2$ covering a map $\vartheta : B_1 \to B_2$ such that for each $p \in B_1$ the restriction of θ to the fiber $\pi_1^{-1}(p)$ is a complex linear function between the complex vector spaces $\pi_1^{-1}(p)$ and $\pi_2^{-1}(\vartheta(p))$.

Finally, as in the case of real vector bundle morphisms, we can define an isomorphism as an invertible morphism.

Definition 43.2.3 (Complex vector bundle isomorphism)**.** A *complex vector bundle isomorphism* is a bijective complex vector bundle morphism whose inverse is also a complex vector bundle morphism. If a complex vector bundle morphism between two complex vector bundles exists, these bundles are called *isomorphic*.

43.3 Complexification of real vector bundles

We have already learned that every manifold is naturally equipped with several (real) vector bundles. To discuss the case of complex vector bundles, we now show that every real vector bundle gives rise to a complex one by means of the following construction.

Definition 43.3.1 (Complexification of a vector bundle)**.** Let $(E, B, \pi, \mathbb{R}^k)$ be a real vector bundle of rank k. Its *complexification* is the vector bundle $E^{\mathbb{C}} = E \otimes \mathbb{C}$, whose fibers are the vector spaces $E_p^{\mathbb{C}} = E_p \otimes \mathbb{C}$ for all $p \in B$

To illustrate this definition, let $(\epsilon_\mu, \mu = 1, \ldots, k)$ be a local basis of E. We may regard $\mathbb{C}$ as a real, two-dimensional vector space with basis $(1, i)$. On the tensor product, regarded as a real vector bundle of rank $2k$, we may thus use the induced basis $(\epsilon_1, \ldots, \epsilon_k, i\epsilon_1, \ldots, i\epsilon_k)$. For a vector $w = u^\mu \epsilon_\mu + iv^\mu \epsilon_\mu \in E$ expressed in this basis, we define multiplication with a complex number $z = x + iy \in \mathbb{C}$ in the natural way as

$$zw = (xu^\mu - yv^\mu)\epsilon_\mu + i(xv^\mu + yu^\mu)\epsilon_\mu \,. \tag{43.3.1}$$

Alternatively, one may regard the tensor product as the direct sum

$$E \otimes \mathbb{C} \cong E \oplus iE \,, \tag{43.3.2}$$

whose basis is likewise given by $(\epsilon_1, \ldots, \epsilon_k, i\epsilon_1, \ldots, i\epsilon_k)$, and the complex multiplication is defined as above. One may expect that this construction yields a complex vector bundle. We prove this more formally.

Theorem 43.3.1. *The complexification $E^{\mathbb{C}}$ of a real vector bundle $(E, B, \pi, \mathbb{R}^k)$ of rank k is a complex vector bundle $(E^{\mathbb{C}}, B, \pi^{\mathbb{C}}, \mathbb{C}^k)$ of rank k.*

Proof. We have to construct local trivializations of $E^{\mathbb{C}}$, and show that these restrict to complex linear functions on every fiber. Let $\phi : \pi^{-1}(U) \to U \times \mathbb{R}^k$ be a local trivialization of E over $U \subset B$. Consider $\mathbb{C}^k \cong \mathbb{R}^k \otimes \mathbb{C}$, and let $\phi^{\mathbb{C}} : \pi^{\mathbb{C}\,-1}(U) \to U \times \mathbb{C}^k$ be defined via the tensor product, i.e., for $w = u + iv \in E_p^{\mathbb{C}}$ with $u, v \in E_p$ for $p \in U$ define

$$\phi^{\mathbb{C}}(w) = \phi^{\mathbb{C}}(u + iv) = \phi(u) + i\phi(v) \in \{p\} \times \mathbb{C}^k\,. \tag{43.3.3}$$

Then this preserves the complex multiplication on every fiber, since

$$\begin{aligned} \phi^{\mathbb{C}}(zw) &= \phi^{\mathbb{C}}(xu - yv + ixv + iyu) \\ &= \phi^{\mathbb{C}}((x + iy)(u + iv)) \\ &= \phi(xu - yv) + i\phi(xv + yu) \\ &= x\phi(u) - y\phi(v) + i[x\phi(v) + y\phi(u)] \\ &= (x + iy)[\phi(u) + i\phi(v)] \\ &= z\phi^{\mathbb{C}}(w) \end{aligned} \tag{43.3.4}$$

for all $z = x + iy \in \mathbb{C}$, and hence restricts to complex linear functions. ■

Similarly to the complexification of a real vector space, also the complexification of a real vector bundle carries an additional structure due to the decomposition (43.3.2) as a tensor product or a direct sum, which is not present for general complex vector bundles. This additional structure is induced by the vector bundle homomorphisms we found for direct sum bundles in section 4.2 - in particular, the projection maps. These allow us to define the following.

Definition 43.3.2 (Real and imaginary part)**.** Let $E^{\mathbb{C}}$ be the complexification of a real vector bundle. For each $w = u + iv \in E^{\mathbb{C}}$, the *real and imaginary part* are defined via the projection maps

$$\begin{array}{rccc} \Re : & E^{\mathbb{C}} & \to & E \\ & u + iv & \mapsto & u \end{array} \tag{43.3.5}$$

and

$$\begin{array}{rccc} \Im : & E^{\mathbb{C}} & \to & E \\ & u + iv & \mapsto & v \end{array}\,. \tag{43.3.6}$$

Making also use of the inclusion maps, we further have the following.

Definition 43.3.3 (Complex conjugate)**.** Let $E^{\mathbb{C}}$ be the complexification of a real vector bundle. For each $w = u + iv \in E^{\mathbb{C}}$, the *complex conjugate* is defined via the map

$$\begin{array}{rccc} \bar{\bullet} : & E^{\mathbb{C}} & \to & E^{\mathbb{C}} \\ & u + iv & \mapsto & u - iv \end{array}\,. \tag{43.3.7}$$

It follows immediately from their definition, as well as theorems 4.2.2 and 4.2.3, that these maps are real vector bundle homomorphisms, the last one being a real vector bundle isomorphism. However, note that it is *not* a complex vector bundle isomorphism, since it is not complex linear on each fiber, but antilinear. Finally, it is clear that we can also compose these maps with a section $\sigma : B \to E^{\mathbb{C}}$, and thus define the real and imaginary part as well as the complex conjugate of any section as

$$\Re\sigma = \Re \circ \sigma\,, \quad \Im\sigma = \Im \circ \sigma\,, \quad \bar{\sigma} = \bar{\bullet} \circ \sigma\,, \tag{43.3.8}$$

by a simplification of notation.

43.4 Complex frame bundles

Definition 43.4.1 (Complex frame bundle)**.** Let M be a manifold of dimension $\dim M = 2n$ equipped with an almost complex structure J. A *complex frame* at $p \in M$ is a bijective $\mathbb{R}$-linear function $f : \mathbb{C}^n \to T_pM$ such that $J \circ f = f \circ (i\bullet)$. The set of all complex frames constitutes the *complex frame bundle* with projection mapping $f : \mathbb{C}^n \to T_pM$ to $p \in M$.

Theorem 43.4.1. *The complex frame bundle of a manifold of dimension* $2n$ *is a principal fiber bundle with structure group* $\mathrm{GL}(n, \mathbb{C})$*.*

Proof. We show that $\mathrm{GL}(n, \mathbb{C})$ acts freely and transitively on the fibers over $p \in M$:

- Let $f : \mathbb{C}^n \to T_pM$ be a complex frame at p and $A \in \mathrm{GL}(n, \mathbb{C})$. Then

  $$J((f \circ A)(v)) = J(f(Av)) = f(iAv) = f(A(iv)) = (f \circ A)(iv) \tag{43.4.1}$$

  for all $v \in \mathbb{C}^n$, since A is $\mathbb{C}$-linear, and so $f \circ A$ is a complex frame. Hence, $\mathrm{GL}(n, \mathbb{C})$ acts on the fibers.

- Let $f, f' : \mathbb{C}^n \to T_pM$ be complex frames at p. Then define $A = f^{-1} \circ f' : \mathbb{C}^n \to \mathbb{C}^n$ as the unique bijection satisfying $f' = f \circ A$. Observe that A is $\mathbb{R}$-linear, $A \in \mathrm{GL}(2n, \mathbb{R})$, since f and f' are $\mathbb{R}$-linear bijections. Further, A must be $\mathbb{C}$-linear, since

  $$A(iv) = f^{-1}(f'(iv)) = f^{-1}(J(f'(v))) = -f^{-1}(J^{-1}(f'(v))) = -\frac{1}{i}f^{-1}(f'(v)) = iAv\,. \tag{43.4.2}$$

  for all $v \in \mathbb{C}^n$, and so $A \in \mathrm{GL}(n, \mathbb{C})$. This element is unique, and thus the action is free and transitive.

To complete the proof, one still needs to construct the local trivializations of the complex frame bundle and show its bundle structure. We will not show this here, and remark that this can be done by using an atlas of M. ■

Theorem 43.4.2. *Let* M *be a manifold of dimension* $\dim M = 2n$*. There exists a one-to-one correspondence between almost complex structures on* M *and reductions of the frame bundle via the inclusion* $\mathrm{GL}(n, \mathbb{C}) \hookrightarrow \mathrm{GL}(2n, \mathbb{R})$*.*

Proof. ▶. . . ◀ ■

43.5 Complex structures

While the algebraic structure of an almost complex structure J already yields a number of useful properties, it is not yet sufficient in order to arrive at the notion of a *complex* geometry. This step requires another property, which relates the values of J at nearby points. This is captured in the following definition.

Definition 43.5.1 (Complex structure)**.** An almost complex structure J is called a *complex structure* if and only if it is integrable, i.e., if its Nijenhuis tensor N_J vanishes.

The importance of the integrability will become clear in the following section. For now we study a simple example.

Example 43.5.1 (Complex structure on the sphere). We consider again the almost complex structure J given on the sphere $M = S^2$ defined in example 43.1.1. To check the integrability of J, we make use of the charts constructed in example 1.2.1. Defining coordinates (v^1, v^2) as

$$(v^1, v^2) = \left(\frac{x^1}{1+x^3}, \frac{x^2}{1+x^3}\right) \tag{43.5.1}$$

on $M \setminus \{(0,0,-1)\}$, we have covered the sphere except for one point. The inverse mapping is given by

$$\vec{x} = (x^1, x^2, x^3) = \left(\frac{2v^1}{1+(v^1)^2+(v^2)^2}, \frac{2v^2}{1+(v^1)^2+(v^2)^2}, \frac{2}{1+(v^1)^2+(v^2)^2} - 1\right). \tag{43.5.2}$$

Using these coordinates to write a tangent vector as $w = w^1\partial_1 + w^2\partial_2$, we have

$$\vec{u} = (u^1, u^2, u^3) = \frac{\left(2(1-(v^1)^2+(v^2)^2)w^1 - 4v^1v^2w^2, 2(1+(v^1)^2-(v^2)^2)w^2 - 4v^1v^2w^1, -4(v^1w^1+v^2w^2)\right)}{\left(1+(v^1)^2+(v^2)^2\right)^2}. \tag{43.5.3}$$

One easily checks that indeed $\vec{x} \cdot \vec{u} = 0$. Further, one finds

$$\vec{u}' = \vec{x} \times \vec{u} = \frac{\left(-2(1-(v^1)^2+(v^2)^2)w^2 - 4v^1v^2w^1, 2(1+(v^1)^2-(v^2)^2)w^1 + 4v^1v^2w^2, 4(v^1w^2-v^2w^1)\right)}{\left(1+(v^1)^2+(v^2)^2\right)^2}, \tag{43.5.4}$$

which we can now also write as

$$w' = w'^1\partial_1 + w'^2\partial_2 = w^1\partial_2 - w^2\partial_1 . \tag{43.5.5}$$

Hence, we have

$$J = \partial_2 \otimes \mathrm{d}v^1 - \partial_1 \otimes \mathrm{d}v^2 . \tag{43.5.6}$$

With this coordinate expression for J, we can now calculate the Nijenhuis tensor N_J, making use of theorem 17.6.3. Taking a closer look at the formula (17.6.8), we see that the first term vanishes, since the commutator $[\partial_a, \partial_b]$ of coordinate vector fields vanishes. Also the last term vanishes, since the exterior derivative $\mathrm{dd}v^a$ of the basis one-forms vanishes. Finally, also the Lie derivatives $\mathcal{L}_{\partial_a}\mathrm{d}v^b$ vanish, and hence we find $N_J = 0$. Repeating this calculation using another chart which covers all points except $(0,0,1)$, defining coordinates

$$(\tilde{v}^1, \tilde{v}^2) = \left(\frac{x^1}{1-x^3}, \frac{x^2}{1-x^3}\right), \tag{43.5.7}$$

one finds that N_J vanishes everywhere. Thus, the almost complex structure J constructed on the sphere in example 43.1.1 is integrable, and thus a complex structure.

43.6 Complex manifolds

The objects defined in the previous sections are in some way reminiscent of the operation $z \mapsto iz$ on the complex vector space $\mathbb{C}^n$, which likewise squares to $-\mathrm{id}_{\mathbb{C}^n}$. It is therefore not surprising that there exists another, closely related approach, which allows to locally relate a manifold to the complex vector space $\mathbb{C}^n$. To see this, we start with a suitable definition.

Definition 43.6.1 (Complex manifold)**.** A *complex manifold* is a manifold M of even dimension $\dim M = 2n$ for which there exists an atlas $\mathcal{A}$, such that for any two charts (U_1, ϕ_1) and (U_2, ϕ_2), with $\tilde{\phi}_i : U_i \to \mathbb{C}^n$ via the identification $\mathbb{C}^n \cong \mathbb{R}^{2n}$, the transition functions

$$\tilde{\phi}_{12} : \tilde{\phi}_1(U_1 \cap U_2) \to \tilde{\phi}_2(U_1 \cap U_2)\,, \quad \tilde{\phi}_{12} = \tilde{\phi}_2|_{U_1 \cap U_2} \circ \tilde{\phi}_1^{-1}|_{\tilde{\phi}_1(U_1 \cap U_2)} \tag{43.6.1}$$

are holomorphic.

The complex charts $\tilde{\phi} : U \to \mathbb{C}^n$ define local complex coordinates $(z^a, a = 1, \ldots, n)$ on $U \subset M$, which can be decomposed as $z^a = x^a + iy^a$ to give a distinguished class of real coordinates (x^a, y^a). It is often useful to work either with the complex or distinguished real coordinates, as we will see later.

The relation between complex manifolds and the structures we have defined earlier in this chapter is given by the following important theorem.

Theorem 43.6.1 (Newlander-Nirenberg)**.** *There exists a one-to-one correspondence between complex manifolds and manifolds equipped with a complex structure.*

Proof. We first show that every complex manifold M naturally possesses a complex structure. For $p \in M$, let $\tilde{\phi} : U \to \mathbb{C}^n$ be a complex chart, and $\tilde{\phi}_* : TU \to T\mathbb{C}^n$ its differential. Note that $\tilde{\phi}_*$ establishes a bijection between the tangent spaces T_pM and $T_{\tilde{\phi}(p)}\mathbb{C}^n \cong \mathbb{C}^n$. The latter is a complex vector space, and hence allows multiplication by the imaginary unit i. For $\xi \in T_pM$, we may thus define $J(\xi) = \tilde{\phi}_*^{-1}(i\tilde{\phi}_*(\xi))$. Note that this definition is independent of the choice of the chart, since for any other chart $\tilde{\psi} : V \to \mathbb{C}^n$ the transition functions $\tilde{\psi} \circ \tilde{\phi}^{-1}|_{\tilde{\phi}(U \cap V)}$ and $\tilde{\phi} \circ \tilde{\psi}^{-1}|_{\tilde{\psi}(U \cap V)}$ are holomorphic, and hence

$$\tilde{\psi}_*^{-1}(i\tilde{\psi}_*(\xi)) = \blacktriangleright \ldots \blacktriangleleft \tag{43.6.2}$$

Doing this for every $p \in M$, we obtain a map $J : TM \to TM$, covering the identity on M. This map is $\mathbb{R}$-linear on every tangent space T_xM, since the differentials $\tilde{\phi}_*$ restrict to linear functions on every tangent space, and so J is a vector bundle morphism. Further, we have

$$(J \circ J)(\xi) = \tilde{\phi}_*^{-1}(i\tilde{\phi}_*(\tilde{\phi}_*^{-1}(i\tilde{\phi}_*(\xi)))) = \tilde{\phi}_*^{-1}(-\tilde{\phi}_*(\xi)) = -\xi\,, \tag{43.6.3}$$

and so $J \circ J = -\mathrm{id}_{TM}$, so that J constitutes an almost complex structure. Finally, its Nijenhuis tensor is given by

$$N_J = \blacktriangleright \ldots \blacktriangleleft \tag{43.6.4}$$

which shows that J is integrable, and hence a complex structure.

Conversely, let J be a complex structure on a manifold M of dimension $2n$. ►...◄ ■

The complex structure J can most easily be expressed using the distinguished coordinates (x^a, y^a) we introduced before. In the induced coordinate basis, the differential $\tilde{\phi}_*$ reads

$$\tilde{\phi}_* \left(\xi^a \frac{\partial}{\partial x^a} + \eta^a \frac{\partial}{\partial y^a} \right) = (\xi^a + i\eta^a) \mathbb{e}_a \,, \tag{43.6.5}$$

where $(\mathbb{e}_a)$ is the canonical basis of $\mathbb{C}^n$. This is obviously linear and bijective, and we have

$$J \left(\xi^a \frac{\partial}{\partial x^a} + \eta^a \frac{\partial}{\partial y^a} \right) = \tilde{\phi}_*^{-1}((i\xi^a - \eta^a)\mathbb{e}_a) = -\eta^a \frac{\partial}{\partial x^a} + \xi^a \frac{\partial}{\partial y^a} \,, \tag{43.6.6}$$

so that on the basis elements J yields

$$J \left(\frac{\partial}{\partial x^a} \right) = \frac{\partial}{\partial y^a} \,, \quad J \left(\frac{\partial}{\partial y^a} \right) = -\frac{\partial}{\partial x^a} \,. \tag{43.6.7}$$

In other words, we can write J as

$$J = \frac{\partial}{\partial y^a} \otimes \mathrm{d}x^a - \frac{\partial}{\partial x^a} \otimes \mathrm{d}y^a \,. \tag{43.6.8}$$

Now one easily calculates the Nijenhuis tensor

$$\begin{aligned} 2N_J &= [\![J, J]\!] \\ &= \left[\!\!\left[\frac{\partial}{\partial y^a} \otimes \mathrm{d}x^a, \frac{\partial}{\partial y^b} \otimes \mathrm{d}x^b \right]\!\!\right] + \left[\!\!\left[\frac{\partial}{\partial x^a} \otimes \mathrm{d}y^a, \frac{\partial}{\partial x^b} \otimes \mathrm{d}y^b \right]\!\!\right] \\ &\quad - \left[\!\!\left[\frac{\partial}{\partial x^a} \otimes \mathrm{d}y^a, \frac{\partial}{\partial y^b} \otimes \mathrm{d}x^b \right]\!\!\right] - \left[\!\!\left[\frac{\partial}{\partial y^a} \otimes \mathrm{d}x^a, \frac{\partial}{\partial x^b} \otimes \mathrm{d}y^b \right]\!\!\right] . \end{aligned} \tag{43.6.9}$$

To see that this vanishes, we make use of theorem 17.6.3. Note that each term in the formula (17.6.8) contains as a factor either the Lie bracket of the involved vector fields, the Lie derivative of one of the involved differential forms or their exterior derivative. The latter obviously vanishes in this case, since $\mathrm{d}\mathrm{d}x^a = \mathrm{d}\mathrm{d}y^a = 0$. The same holds for the Lie bracket of coordinate vector fields, and so

$$\left[\frac{\partial}{\partial x^a}, \frac{\partial}{\partial x^b} \right] = \left[\frac{\partial}{\partial y^a}, \frac{\partial}{\partial y^b} \right] = \left[\frac{\partial}{\partial x^a}, \frac{\partial}{\partial y^b} \right] = 0 \,. \tag{43.6.10}$$

For the Lie derivatives of $\mathrm{d}x^a$ and $\mathrm{d}y^a$, one can use Cartan's magic formula, and realize that the interior product with any of the given coordinate vector fields is constant, and so its exterior derivative vanishes. Hence, all terms in the Frölicher-Nijenhuis bracket vanish.

Example 43.6.1 (Riemann sphere). Consider the space

$$M = \mathbb{C}P^1 = \{[w, z], (w, z) \in \mathbb{C}^2\} \,, \tag{43.6.11}$$

where $[w, z]$ denotes the equivalence class

$$[w, z] = \{(\lambda w, \lambda z), \lambda \in \mathbb{C} \setminus \{0\}\} \,. \tag{43.6.12}$$

We can construct two complex charts as follows. First, we consider the chart (U_1, ϕ_1) with

$$\begin{array}{cccc} \phi_1 & : & U_1 & \to & \mathbb{C} \\ & & [w, z] & \mapsto & u = \frac{z}{w} \end{array} \,, \tag{43.6.13}$$

which is defined on the domain

$$U_1 = M \setminus \{[0, 1]\} \tag{43.6.14}$$

that lacks a single point. Another chart (U_2, ϕ_2) is given by

$$\begin{array}{cccc} \phi_2 & : & U_2 & \to & \mathbb{C} \\ & & [w, z] & \mapsto & \tilde{u} = \frac{w}{z} \end{array} \,, \tag{43.6.15}$$

and is defined on the domain

$$U_2 = M \setminus \{[1,0]\} . \tag{43.6.16}$$

These two charts cover M. On their overlap one finds the transition function

$$\begin{array}{rccc} \phi_2 \circ \phi_1^{-1} & : & \mathbb{C} \setminus \{0\} & \to & \mathbb{C} \setminus \{0\} \\ & & u & \mapsto & \tilde{u} = \frac{1}{u} \end{array} , \tag{43.6.17}$$

which is holomorphic. Hence, M is a complex manifold. To get an understanding of the geometry of this space, we write these coordinates as

$$u = x + iy , \quad \tilde{u} = \tilde{x} + i\tilde{y} , \tag{43.6.18}$$

and so the transition function becomes

$$\tilde{x} + i\tilde{y} = \frac{x - iy}{x^2 + y^2} . \tag{43.6.19}$$

With the identification

$$v^1 = x , \quad v^2 = y , \quad \tilde{v}^1 = \tilde{x} , \quad \tilde{v}^2 = -\tilde{y} , \tag{43.6.20}$$

we see that this is the transition function of the sphere, given in example 1.2.1. Also the induced complex structure (43.6.8) agrees with (43.5.6). This complex manifold is also known as the *Riemann sphere*.

43.7 Holomorphic maps

The complex structure allows the definition of a certain class of maps, which we will also use later.

Definition 43.7.1 (Holomorphic map)**.** Let M, N be complex manifolds. A *holomorphic map* from M to N is a function $f : M \to N$ such that for each point $p \in M$ exist charts (U, ϕ) of M and (V, χ) on N such that:

- $p \in U$ and $f(U) \subset V$.
- The function $\tilde{\chi} \circ f \circ \tilde{\phi}^{-1} : \tilde{\phi}(U) \to \tilde{\chi}(V)$ is holomorphic.

Note that, using the notation of definition 43.6.1, $\tilde{\phi}(U)$ and $\tilde{\chi}(V)$ are subsets of $\mathbb{C}^m$ and $\mathbb{C}^n$, where $\dim M = 2m$ and $\dim N = 2n$ are the real dimensions of M and N, and so there exists a well-defined notion of holomorphic functions between these sets. Denoting local coordinates by $z^a = x^a + iy^a$ on U and $w^\mu = u^\mu + iv^\mu$ on V, a function is holomorphic if and only if

$$\frac{\partial w^\mu}{\partial \bar{z}^a} = 0 , \tag{43.7.1}$$

or equivalently, if

$$\frac{\partial u^\mu}{\partial x^a} = \frac{\partial v^\mu}{\partial y^a} , \quad \frac{\partial u^\mu}{\partial y^a} = -\frac{\partial v^\mu}{\partial x^a} . \tag{43.7.2}$$

As in the case of real manifolds, one also considers a class of maps which yields an equivalence relation among complex manifolds, by the following definition.

Definition 43.7.2 (Biholomorphic map)**.** A holomorphic map $f : M \to N$ which is bijective and whose inverse $f^{-1} : N \to M$ is again a holomorphic map, is called a *biholomorphic map*.

43.8 Holomorphic vector bundles

A particularly interesting class of complex vector bundles can be defined if also the base manifold is a complex manifold, and the same holds for the total space. We define these as follows.

Definition 43.8.1 (Holomorphic vector bundle)**.** A *holomorphic vector bundle* of rank $k \in \mathbb{N}$ is a complex vector bundle $(E, B, \pi, \mathbb{C}^k)$ with complex manifolds E and B such that the projection π is a holomorphic map and such that the local trivializations $\phi : \pi^{-1}(U) \to U \times \mathbb{C}^k$ are biholomorphic maps.

Theorem 43.8.1. *A complex vector bundle $(E, B, \pi, \mathbb{C}^k)$ over a complex manifold B is holomorphic if and only if all transition functions are holomorphic.*

Proof. ►...◄ ■

Definition 43.8.2 (Holomorphic vector bundle morphism)**.** Let $(E_1, B_1, \pi_1, \mathbb{C}^{k_1})$ and $(E_2, B_2, \pi_2, \mathbb{C}^{k_2})$ be holomorphic vector bundles. A *holomorphic vector bundle morphism* (or *holomorphic vector bundle homomorphism*) is a complex vector bundle morphism $\theta : E_1 \to E_2$ which is also a holomorphic map.

Definition 43.8.3 (Holomorphic vector bundle isomorphism)**.** A *holomorphic vector bundle isomorphism* is a biholomorphic complex vector bundle isomorphism, i.e., a bijective holomorphic vector bundle morphism whose inverse is also a holomorphic vector bundle morphism. If a holomorphic vector bundle morphism between two holomorphic vector bundles exists, these bundles are called *isomorphic*.

43.9 Holomorphic tangent bundle

We now come to an important example for a holomorphic vector bundle, which we can construct explicitly. Recall from chapter 7 that every differentiable manifold M possesses a canonically defined vector bundle, which is the tangent bundle TM, and which is one of the most elementary structures in differential geometry. For a complex manifold, a role of similar importance is played by the holomorphic tangent bundle. In fact, it can also be defined given an *almost* complex structure, which is how we will proceed.

Definition 43.9.1 (Holomorphic tangent bundle)**.** Let M be a manifold equipped with an almost complex structure J. Its *holomorphic tangent bundle* is the complex vector bundle

$$T^{(1,0)}M = \{\Xi \in T^{\mathbb{C}}M, J(\Xi) = i\Xi\}\,, \tag{43.9.1}$$

while its *antiholomorphic tangent bundle* is the complex vector bundle

$$T^{(0,1)}M = \{\Xi \in T^{\mathbb{C}}M, J(\Xi) = -i\Xi\}\,. \tag{43.9.2}$$

The spaces of sections of these bundles will be denoted $\mathrm{Vect}^{(1,0)}(M) = \Gamma(T^{(1,0)}M)$ and $\mathrm{Vect}^{(0,1)}(M) = \Gamma(T^{(0,1)}M)$.

We then show a few properties of the holomorphic and antiholomorphic tangent bundles. By theorem 43.3.1, $T^{\mathbb{C}}M$ is a complex vector bundle of rank $\dim M = 2n$. Since the holomorphic and antiholomorphic tangent bundles are constructed from eigenvectors, one can further show the following.

Theorem 43.9.1. *The holomorphic and antiholomorphic tangent bundles of a manifold M of dimension $2n$ equipped with an almost complex structure J are complex vector bundles of complex rank n.*

Proof. For each $p \in M$ the eigenspaces of J in $T_p^{\mathbb{C}}M$ are again complex vector spaces, since J acts by complex linear extension, and so $J(z\Xi) = zJ(\Xi)$ for all $\Xi \in T_p^{\mathbb{C}}M$ and $z \in \mathbb{C}$. Hence, the fibers of $T^{(1,0)}M$ and $T^{(0,1)}M$ are complex vector spaces. Further, we need to construct the local trivializations and check that these are complex linear on each fiber. Let $\phi : \tau^{-1}(U) \to U \times \mathbb{R}^{2n}$ be a local trivialization of the tangent bundle $\tau : TM \to M$ defined on $U \subset M$ and $\phi^{\mathbb{C}} : \tau^{\mathbb{C}\,-1}(U) \to U \times \mathbb{C}^{2n}$ its complexification. Then we can define the real and imaginary part

$$\Re\phi^{\mathbb{C}}(\xi + i\eta) = \phi(\xi)\,, \quad \Im\phi^{\mathbb{C}}(\xi + i\eta) = \phi(\eta) \tag{43.9.3}$$

for $\xi, \eta \in T_pM$ with $p \in U$. Note that $\xi \in T^{(1,0)}M$ if and only if

$$0 = J(\xi + i\eta) - i(\xi + i\eta) = J(\xi) + \eta + i(J(\eta) - \xi)\,. \tag{43.9.4}$$

Since the real and imaginary parts are independent, this is equivalent to $\eta = -J(\xi)$ and $\xi = J(\eta)$, which are equivalent conditions, since $J \circ J = -\mathrm{id}_{TM}$. Similarly, one finds $\xi + i\eta \in T^{(0,1)}M$ if and only if

$$0 = J(\xi + i\eta) + i(\xi + i\eta) = J(\xi) - \eta + i(J(\eta) + \xi)\,. \tag{43.9.5}$$

►...◄ ■

Further, it follows from the condition $J \circ J = -\mathrm{id}_{TM}$ on an almost complex structure that these two bundles together yield the complexified tangent bundle as follows.

Theorem 43.9.2. *The complexified tangent bundle of a manifold M equipped with an almost complex structure J is given by $T^{\mathbb{C}}M = T^{(1,0)}M \oplus T^{(0,1)}M$.*

Proof. ►...◄ ■

The construction of the holomorphic and antiholomorphic tangent bundle follows a curious principle. First, we complexify the tangent bundle, hence combining two copies of this real vector bundle into a complex vector bundle by taking their direct sum $TM \oplus iTM$. Then we decompose this newly created vector bundle into a different direct sum $T^{(1,0)}M \oplus T^{(0,1)}M$, where now both parts are complex vector bundles. Viewing them as real vector bundles again, one might expect that these are still isomorphic to the original bundle, i.e., the tangent bundle TM. We now check that this is the case.

Theorem 43.9.3. *The maps $\varphi = \pi_i \circ \iota : TM \to T^{(1,0)}M$ and $\bar{\varphi} = \pi_{-i} \circ \iota : TM \to T^{(0,1)}M$, where $\iota : TM \hookrightarrow T^{\mathbb{C}}M$ is the canonical inclusion and $\pi_{\pm i}$ are the projectors onto the eigenspaces of J, are (real) vector bundle isomorphisms, such that the diagrams*

$$\begin{array}{ccc} TM & \xrightarrow{\varphi} & T^{(1,0)}M \\ {\scriptstyle J}\downarrow & & \downarrow{\scriptstyle i\cdot\bullet} \\ TM & \xrightarrow[\varphi]{} & T^{(1,0)}M \end{array} \tag{43.9.6}$$

and

$$\begin{array}{ccc} TM & \xrightarrow{\bar{\varphi}} & T^{(0,1)}M \\ {\scriptstyle J}\downarrow & & \downarrow{\scriptstyle -i\cdot\bullet} \\ TM & \xrightarrow[\bar{\varphi}]{} & T^{(0,1)}M \end{array} \tag{43.9.7}$$

commute.

Proof. Using the general formula for projectors onto eigenspaces, one calculates

$$\pi_i = \frac{J + i\mathrm{id}_{T^{\mathbb{C}}M}}{2i} = \frac{\mathrm{id}_{T^{\mathbb{C}}M} - iJ}{2} \tag{43.9.8}$$

and

$$\pi_{-i} = \frac{J - i\mathrm{id}_{T^{\mathbb{C}}M}}{-2i} = \frac{\mathrm{id}_{T^{\mathbb{C}}M} + iJ}{2}\,, \tag{43.9.9}$$

so that

$$\begin{array}{cccc} \varphi & : & TM & \to & T^{(1,0)}M \\ & & \xi & \mapsto & \frac{\xi - iJ(\xi)}{2} \end{array} \tag{43.9.10}$$

and

$$\begin{array}{cccc} \bar{\varphi} & : & TM & \to & T^{(0,1)}M \\ & & \xi & \mapsto & \frac{\xi + iJ(\xi)}{2} \end{array}\,. \tag{43.9.11}$$

One finds that the images of these maps indeed lie in the respective bundle, since

$$J(\varphi(\xi)) = \frac{J(\xi) - iJ^2(\xi)}{2} = \frac{J(\xi) + i\xi}{2} = i\varphi(\xi) \tag{43.9.12}$$

and

$$J(\bar{\varphi}(\xi)) = \frac{J(\xi) + iJ^2(\xi)}{2} = \frac{J(\xi) - i\xi}{2} = -i\bar{\varphi}(\xi)\,. \tag{43.9.13}$$

These maps are smooth, preserve the fibers and act linearly on them. To see that they are vector bundle isomorphisms, and hence bijective, note that an element of $T^{\mathbb{C}}M$ is of the general form $\xi + i\eta$, where $\xi, \eta \in TM$, and that it belongs to $T^{(1,0)}M$ if and only if

$$0 = J(\xi + i\eta) - i(\xi + i\eta) = J(\xi) + \eta + i(J(\eta) - \xi)\,. \tag{43.9.14}$$

Since the real and imaginary parts are independent, this is equivalent to $\eta = -J(\xi)$ and $\xi = J(\eta)$, which are equivalent conditions, since $J \circ J = -\mathrm{id}_{TM}$. Hence, $\xi + i\eta \in T^{(1,0)}M$ if and only if it is of the form $\varphi(\xi)$ for some $\xi \in TM$. Similarly, one finds $\xi + i\eta \in T^{(0,1)}M$ if and only if

$$0 = J(\xi + i\eta) + i(\xi + i\eta) = J(\xi) - \eta + i(J(\eta) + \xi)\,, \tag{43.9.15}$$

so that it must be of the form $\bar{\varphi}(\xi)$ for some $\xi \in TM$. Finally, commutativity of the diagrams follows from

$$\varphi(J(\xi)) = \frac{J(\xi) - iJ^2(\xi)}{2} = \frac{J(\xi) + i\xi}{2} = i\varphi(\xi) \tag{43.9.16}$$

and

$$\bar{\varphi}(J(\xi)) = \frac{J(\xi) + iJ^2(\xi)}{2} = \frac{J(\xi) - i\xi}{2} = -i\bar{\varphi}(\xi)\,. \tag{43.9.17}$$

■

For the properties above we have only assumed that J is an *almost* complex structure, i.e., we have made no assumptions on its integrability. We now pose the question whether the holomorphic tangent bundle also possesses properties, which hold only if J is also integrable, and hence a complex structure. Indeed, we find a property which is even in one-to-one correspondence with the integrability of J, which we see as follows.

Theorem 43.9.4. *An almost complex structure J is integrable, $N_J = 0$, if and only if any of the following equivalent conditions holds:*

1. $[X, Y] \in \mathrm{Vect}^{(1,0)}(M)$ *for all* $X, Y \in \mathrm{Vect}^{(1,0)}(M)$.
2. $[X, Y] \in \mathrm{Vect}^{(0,1)}(M)$ *for all* $X, Y \in \mathrm{Vect}^{(0,1)}(M)$.

Proof. By theorem 17.6.6 we have for all $X, Y \in \mathrm{Vect}^{(1,0)}(M)$

$$\begin{aligned} N_J(X,Y) &= [JX,JY] + J^2[X,Y] - J([JX,Y] + [X,JY]) \\ &= [iX,iY] - [X,Y] - J([iX,Y] - [X,iY]) \\ &= -2i(J[X,Y] - i[X,Y])\,, \end{aligned} \tag{43.9.18}$$

which vanishes if and only if $J[X,Y] = i[X,Y]$, hence $[X,Y] \in \mathrm{Vect}^{(1,0)}(M)$. Analogously, for $X, Y \in \mathrm{Vect}^{(0,1)}(M)$ we have

$$\begin{aligned} N_J(X,Y) &= [JX,JY] + J^2[X,Y] - J([JX,Y] + [X,JY]) \\ &= [-iX,-iY] - [X,Y] - J([-iX,Y] - [X,-iY]) \\ &= 2i(J[X,Y] + i[X,Y])\,, \end{aligned} \tag{43.9.19}$$

which vanishes if and only if $J[X,Y] = -i[X,Y]$, hence $[X,Y] \in \mathrm{Vect}^{(0,1)}(M)$. ∎

Recall from theorem 43.6.1 that J is integrable if and only if M is a complex manifold. Given local complex coordinates ($z^a = x^a + iy^a$), one can easily construct bases for the holomorphic and antiholomorphic tangent bundles. From the coordinate expression (43.6.8) of the complex structure, which also holds for its extension to the complexified tangent bundle $T^{\mathbb{C}}M$, one finds

$$J\left(\frac{\partial}{\partial x^a} - i\frac{\partial}{\partial y^a}\right) = i\frac{\partial}{\partial x^a} + \frac{\partial}{\partial y^a} = i\left(\frac{\partial}{\partial x^a} - i\frac{\partial}{\partial y^a}\right)\,, \tag{43.9.20a}$$

$$J\left(\frac{\partial}{\partial x^a} + i\frac{\partial}{\partial y^a}\right) = -i\frac{\partial}{\partial x^a} + \frac{\partial}{\partial y^a} = -i\left(\frac{\partial}{\partial x^a} + i\frac{\partial}{\partial y^a}\right)\,. \tag{43.9.20b}$$

Hence, defining

$$\frac{\partial}{\partial z^a} = \frac{1}{2}\left(\frac{\partial}{\partial x^a} - i\frac{\partial}{\partial y^a}\right)\,, \quad \frac{\partial}{\partial \bar{z}^a} = \frac{1}{2}\left(\frac{\partial}{\partial x^a} + i\frac{\partial}{\partial y^a}\right)\,, \tag{43.9.21}$$

we see that the former constitutes a basis of $T^{(1,0)}M$, while the latter constitutes a basis of $T^{(0,1)}M$. This allows us to express sections of these bundles using (complex) component functions Z^a. The reason for the particular choice of basis elements will become clear when we discuss the dual bundles.

In theorem 43.9.1 we have shown that the holomorphic tangent bundle and its antiholomorphic counterpart are complex vector bundles. In the case of a complex manifold, an even stronger statement holds, albeit only for the holomorphic tangent bundle, which justifies its nomenclature.

Theorem 43.9.5. *The holomorphic tangent bundle $T^{(1,0)}M$ of a complex manifold M is a holomorphic vector bundle.*

Proof. ▶...◀ ■

Since the total space of a holomorphic vector bundle is again a complex manifold, one may consider holomorphic sections. In the case of the holomorphic tangent bundle, these deserve their own name.

Definition 43.9.2 (Holomorphic vector field)**.** A *holomorphic vector field* on a complex manifold M is a holomorphic section $Z : M \to T^{(1,0)}M$ of the holomorphic tangent bundle $T^{(1,0)}M$.

For practical purposes, it is also helpful to derive the coordinate expression of holomorphic vector fields. ▶...◀

43.10 Complex differential forms

In analogy to the tangent bundle, one may also decompose the cotangent bundle into holomorphic and antiholomorphic subbundles, to be defined as follows. Note that again we start by assuming only an *almost* complex structure, which is sufficient to define some of the notions we discuss here, and only later come to the special case of a complex structure, i.e., consider integrability.

Definition 43.10.1 (Holomorphic cotangent bundle)**.** Let M be a manifold equipped with an almost complex structure J, and denote by $J^* : T^*M \to T^*M$ the dual vector bundle morphism defined such that

$$\langle J(\xi), \alpha\rangle = \langle \xi, J^*(\alpha)\rangle \tag{43.10.1}$$

for all $(\xi, \alpha) \in TM \times_M T^*M$. The *holomorphic cotangent bundle* is the complex vector bundle

$$T^{*(1,0)}M = \{\alpha \in T^{*\mathbb{C}}M, J^*(\alpha) = i\alpha\}\,, \tag{43.10.2}$$

while its *antiholomorphic cotangent bundle* is the complex vector bundle

$$T^{*(0,1)}M = \{\alpha \in T^{*\mathbb{C}}M, J^*(\alpha) = -i\alpha\}\,, \tag{43.10.3}$$

One may expect that the holomorphic and antiholomorphic cotangent bundles share some of the properties of the similarly defined tangent bundle analogous. We start with the following, elementary property.

Theorem 43.10.1. *The holomorphic and antiholomorphic cotangent bundles of a manifold M of dimension $2n$ equipped with an almost complex structure J are complex vector bundles of complex rank n.*

Proof. ▶...◀ ■

Next, we find that we can decompose the complexified cotangent bundle in full analogy to the tangent bundle.

Theorem 43.10.2. *The complexified cotangent bundle of a manifold M equipped with an almost complex structure J is given by $T^{*\mathbb{C}}M = T^{*(1,0)}M \oplus T^{*(0,1)}M$.*

Proof. ▶ . . . ◀ ■

The decomposition of the cotangent bundle is also inherited by its exterior powers. In other words, we can decompose the r-th exterior power of the complexified cotangent bundle into certain subbundles, which we define as follows.

Definition 43.10.2 (Complex differential form)**.** Let M be a manifold equipped with an almost complex structure J. A *complex differential* (p, q)*-form* is a section of the bundle

$$\Lambda^p T^{*(1,0)} M \otimes \Lambda^q T^{*(0,1)} M \,. \tag{43.10.4}$$

The space of all (p, q)-forms on M is denoted $\Omega^{(p,q)}(M)$.

This now leads us to the following decomposition.

Theorem 43.10.3. *The complexified exterior power bundle $\Lambda^k T^* M \otimes \mathbb{C}$ is given by the direct sum*

$$\Lambda^k T^* M \otimes \mathbb{C} = \bigoplus_{p+q=k} \Lambda^p T^{*(1,0)} M \otimes \Lambda^q T^{*(0,1)} M \,. \tag{43.10.5}$$

Proof. ▶ . . . ◀ ■

We also remark that the (complex) ranks of the vector bundles on the right hand side of the relation (43.10.5) are given by

$$\begin{aligned} \sum_{p+q=k} \binom{n}{p}\binom{n}{q} &= \sum_{j=0}^{k} \frac{n!}{j!(n-j)!} \frac{n!}{(k-j)!(n-k+j)!} \\ &= \blacktriangleright \ldots \blacktriangleleft \end{aligned} \tag{43.10.6}$$

Further, this decomposition allows us to generalize the action of the dual almost complex structure to complex differential forms of arbitrary rank. By an abuse of notation, we define it as follows.

▶ . . . ◀

43.11 Dolbeault operators

We have seen that at almost complex structure induces a decomposition of the space of complex differential forms based on theorem 43.10.3 into spaces which are generated by the holomorphic and antiholomorphic cotangent bundles. One may then ask how certain differential forms are decomposed, such as the exterior derivative of a complex differential form, i.e., whether the exterior covariant derivative is compatible with the decomposition. In order to study this question, we first define the following two operators.

Definition 43.11.1 (Dolbeault operators)**.** Let M be a manifold equipped with an almost complex structure J, and denote by $\pi^{p,q} : \Omega^{p+q}(M) \otimes \mathbb{C} \to \Omega^{(p,q)}(M)$ the projector onto the spaces of complex (p, q)-forms. On the space $\Omega^{(p,q)}(M)$ of complex (p, q)-forms the *Dolbeault operators* are given by

$$\partial = \pi^{p+1,q} \circ \mathrm{d} : \Omega^{(p,q)}(M) \to \Omega^{(p+1,q)}(M) \,, \tag{43.11.1a}$$

$$\bar{\partial} = \pi^{p,q+1} \circ \mathrm{d} : \Omega^{(p,q)}(M) \to \Omega^{(p,q+1)}(M) \,. \tag{43.11.1b}$$

Note that in the definition above we have *not* made any assumption on the sum of the Dolbeault operators, and we have not assumed anything about the image of the exterior derivative acting on a complex (p,q)-form. We only know that it increases the total degree $p+q$ of the differential form by one, but we have not made any statement on the individual holomorphic and antiholomorphic degrees p and q. In fact, it turns out that the latter is closely linked to the question whether the almost complex structure J is integrable, and hence a complex structure. This can be seen as follows.

Theorem 43.11.1. *Let M be a manifold equipped with an almost complex structure J. Then J is a complex structure if and only if the Dolbeault operators satisfy $\mathrm{d} = \partial + \bar{\partial}$.*

Proof. Note first that $\mathrm{d} = \partial + \bar{\partial}$ if and only if

$$\mathrm{d}\omega \in \Omega^{(p+1,q)}(M) \oplus \Omega^{(p,q+1)}(M) \tag{43.11.2}$$

for all $\omega \in \Omega^{(p,q)}(M)$. If it holds for general $p,q \in \mathbb{N}$, then in clearly holds for $p+q=1$. Conversely, if it holds for $p+q=1$, then we may consider ω of the form

$$\omega = f\,\zeta_1 \wedge \ldots \wedge \zeta_p \wedge \zeta'_1 \wedge \ldots \zeta'_q\,, \tag{43.11.3}$$

where $f \in \Omega^{(0,0)}(M)$, $\zeta_k \in \Omega^{(1,0)}(M)$ and $\zeta'_k \in \Omega^{(0,1)}(M)$, since $\Omega^{(p,q)}(M)$ is spanned by such terms. Clearly, we have $\mathrm{d}f \in \Omega^{1\mathbb{C}}(M) = \Omega^{(1,0)}(M) \oplus \Omega^{(0,1)}(M)$. Further, by the assumption, we have

$$\mathrm{d}\zeta_k \in \Omega^{(2,0)}(M) \oplus \Omega^{(1,1)}(M)\,, \quad \mathrm{d}\zeta'_k \in \Omega^{(1,1)}(M) \oplus \Omega^{(0,2)}(M)\,. \tag{43.11.4}$$

From the Leibniz rule thus follows that $\mathrm{d}\omega \in \Omega^{(p+1,q)}(M) \oplus \Omega^{(p,q+1)}(M)$. Hence, in the following it will be sufficient to consider $p+q=1$.

Let $\omega \in \Omega^{(0,1)}(M)$ and $X,Y \in \mathrm{Vect}^{(1,0)}(M)$, so that $\iota_X\omega = \iota_Y\omega = 0$. By $\mathbb{C}$-linear extension of theorem 9.4.2, and in particular the relation (9.4.5), we thus have

$$\iota_Y\iota_X\mathrm{d}\omega = X(\iota_Y\omega) - Y(\iota_X\omega) - \iota_{[X,Y]}\omega = -\iota_{[X,Y]}\omega\,. \tag{43.11.5}$$

The left hand side vanishes if and only if $\mathrm{d}\omega \in \Omega^{(1,1)}(M) \oplus \Omega^{(0,2)}(M)$, and so $\mathrm{d}\omega = \partial\omega + \bar{\partial}\omega$, while the right hand side vanishes if and only if $[X,Y] \in \mathrm{Vect}^{(1,0)}(M)$, hence, if and only if the almost complex structure on M is integrable, by theorem 43.9.4. The same formula holds also if we assume $\omega \in \Omega^{(1,0)}(M)$ and $X,Y \in \mathrm{Vect}^{(0,1)}(M)$ instead. ■

We thus find that for a (p,q)-form $\omega \in \Omega^{(p,q)}(M)$ we have $\mathrm{d}\omega = \partial\omega + \bar{\partial}\omega$, where $\partial\omega \in \Omega^{(p+1,q)}(M)$ and $\bar{\partial}\omega \in \Omega^{(p,q+1)}(M)$. Using this result, we find another helpful property of the Dolbeault operators on a complex manifold.

Theorem 43.11.2. *Let M be a manifold equipped with an almost complex structure J. Then J is a complex structure if and only if the Dolbeault operators satisfy any of, and hence all of, the relations $\partial^2 = \partial\bar{\partial} + \bar{\partial}\partial = \bar{\partial}^2 = 0$.*

Proof. Let $\omega \in \Omega^{(p,q)}(M)$ be a complex (p,q)-form, and assume that J is integrable, so that $\mathrm{d} = \partial + \bar{\partial}$. Then we have

$$0 = \mathrm{d}^2\omega = \partial^2\omega + (\partial\bar{\partial} + \bar{\partial}\partial)\omega + \bar{\partial}^2\omega\,, \tag{43.11.6}$$

where each term on the right hand side must vanish individually, since they belong to different subspaces $\Omega^{(p',q')}(M)$ with $(p',q') \in \{(p+2,q),(p+1,q+1),(p,q+2)\}$.

Conversely, assume that $\partial^2 = 0$. For any $f \in \Omega^{(0,0)}(M)$ and $X, Y \in \mathrm{Vect}^{(1,0)}(M)$ we then have

$$\begin{aligned} 0 &= \iota_Y \iota_X \partial^2 f \\ &= \iota_Y \iota_X \pi^{(2,0)} \mathrm{d}\partial f \\ &= \iota_Y \iota_X \mathrm{d}\partial f \\ &= X(\iota_Y \partial f) - Y(\iota_X \partial f) - \iota_{[X,Y]}\partial f \\ &= X(\iota_Y \mathrm{d}f) - Y(\iota_X \mathrm{d}f) - \iota_{[X,Y]}\partial f \\ &= \iota_Y \iota_X \mathrm{d}^2 f + \iota_{[X,Y]}\mathrm{d}f - \iota_{[X,Y]}\partial f \\ &= \iota_{[X,Y]}\bar{\partial} f\,, \end{aligned} \tag{43.11.7}$$

where we used the fact that $\mathrm{d}f = \partial f + \bar{\partial} f$, since f is a $(0,0)$-form, as well as $\iota_X \bar{\partial} f = \iota_Y \bar{\partial} f = 0$. Since f is arbitrary, it follows that $[X, Y] \in \mathrm{Vect}^{(1,0)}(M)$, and so J is integrable.

Starting from the assumption $\bar{\partial}^2 = 0$, one proceeds similarly, but with $X, Y \in \mathrm{Vect}^{(0,1)}(M)$. ■

To illustrate this result, it is helpful to express the complex differential forms using a coordinate basis of the cotangent bundle. For this purpose, let $z^a = x^a + iy^a$ be local complex coordinates. Considering at a point $p \in M$ the elements

$$\xi^a \frac{\partial}{\partial x^a} + \eta^a \frac{\partial}{\partial y^a} \in T_pM\,, \quad \alpha_a \mathrm{d}x^a + \beta_a \mathrm{d}y^a \in T_p^*M\,, \tag{43.11.8}$$

as well as the complex structure given by the relation (43.6.8), the canonical pairing between the tangent and cotangent bundles yields

$$\left\langle J\left(\xi^a \frac{\partial}{\partial x^a} + \eta^a \frac{\partial}{\partial y^a}\right), \alpha_a \mathrm{d}x^a + \beta_a \mathrm{d}y^a \right\rangle = \xi^a\beta_a - \eta^a\alpha_a = \left\langle \xi^a \frac{\partial}{\partial x^a} + \eta^a \frac{\partial}{\partial y^a}, J^*(\alpha_a \mathrm{d}x^a + \beta_a \mathrm{d}y^a)\right\rangle\,, \tag{43.11.9}$$

and so we find

$$J^* = \mathrm{d}x^a \otimes \frac{\partial}{\partial y^a} - \mathrm{d}y^a \otimes \frac{\partial}{\partial x^a}\,. \tag{43.11.10}$$

By extension to the complexified cotangent bundle, we thus find

$$J^*(\mathrm{d}x^a + i\mathrm{d}y^a) = i\mathrm{d}x^a - \mathrm{d}y^a = i(\mathrm{d}x^a + i\mathrm{d}y^a)\,, \tag{43.11.11a}$$

$$J^*(\mathrm{d}x^a - i\mathrm{d}y^a) = -i\mathrm{d}x^a - \mathrm{d}y^a = -i(\mathrm{d}x^a - i\mathrm{d}y^a)\,. \tag{43.11.11b}$$

Hence, we may define the basis elements

$$\mathrm{d}z^a = \mathrm{d}x^a + i\mathrm{d}y^a\,, \quad \mathrm{d}\bar{z}^a = \mathrm{d}x^a - i\mathrm{d}y^a\,, \tag{43.11.12}$$

with the former giving a basis of $T^{*(1,0)}M$, while the latter form a basis of $T^{*(0,1)}M$. Note that in contrast to the basis elements (43.9.21) of the holomorphic and antiholomorphic tangent bundles, we have not introduced a factor $\frac{1}{2}$ in this case. This has two reasons. The first reason is that one should be able to obtain the basis element $\mathrm{d}z^a$ by acting with the exterior derivative d on the coordinate function $z^a = x^a + iy^a$. The second reason is the canonical pairing, which reads

$$\left\langle \frac{\partial}{\partial z^a}, \mathrm{d}z^b \right\rangle = \frac{1}{2}\left\langle \frac{\partial}{\partial x^a} - i\frac{\partial}{\partial y^a}, \mathrm{d}x^b + i\mathrm{d}y^b \right\rangle = \delta_a^b\,, \tag{43.11.13a}$$

$$\left\langle \frac{\partial}{\partial z^a}, \mathrm{d}\bar{z}^b \right\rangle = \frac{1}{2}\left\langle \frac{\partial}{\partial x^a} - i\frac{\partial}{\partial y^a}, \mathrm{d}x^b - i\mathrm{d}y^b \right\rangle = 0\,, \tag{43.11.13b}$$

$$\left\langle \frac{\partial}{\partial \bar{z}^a}, \mathrm{d}z^b \right\rangle = \frac{1}{2}\left\langle \frac{\partial}{\partial x^a} + i\frac{\partial}{\partial y^a}, \mathrm{d}x^b + i\mathrm{d}y^b \right\rangle = 0\,, \tag{43.11.13c}$$

$$\left\langle \frac{\partial}{\partial \bar{z}^a}, \mathrm{d}\bar{z}^b \right\rangle = \frac{1}{2}\left\langle \frac{\partial}{\partial x^a} + i\frac{\partial}{\partial y^a}, \mathrm{d}x^b - i\mathrm{d}y^b \right\rangle = \delta_a^b\,, \tag{43.11.13d}$$

so that we have constructed the dual basis. One now easily checks that in this basis the complex structure is given in diagonal form in terms of its eigenvalues $\pm i$ as

$$\begin{aligned}
J &= i\left(\frac{\partial}{\partial z^a} \otimes \mathrm{d}z^a - \frac{\partial}{\partial \bar{z}^a} \otimes \mathrm{d}\bar{z}^a\right) \\
&= \frac{i}{2}\left(\frac{\partial}{\partial x^a} - i\frac{\partial}{\partial y^a}\right) \otimes (\mathrm{d}x^a + i\mathrm{d}y^a) - \frac{i}{2}\left(\frac{\partial}{\partial x^a} + i\frac{\partial}{\partial y^a}\right) \otimes (\mathrm{d}x^a - i\mathrm{d}y^a) \\
&= \frac{\partial}{\partial y^a} \otimes \mathrm{d}x^a - \frac{\partial}{\partial x^a} \otimes \mathrm{d}y^a \,.
\end{aligned} \tag{43.11.14}$$

Further, we find that a complex (p, q)-form $\omega \in \Omega^{(p,q)}(M)$ takes the form

$$\omega = \frac{1}{p!q!}\omega_{a_1\cdots a_p b_1\cdots b_q}\mathrm{d}z^{a_1} \wedge \ldots \wedge \mathrm{d}z^{a_p} \wedge \mathrm{d}\bar{z}^{b_1} \wedge \ldots \wedge \mathrm{d}\bar{z}^{b_q} \,, \tag{43.11.15}$$

where the coefficients are antisymmetric in each type of indices separately,

$$\omega_{a_1\cdots a_p b_1\cdots b_q} = \omega_{[a_1\cdots a_p][b_1\cdots b_q]} \,. \tag{43.11.16}$$

To express the exterior derivative in the complex coordinate basis, consider first a complex function $f \in C^\infty(M, \mathbb{C})$, for which holds

$$\begin{aligned}
\mathrm{d}f &= \frac{\partial f}{\partial x^a}\mathrm{d}x^a + \frac{\partial f}{\partial y^a}\mathrm{d}y^a \\
&= \left(\frac{\partial f}{\partial z^a} + \frac{\partial f}{\partial \bar{z}^a}\right)\frac{\mathrm{d}z^a + \mathrm{d}\bar{z}^a}{2} - \frac{1}{i}\left(\frac{\partial f}{\partial z^a} - \frac{\partial f}{\partial \bar{z}^a}\right)\frac{\mathrm{d}z^a - \mathrm{d}\bar{z}^a}{2i} \\
&= \frac{\partial f}{\partial z^a}\mathrm{d}z^a + \frac{\partial f}{\partial \bar{z}^a}\mathrm{d}\bar{z}^a \,,
\end{aligned} \tag{43.11.17}$$

so that the usual formula holds also in the complex coordinates, due to the particular definition of the coordinate bases. For a complex (p, q)-form $\omega \in \Omega^{(p,q)}(M)$ we thus find

$$\mathrm{d}\omega = \frac{1}{p!q!}\left(\frac{\partial\omega_{a_1\cdots a_p b_1\cdots b_q}}{\partial z^c}\mathrm{d}z^c + \frac{\partial\omega_{a_1\cdots a_p b_1\cdots b_q}}{\partial \bar{z}^c}\mathrm{d}\bar{z}^c\right) \wedge \mathrm{d}z^{a_1} \wedge \ldots \wedge \mathrm{d}z^{a_p} \wedge \mathrm{d}\bar{z}^{b_1} \wedge \ldots \wedge \mathrm{d}\bar{z}^{b_q} \,, \tag{43.11.18}$$

so that we read off the Dolbeault operators

$$\partial\omega = \frac{1}{p!q!}\frac{\partial\omega_{a_1\cdots a_p b_1\cdots b_q}}{\partial z^c}\mathrm{d}z^c \wedge \mathrm{d}z^{a_1} \wedge \ldots \wedge \mathrm{d}z^{a_p} \wedge \mathrm{d}\bar{z}^{b_1} \wedge \ldots \wedge \mathrm{d}\bar{z}^{b_q} \,, \tag{43.11.19a}$$

$$\bar{\partial}\omega = \frac{1}{p!q!}\frac{\partial\omega_{a_1\cdots a_p b_1\cdots b_q}}{\partial \bar{z}^c}\mathrm{d}\bar{z}^c \wedge \mathrm{d}z^{a_1} \wedge \ldots \wedge \mathrm{d}z^{a_p} \wedge \mathrm{d}\bar{z}^{b_1} \wedge \ldots \wedge \mathrm{d}\bar{z}^{b_q} \,. \tag{43.11.19b}$$

Definition 43.11.2 (Conjugate exterior derivative)**.** Let M be a manifold equipped with an almost complex structure J. The *conjugate exterior derivative* is the operator

$$\mathrm{d}^c = J^{-1} \circ \mathrm{d} \circ J = -J \circ \mathrm{d} \circ J \tag{43.11.20}$$

acting on (complex) differential forms on M.

Theorem 43.11.3. *Let M be a manifold equipped with an almost complex structure J. Then J is a complex structure if and only if the Dolbeault operators satisfy $\mathrm{d}^c = -i(\partial - \bar{\partial})$.*

Proof. ►. . . ◄ ■

Chapter 44

(Almost) Hermitian manifolds

44.1 Hermitian metrics

While (almost) complex manifolds by themselves already possess a rich geometric structure which is worth studying, they can also be equipped with additional structure. Demanding that this additional structure is compatible with the (almost) structure then leads to a number of new properties, which are not present if only one of them is given. An important example is that of a Riemannian (and thus positive definite) metric. The notion of compatibility we demand here is the following.

Definition 44.1.1 (Compatible metric)**.** Let M be manifold equipped with an almost complex structure J. A Riemannian metric g is called *compatible* with J if and only if

$$g(JX, JY) = g(X, Y) \tag{44.1.1}$$

for all vector fields $X, Y \in \mathrm{Vect}(M)$.

Hence, one demands that J preserves the metric, and thus constitutes an orthogonal endomorphism on the tangent bundle. To reveal the additional structure which can be obtained from this definition, we first show that it is actually equivalent to another structure we can define.

Theorem 44.1.1. *Let M be a manifold equipped with an almost complex structure J. Then there exists a one-to-one correspondence between compatible Riemannian metrics and non-degenerate two-forms $\omega \in \Omega^2(M)$ which satisfy*

$$\omega(JX, JY) = \omega(X, Y) \tag{44.1.2}$$

for all $X, Y \in \mathrm{Vect}(M)$ and $\omega(u, Ju) > 0$ for all non-zero $u \in TM$, which is given by

$$\omega(X, Y) = g(JX, Y) \quad \Leftrightarrow g(X, Y) = \omega(X, JY)\,. \tag{44.1.3}$$

Proof. ►...◄ ■

It is remarkable that under the condition of compatibility with the almost complex structure the same geometric structure can be defined either by a metric g, which is a symmetric tensor field, or a two-form ω, which is an antisymmetric tensor field. This is due to the fact that the compatibility conditions restrict the component of both tensor fields such that only a common subset of independent components remains, while all other components are uniquely determined from the compatibility. We can take this even further, and define the following.

Definition 44.1.2 (Hermitian metric)**.** Let M be a manifold. A *Hermitian metric* on M is a section h of the bundle $\mathrm{Sym}^2 T^{*\mathbb{C}}M$ such that for each $p \in M$, $h_p : T_p^{\mathbb{C}}M \times T_p^{\mathbb{C}}M \to \mathbb{C}$ is a positive definite Hermitian form on $T_p^{\mathbb{C}}M$, i.e., it satisfies

$$h_p(w, w') = \overline{h_p(w', w)} \tag{44.1.4}$$

for all $w, w' \in T_p^{\mathbb{C}}M$ and

$$h_p(w, w) > 0 \tag{44.1.5}$$

for all non-zero $w \in T_p^{\mathbb{C}}M$.

With this definition in place, we can show the following remarkable result.

Theorem 44.1.2. *Let M be a manifold equipped with an almost complex structure J. Then there exists a one-to-one correspondence between compatible Riemannian metrics g and Hermitian metrics h which satisfy*

$$h(JX, JY) = h(X, Y) \tag{44.1.6}$$

for all complex vector fields $X, Y \in \mathrm{Vect}^{\mathbb{C}}(M)$, which is given by complex linear extension of

$$h(X, Y) = g(X, Y) - ig(JX, Y)\,, \tag{44.1.7}$$

or conversely by

$$g(X, Y) = \Re h(X, Y)\,, \quad \omega(X, Y) = -\Im h(X, Y) \tag{44.1.8}$$

for all $X, Y \in \mathrm{Vect}(M)$.

Proof. ▶. . . ◀ ■

Definition 44.1.3 ((Almost) Hermitian manifold)**.** A manifold M equipped with an almost complex structure J and a Hermitian metric h is called an *almost Hermitian manifold*. If J is integrable, i.e., a complex structure, then it is called a *Hermitian manifold*.

44.2 Unitary frame bundle

44.3 Volume form

Definition 44.3.1 (Hermitian volume form)**.** Let M be an almost Hermitian manifold of dimension $2n$. Its *volume form* is the $(2n)$-form

$$\mathrm{vol}_\omega = \frac{1}{n!} \underbrace{\omega \wedge \ldots \wedge \omega}_{n \text{ times}} \in \Omega^{2n}(M)\,. \tag{44.3.1}$$

44.4 Chern connection

44.5 Differential forms on (almost) Hermitian manifolds

Definition 44.5.1 (Lefschetz operator)**.** The *Lefschetz operator* on an almost Hermitian manifold M is the operator $L : \Omega^{(p,q)}(M) \to \Omega^{(p+1,q+1)}(M)$ defined by $\alpha \mapsto L\alpha = \omega \wedge \alpha$, where ω is the associated two-form.

Definition 44.5.2 (Dual Lefschetz operator)**.** The *dual Lefschetz operator* on an almost Hermitian manifold M is the unique operator $\Lambda : \Omega^{(p,q)}(M) \to \Omega^{(p-1,q-1)}(M)$ satisfying

$$\langle \Lambda\alpha, \beta\rangle_g = \langle \alpha, L\beta\rangle_g \tag{44.5.1}$$

for all $\alpha \in \Omega^{(p,q)}(M)$ and $\beta \in \Omega^{p+q-2}(M)$.

Theorem 44.5.1. *The Lefschetz operator and its dual on an almost Hermitian manifold of dimension $2n$ satisfy*

$$L\Lambda\alpha - \Lambda L\alpha = (k-n)\alpha \tag{44.5.2}$$

for all $\alpha \in \Omega^k(M)$.

Proof. ►...◄ ■

Theorem 44.5.2. *The repeated application*

$$L^{n-k}\alpha \tag{44.5.3}$$

of the Lefschetz operator to a k-form $\alpha \in \Omega^k(M)$ induces a bijection between $\Omega^k(M)$ and $\Omega^{2n-k}(M)$ for $0 \le k \le n$, and

$$\Lambda^{k-n}\alpha \tag{44.5.4}$$

induces a bijection between $\Omega^k(M)$ and $\Omega^{2n-k}(M)$ for $n \le k \le 2n$.

Proof. ►...◄ ■

Definition 44.5.3 (Adjoint derivative operators)**.** ►$\mathrm{d}^*, \mathrm{d}^{c*}, \partial^*, \bar\partial^*$...◄

Definition 44.5.4 (Laplace operators)**.** The *Laplace operators* on an almost Hermitian manifold are defined by

$$\triangle_{\mathrm{d}} = \mathrm{d}\mathrm{d}^* + \mathrm{d}^*\mathrm{d}\,, \quad \triangle_{\partial} = \partial\partial^* + \partial^*\partial\,, \quad \triangle_{\bar\partial} = \bar\partial\bar\partial^* + \bar\partial^*\bar\partial\,. \tag{44.5.5}$$

Theorem 44.5.3. *The Laplace operators on a Hermitian manifold are related by*

$$\triangle_{\mathrm{d}} = 2\triangle_{\partial} = 2\triangle_{\bar\partial}\,. \tag{44.5.6}$$

Proof. ▶. . . ◀ ■

Theorem 44.5.4. *The Laplace operators on a Hermitian manifold satisfy*

$$[\triangle_{\mathrm{d}}, \mathrm{d}] = [\triangle_{\mathrm{d}}, \mathrm{d}^c] = [\triangle_{\mathrm{d}}, \mathrm{d}^*] = [\triangle_{\mathrm{d}}, \mathrm{d}^{c*}] = [\triangle_{\mathrm{d}}, \partial] = [\triangle_{\mathrm{d}}, \bar{\partial}] = [\triangle_{\mathrm{d}}, \partial^*] = [\triangle_{\mathrm{d}}, \bar{\partial}^*] = [\triangle_{\mathrm{d}}, \star] = 0\,. \tag{44.5.7}$$

Proof. ▶. . . ◀ ■

44.6 Kähler manifolds

A particular class of Hermitian manifolds which play an important role in different areas of physics, most prominently in string theory, are Kähler manifolds. Since Hermitian manifolds are characterized by the existence of several compatible geometric structures, there exists a number of equivalent definitions of Kähler manifolds. We will start from the following, most commonly encountered definition.

Definition 44.6.1 ((Almost) Kähler manifold)**.** An *almost Kähler manifold* is a manifold M equipped with a symplectic form ω (called the *Kähler form*) and an almost complex structure J such that $g(\bullet, \bullet) = \omega(\bullet, J\bullet)$ is a Riemannian metric. If J is integrable, i.e., a complex structure. then it is called a Kähler manifold.

This definition is reminiscent of one of the defining properties on an almost Hermitian manifold in terms of a non-degenerate two-form which is compatible with the almost complex structure, but differs in two aspects. First, recall from chapter 35 that a symplectic form is not only non-degenerate, but also closed, $\mathrm{d}\omega = 0$. Further, we did not require that ω is compatible with the almost complex structure, but we demanded the existence of a Riemannian metric. We will now show that this second aspect turns out to be an equivalent condition.

Theorem 44.6.1. *Let M be an almost Kähler manifold with symplectic form ω and metric g. Then both ω and g are compatible with the almost complex structure,*

$$\omega(JX, JY) = \omega(X, Y)\,, \quad g(JX, JY) = g(X, Y) \tag{44.6.1}$$

for all vector fields $X, Y \in \mathrm{Vect}(M)$.

Proof. Using the fact that a Riemannian metric is symmetric by definition and that $J^2 = -\mathrm{id}_{TM}$ we find

$$\begin{aligned} \omega(JX, JY) &= g(JX, Y) \\ &= g(Y, JX) \\ &= \omega(Y, J^2X) \\ &= -\omega(Y, X) \\ &= \omega(X, Y)\,. \end{aligned} \tag{44.6.2}$$

Similarly, from the antisymmetry of ω follows

$$\begin{aligned} g(JX, JY) &= \omega(JX, J^2Y) \\ &= -\omega(JX, Y) \\ &= \omega(Y, JX) \\ &= g(Y, X) \\ &= g(X, Y)\,. \quad \blacksquare \end{aligned} \tag{44.6.3}$$

With this result in mind, we can recall that g and ω are equivalent descriptions of an almost Hermitian manifold. We therefore see that indeed (almost) Kähler manifolds are special cases of (almost) Hermitian manifolds. The following is thus straightforward to show.

Theorem 44.6.2. *A (almost) Kähler manifold is a (almost) Hermitian manifold with a Hermitian metric h whose associated two-form is closed,* $\mathrm{d}\Im h = 0$.

Proof. ▶... ◀ ■

In the case of a Kähler manifold, we can also use the distinguished coordinates (x^a, y^a) to derive a coordinate expression for the Kähler form ω from definition 44.6.1. A general two-form is given by

$$\omega = \frac{1}{2}p_{ab}\mathrm{d}x^a \wedge \mathrm{d}x^b + q_{ab}\mathrm{d}x^a \wedge \mathrm{d}y^b + \frac{1}{2}r_{ab}\mathrm{d}y^a \wedge \mathrm{d}y^b\,, \tag{44.6.4}$$

where $p_{ab} = -p_{ba}$ and $r_{ab} = -r_{ba}$. Now one has

$$\begin{aligned} g\left(t^a\frac{\partial}{\partial x^a} + u^a\frac{\partial}{\partial y^a}, v^a\frac{\partial}{\partial x^a} + w^a\frac{\partial}{\partial y^a}\right) &= \omega\left(t^a\frac{\partial}{\partial x^a} + u^a\frac{\partial}{\partial y^a}, v^a\frac{\partial}{\partial y^a} - w^a\frac{\partial}{\partial x^a}\right) \\ &= u^a v^b r_{ab} - t^a w^b p_{ab} + (w^a u^b + t^a v^b)q_{ab}\,. \end{aligned} \tag{44.6.5}$$

We see that this is symmetric if and only if $p_{ab} = r_{ab}$ and $q_{ab} = q_{ba}$. Using the complex coordinate basis, we can then also write

$$\begin{aligned} \omega &= \frac{1}{2}p_{ab}(\mathrm{d}x^a \wedge \mathrm{d}x^b + \mathrm{d}y^a \wedge \mathrm{d}y^b) + q_{ab}\mathrm{d}x^a \wedge \mathrm{d}y^b \\ &= \frac{1}{8}p_{ab}[(\mathrm{d}z^a + \mathrm{d}\bar{z}^a) \wedge (\mathrm{d}z^b + \mathrm{d}\bar{z}^b) - (\mathrm{d}z^a - \mathrm{d}\bar{z}^a) \wedge (\mathrm{d}z^b - \mathrm{d}\bar{z}^b)] + \frac{1}{4i}q_{ab}(\mathrm{d}z^a + \mathrm{d}\bar{z}^a) \wedge (\mathrm{d}z^b - \mathrm{d}\bar{z}^b) \\ &= \frac{1}{2}p_{ab}\mathrm{d}z^a \wedge \mathrm{d}\bar{z}^b + \frac{i}{2}q_{ab}\mathrm{d}z^a \wedge \mathrm{d}\bar{z}^b \\ &= \frac{1}{2}(p_{ab} + iq_{ab})\mathrm{d}z^a \wedge \mathrm{d}\bar{z}^b\,, \\ &= \frac{i}{2}h_{ab}\mathrm{d}z^a \wedge \mathrm{d}\bar{z}^b\,. \end{aligned} \tag{44.6.6}$$

Keeping in mind that p_{ab} is antisymmetric and q_{ab} is symmetric, and both are real, we find that h_{ab} is Hermitian, $h_{ab} = \overline{h_{ba}}$. ▶... ◀

Note that we now have two integrability conditions, $N_J = 0$ and $\mathrm{d}\omega = 0$, which distinguish Kähler manifolds among almost Hermitian manifolds, as shown on the outer edges of figure 44.1. It turns out that demanding both conditions simultaneously can also equivalently be expressed in terms of a few other conditions, which we show as follows.

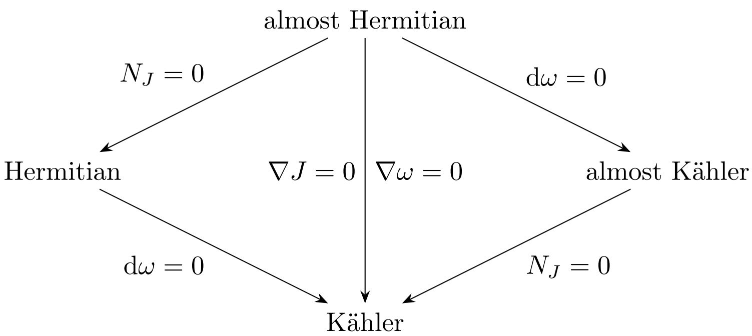


Figure 44.1: Conditions relating (almost) Hermitian and (almost) Kähler manifolds.

Theorem 44.6.3. *Let M be an almost Hermitian manifold of dimension $2n$. The following conditions are equivalent:*

1. *M is a Kähler manifold: $N_J = 0$ and $\mathrm{d}\omega = 0$.*
2. *The almost complex structure is covariantly constant: $\nabla J = 0$.*
3. *The associated two-form is covariantly constant: $\nabla\omega = 0$.*
4. *The holonomy group of ∇ is contained in the unitary group $\mathrm{U}(n)$.*

Proof. ►...◄ ■

As a consequence of the integrability conditions, Kähler manifolds have a few useful properties, which we will study next. The most important one is the existence of a potential for the Kähler form, similarly to the symplectic potential for a general symplectic form. In this case we can make use of the complex structure to define the following notion.

Definition 44.6.2 (Kähler potential)**.** Let M be a Kähler manifold with symplectic structure ω and $U \subset M$. A (local) *Kähler potential* on U is a function $\rho \in C^\infty(U, \mathbb{R})$ such that

$$\omega = \frac{i}{2}\partial\bar{\partial}\rho \tag{44.6.7}$$

on U.

This leads to the following statement.

Theorem 44.6.4. *Let M be a Kähler manifold with symplectic structure ω. Then for every $p \in M$ there exists an open neighborhood $U \subset M$ of p and a Kähler potential ρ on U satisfying (44.6.7).*

Proof. ►...◄ ■

Theorem 44.6.5 (Kähler identities)**.** *The Lefschetz operator and its adjoint and the derivative operators on a Kähler manifold satisfy the identities*

$$[\partial, L] = 0\,, \qquad [\partial, \Lambda] = -i\bar{\partial}^*\,, \tag{44.6.8a}$$
$$[\bar{\partial}, L] = 0\,, \qquad [\bar{\partial}, \Lambda] = i\partial^*\,, \tag{44.6.8b}$$
$$[\partial^*, L] = -i\bar{\partial}\,, \qquad [\partial^*, \Lambda] = 0\,, \tag{44.6.8c}$$
$$[\bar{\partial}, L] = i\partial\,, \qquad [\bar{\partial}, \Lambda] = 0\,, \tag{44.6.8d}$$
$$[\mathrm{d}, L] = 0\,, \qquad [\mathrm{d}, \Lambda] = \mathrm{d}^{c*}\,, \tag{44.6.8e}$$
$$[\mathrm{d}^c, L] = 0\,, \qquad [\mathrm{d}^c, \Lambda] = \blacktriangleright \ldots \blacktriangleleft\,, \tag{44.6.8f}$$
$$[\mathrm{d}^*, L] = \blacktriangleright \ldots \blacktriangleleft\,, \qquad [\mathrm{d}^*, \Lambda] = 0\,, \tag{44.6.8g}$$
$$[\mathrm{d}^{c*}, L] = \blacktriangleright \ldots \blacktriangleleft\,, \qquad [\mathrm{d}^{c*}, \Lambda] = 0\,, \tag{44.6.8h}$$
$$[\triangle_{\mathrm{d}}, L] = 0\,, \qquad [\triangle_{\mathrm{d}}, \Lambda] = 0\,. \tag{44.6.8i}$$

Proof. ►...◄ ■

Theorem 44.6.6. *The Dolbeault operators and their adjoints on a Kähler manifold satisfy*

$$\bar{\partial}^*\partial + \partial\bar{\partial}^* = 0\,, \quad \partial^*\bar{\partial} + \bar{\partial}\partial^* = 0\,. \tag{44.6.9}$$

Proof. ►...◄ ■

44.7 Calabi-Yau manifolds

Theorem 44.7.1. *Let M be a compact Kähler manifold of dimension $2n$. Then the following conditions are equivalent:*

1. *The line bundle $\Lambda^{2n}T^*M$ is trivial.*
2. *There exists a holomorphic n-form which is nowhere vanishing.*
3. *There exists a $\mathrm{SU}(n)$-reduction of the unitary frame bundle.*
4. *The holonomy group of ∇ is contained in the special unitary group $\mathrm{SU}(n)$.*

Proof. ►. . . ◄ ■

Definition 44.7.1 (Calabi-Yau manifold). A *Calabi-Yau manifold* is a compact Kähler manifold which satisfies any, and hence all, of the conditions given in theorem 44.7.1.

Theorem 44.7.2. *The Ricci curvature on a Calabi-Yau manifold vanishes.*

Proof. ►. . . ◄ ■

Chapter 45

Spin geometry

45.1 Clifford algebras

There are different ways to introduce the mathematical notions, in particular the necessary Lie groups, which allow the description of spinors. The most straightforward construction is based on the notion of a *Clifford algebra*. Here we use the following convention.

Definition 45.1.1 (Clifford algebra)**.** Let V be a vector space over a field $\mathbb{K}$ equipped with a non-degenerate quadratic form $Q : V \to \mathbb{K}$. The *Clifford algebra* $\mathrm{Cl}(V, Q)$ is the quotient algebra of the tensor algebra

$$\bigoplus_{k\in\mathbb{N}} \bigotimes^k V = \mathbb{K} \oplus V \oplus V \otimes V \oplus \ldots \tag{45.1.1}$$

and the two-sided ideal generated by the elements $v \otimes v - Q(v)\mathbb{1}$.

Note that also other conventions are used in the literature, in particular regarding the choice of the ideal:

- In [Har90], the ideal is spanned by the elements $v \otimes v + \langle v, v\rangle$.
- In [LM89], the opposite sign is used, and the ideal is spanned by $v \otimes v + Q(v)\mathbb{1}$.

The most common cases are given if V is a finite-dimensional real or complex vector space. In this case there is a canonical quadratic form, up to the choice of the signature, given by the symmetric bilinear form η. For the real case, we use the following convention.

Definition 45.1.2 (Real Clifford algebra)**.** The *real Clifford algebra* $\mathrm{Cl}_{k,l}(\mathbb{R})$ is the Clifford algebra $\mathrm{Cl}(V, Q)$, where $V = \mathbb{R}^{k+l}$ and

$$Q(v^\mu \mathbb{e}_\mu) = (v^1)^2 + \ldots + (v^k)^2 - (v^{k+1})^2 - \ldots - (v^{k+l})^2 = -\eta_{\mu\nu}v^\mu v^\nu = -\eta(v, v)\,. \tag{45.1.2}$$

Note that we define the quadratic form as $Q(v) = -\eta(v, v)$. The reason for this choice is that we have defined η such that it contains k negative and l positive signs, while in most of the literature on spin geometry the opposite convention is chosen. In the complex case, no such ambiguity exists, and one has the following unique definition.

Definition 45.1.3 (Complex Clifford algebra)**.** The *complex Clifford algebra* $\mathrm{Cl}_n(\mathbb{C})$ is the Clifford algebra $\mathrm{Cl}(V, Q)$, where $V = \mathbb{C}^n$ and

$$Q(v^\mu \mathbb{e}_\mu) = (v^1)^2 + \ldots + (v^n)^2 \,. \tag{45.1.3}$$

Taking into account that the tensor algebra used in the definition 45.1.1 is infinite-dimensional, one may wonder whether the Clifford algebras defined above are finite-dimensional. It turns out that this is indeed the case, and one finds their dimensions as follows.

Theorem 45.1.1. *The real and complex Clifford algebras are finite-dimensional real (complex) algebras of real (complex) dimension 2^{k+l} for $\mathrm{Cl}_{k,l}(\mathbb{R})$ and 2^n for $\mathrm{Cl}_n(\mathbb{C})$.*

Instead of a general, abstract proof, we will explicitly construct the bases of the aforementioned Clifford algebras from a basis $(\mathbb{e}_\mu)$ in which the quadratic form Q is given by the canonical form used in definitions 45.1.2 and 45.1.3. First note that we have

$$\begin{aligned} uv + vu &= (u+v)(u+v) - uu - vv \\ &= [Q(u+v) - Q(u) - Q(v)]\mathbb{1} \\ &= 2\langle u, v\rangle \mathbb{1} \end{aligned} \tag{45.1.4}$$

for $u, v \in V$, where we defined

$$\langle u, v\rangle = \frac{1}{2}[Q(u+v) - Q(u) - Q(v)] \,, \tag{45.1.5}$$

using the relations imposed by the ideal. Hence, given a product $\mathbb{e}_{\mu_1} \cdots \mathbb{e}_{\mu_k}$ of k basis vectors, one may always put them in canonical order and remove duplicates, by performing the substitutions

$$\mathbb{e}_\nu \mathbb{e}_\mu = -\mathbb{e}_\mu \mathbb{e}_\nu \text{ for } \mu < \nu \,, \quad \mathbb{e}_\mu \mathbb{e}_\mu = Q(\mathbb{e}_\mu)\mathbb{1} \,. \tag{45.1.6}$$

Hence, a basis of the real or complex Clifford algebra is given by the products

$$\mathbb{e}_{\mu_1} \cdots \mathbb{e}_{\mu_k} \,, \quad \mu_1 < \ldots < \mu_k \,. \tag{45.1.7}$$

This basis has 2^n elements, if V is of dimension n. It follows that we can write any (real or complex) Clifford algebra $\mathrm{Cl}(V, Q)$ over a vector space V of dimension n in the form

$$\mathrm{Cl}(V, Q) = \bigoplus_{k=0}^{n} \mathrm{Cl}^k(V, Q) \tag{45.1.8}$$

with

$$\mathrm{Cl}^k(V, Q) = \mathrm{span}\{\mathbb{e}_{\mu_1} \cdots \mathbb{e}_{\mu_k} \,, \quad \mu_1 < \ldots < \mu_k\} \,, \tag{45.1.9}$$

and that the (real or complex) dimensions of these subspaces are given by

$$\dim \mathrm{Cl}^k(V, Q) = \binom{n}{k} = \frac{n!}{k!(n-k)!} \,. \tag{45.1.10}$$

Of course, the formula

$$\dim \mathrm{Cl}(V, Q) = \sum_{k=0}^{n} \binom{n}{k} = 2^n \tag{45.1.11}$$

for the dimension of the Clifford algebra holds, and we have the subspaces $\mathrm{Cl}^0(V, Q) \cong \mathbb{K}$ and $\mathrm{Cl}^1(V, Q) \cong V$, where $\cong$ here denotes isomorphisms of vector spaces. In the following, we will canonically identify these subspaces with $\mathbb{K}$ and V, respectively.

45.2 Involutions

Every Clifford algebra is naturally equipped with a number of involutions, which can be used to define certain subalgebras and, as we will see later, certain groups, which will allow us to define the spin geometry we intend to construct. To arrive at these involutions, it is helpful to first define the notion of an automorphism of a Clifford algebra.

Definition 45.2.1 (Clifford algebra automorphism)**.** Let $\mathrm{Cl}(V,Q)$ be a Clifford algebra. An *automorphism* of $\mathrm{Cl}(V,Q)$ is a linear bijection $\phi : \mathrm{Cl}(V,Q) \to \mathrm{Cl}(V,Q)$ which is compatible with the algebra multiplication and preserves the subspace $\mathrm{Cl}^1(V,Q) \cong V$. The set of automorphisms of $\mathrm{Cl}(V,Q)$ will be denoted $\mathrm{Aut}(\mathrm{Cl}(V,Q))$.

It is clear that the automorphisms $\mathrm{Aut}(\mathrm{Cl}(V,Q))$ form a group, where the group multiplication is given by composition, and the group inverse is given by the inverse function, since the composition of two automorphisms is again an automorphism. It turns out that this group has a particular form, and is a Lie group which we already encountered.

Theorem 45.2.1. *The automorphisms* $\mathrm{Aut}(\mathrm{Cl}(V,Q))$ *of a Clifford algebra form a group which is isomorphic to the orthogonal group* $\mathrm{O}(V,Q)$*.*

Proof. We can explicitly construct the automorphism between $\mathrm{Aut}(\mathrm{Cl}(V,Q))$ and $\mathrm{O}(V,Q)$. First, recall that a Clifford automorphism $\phi : \mathrm{Cl}(V,Q) \to \mathrm{Cl}(V,Q)$ preserves the subspace $\mathrm{Cl}^1(V,Q) \cong V$. For elements $u, v \in V$ we can write

$$\begin{aligned}\langle \phi(u), \phi(v)\rangle \mathbb{1} &= \frac{1}{2}[\phi(u)\phi(v) + \phi(v)\phi(u)]\\ &= \frac{1}{2}[\phi(uv) + \phi(vu)]\\ &= \frac{1}{2}\phi(uv + vu)\\ &= \langle u, v\rangle \phi(\mathbb{1})\\ &= \langle u, v\rangle \mathbb{1}\,,\end{aligned} \tag{45.2.1}$$

where we used the fact that ϕ is an algebra automorphism, so that $\phi(uv) = \phi(u)\phi(v)$, $\phi(\mathbb{1}) = \mathbb{1}$ and ϕ is linear. It thus follows that ϕ preserves the inner product $\langle \bullet, \bullet\rangle$, and so it defines an element $A_\phi = \phi|_V \in \mathrm{O}(V,Q)$. Conversely, given $A \in \mathrm{O}(V,Q)$, we can uniquely construct an element $\phi_A \in \mathrm{Aut}(\mathrm{Cl}(V,Q))$ such that $\phi_A|_V = A$, and extending to a Clifford algebra automorphism. ∎

The orthogonal group has a particular element, namely the reflection with respect to the origin, which sends every element of V to its inverse. This obviously preserves the quadratic form, since $Q(-v) = Q(v)$. Hence, by the above theorem, it follows that the reflection induces a Clifford algebra automorphism. We define this automorphism as follows.

Definition 45.2.2 (Canonical automorphism)**.** Let $\mathrm{Cl}(V,Q)$ be a Clifford algebra. The *canonical automorphism* is the unique automorphism $\tilde{\bullet} : \mathrm{Cl}(V,Q) \to \mathrm{Cl}(V,Q), v \mapsto \tilde{v}$ such that $\tilde{v} = -v$ for all $v \in \mathrm{Cl}^1(V,Q)$.

Here we use the notation introduced in [Har90]. Another notation found uses $\hat{\bullet}$ and # symbols [VR16], or $\bullet'$ instead [Gar11]. The canonical automorphism is denoted by α in most other literature [LM89]. Here we prefer the former, in order to avoid potential confusion with the inner automorphism of Lie groups given in definition 15.2.3.

In addition to the canonical automorphism, also two antiautomorphisms, i.e., linear bijections which reverse the algebra product, are canonically defined. We start with the following.

Definition 45.2.3 (Clifford transpose)**.** Let $\mathrm{Cl}(V,Q)$ be a Clifford algebra. The *transpose* is the unique antiautomorphism $\check{\bullet} : \mathrm{Cl}(V,Q) \to \mathrm{Cl}(V,Q), v \mapsto \check{v}$ which maps $v_1 \cdots v_r$ to $v_r \cdots v_1$ for all $v_1, \ldots, v_r \in \mathrm{Cl}^1(V,Q)$.

Again we use the notation from [Har90]. The Clifford transpose is occasionally also denoted with $\bullet^t$. We do not use this notation, in order to avoid confusion with the transpose of a matrix, when we come to the representation of Clifford algebras in terms of matrices.

Combining the transpose and the canonical automorphism, we arrive at the following.

Definition 45.2.4 (Clifford conjugate)**.** Let $\mathrm{Cl}(V,Q)$ be a Clifford algebra. The *conjugate* is the antiautomorphism $\hat{\bullet} : \mathrm{Cl}(V,Q) \to \mathrm{Cl}(V,Q), v \mapsto \hat{v} = \check{\tilde{v}}$.

Also here we follow [Har90] in the notation.

One now easily checks the following property of the three (anti-)automorphisms we introduced above.

Theorem 45.2.2. *The canonical automorphism, transpose and conjugate operations on a Clifford algebra are involutions.*

Proof. We use the fact that any element of a Clifford algebra can be written as a sum of products of elements of $\mathrm{Cl}^1(V,Q)$. Since we are dealing with (anti-)automorphisms by definition, these are in particular linear, and so we can consider their action on a single term in this sum, i.e., a product $v_1 \cdots v_r$ with $v_1, \ldots, v_r \in \mathrm{Cl}^1(V,Q)$. For the canonical automorphism we have

$$\widetilde{v_1 \cdots v_r} = \tilde{v}_1 \cdots \tilde{v}_r = (-v_1) \cdots (-v_r) = (-1)^r v_1 \cdots v_r \,, \tag{45.2.2}$$

and so by applying $\tilde{\bullet}$ again the sign cancels, showing that $\tilde{\bullet}$ is an involution. For the Clifford transpose we obviously have

$$v_1 \cdots v_r \xrightarrow{\check{\bullet}} v_r \cdots v_1 \xrightarrow{\check{\bullet}} v_1 \cdots v_r \,, \tag{45.2.3}$$

showing that this is also an involution. Finally, it is obvious that both operations commute, and so their composition $\hat{\bullet}$ is again an involution. ∎

For an involution on a vector space, the eigenvalues are ± 1, and so we can identify the eigenspaces corresponding to these eigenvalues. Of particular interest are the eigenspaces of the canonical automorphism. We define and denote them as follows.

Definition 45.2.5 (Even and odd subspaces of a Clifford algebra)**.** Let $\mathrm{Cl}(V,Q)$ be a Clifford algebra. Its *even and odd subspaces* are the eigenspaces of the canonical automorphism,

$$\mathrm{Cl}^+(V,Q) = \{v \in \mathrm{Cl}(V,Q), \tilde{v} = v\} \,, \quad \mathrm{Cl}^-(V,Q) = \{v \in \mathrm{Cl}(V,Q), \tilde{v} = -v\} \,. \tag{45.2.4}$$

Looking at the action of the canonical automorphism on the subspaces $\mathrm{Cl}^k(V,Q)$, we can also express the even and odd subspaces as follows.

Theorem 45.2.3. *The even and odd subspaces of a Clifford algebra* $\mathrm{Cl}(V,Q)$ *are given by*

$$\mathrm{Cl}^+(V,Q) = \bigoplus_{k\in 2\mathbb{Z}} \mathrm{Cl}^k(V,Q)\,, \quad \mathrm{Cl}^-(V,Q) = \bigoplus_{k\in 2\mathbb{Z}+1} \mathrm{Cl}^k(V,Q)\,. \tag{45.2.5}$$

Proof. ►...◄ ■

It is instructive to study the action of the (anti-)automorphisms on the basis elements (which is then extended to the whole Clifford algebra by their linearity). For a basis element $\mathbb{e}_{\mu_1}\cdots\mathbb{e}_{\mu_k}$ with $\mu_1 < \ldots < \mu_k$ we have, by definition of the canonical automorphism

$$\widetilde{\mathbb{e}_{\mu_1}\cdots\mathbb{e}_{\mu_k}} = \tilde{\mathbb{e}}_{\mu_1}\cdots\tilde{\mathbb{e}}_{\mu_k} = (-1)^k\mathbb{e}_{\mu_1}\cdots\mathbb{e}_{\mu_k}\,, \tag{45.2.6}$$

which again shows that an element of $\mathrm{Cl}^k(V,Q)$ is even (odd) if and only if k is even (odd). We find similar rules for the two antiautomorphisms we defined. For the transpose we have

$$\mathbb{e}_{\mu_1}\check{\cdots}\mathbb{e}_{\mu_k} = \mathbb{e}_{\mu_k}\cdots\mathbb{e}_{\mu_1} = (-1)^{k(k-1)/2}\mathbb{e}_{\mu_1}\cdots\mathbb{e}_{\mu_k}\,, \tag{45.2.7}$$

and combining both we find the conjugate

$$\mathbb{e}_{\mu_1}\hat{\cdots}\mathbb{e}_{\mu_k} = (-1)^{k(k+1)/2}\mathbb{e}_{\mu_1}\cdots\mathbb{e}_{\mu_k}\,. \tag{45.2.8}$$

The most convenient way to describe these relation is by realizing that the sign depends only on $k \mod 4$. For $v \in \mathrm{Cl}^k(V,Q)$ one then has

$$v \mapsto \left\{\begin{array}{c|ccc} & \tilde{\bullet} & \check{\bullet} & \hat{\bullet} \\ \hline k \mod 4 = 0 & +1 & +1 & +1 \\ k \mod 4 = 1 & -1 & +1 & -1 \\ k \mod 4 = 2 & +1 & -1 & -1 \\ k \mod 4 = 3 & -1 & -1 & +1 \end{array}\right\} v\,. \tag{45.2.9}$$

This leads us to the following property of the canonical automorphism.

Theorem 45.2.4. *The canonical automorphism induces a* $\mathbb{Z}_2$*-grading* $\mathrm{Cl}(V,Q) = \mathrm{Cl}^+(V,Q) \oplus \mathrm{Cl}^-(V,Q)$ *of the Clifford algebra.*

Proof. Let $u, v \in \mathrm{Cl}(V,Q)$ be homogeneous elements, i.e., elements of one of the subspaces $\mathrm{Cl}^\pm(V,Q)$, so that there exists $\sigma_u, \sigma_v \in \{1,-1\}$ such that $\tilde{u} = \sigma_u u$ and $\tilde{v} = \sigma_v v$. Then we have

$$\widetilde{uv} = \tilde{u}\tilde{v} = \sigma_u\sigma_v uv\,, \tag{45.2.10}$$

since $\tilde{\bullet}$ is an automorphism, and hence distributes over products. Hence, uv is homogeneous of degree $\sigma_u\sigma_v$. ■

Note that this holds only for the canonical automorphism, but not for the two antiautomorphisms we have defined. It follows from the grading that the even elements $\mathrm{Cl}^+(V,Q)$ constitute a subalgebra. This will be important to construct the spin groups we will use in the following sections.

45.3 Clifford, pin and spin groups

An element of a Clifford algebra $\mathrm{Cl}(V,Q)$, in general, does not have an inverse, i.e., an element such that their product yields the identity element $\mathbb{1}$. Those elements which possess an inverse form a group, which we define as follows, and whose subgroups we study in this section, following mostly the treatment in [Har90, ch. 10], with elements from [LM89, ch. I, §2]. The aim of this section is to give an overview of the groups which are commonly defined in the literature and their relations, with a particular focus on the spin groups, which we will then use in the following sections.

Definition 45.3.1 (Invertible Clifford elements)**.** Let $\mathrm{Cl}(V, Q)$ be a Clifford algebra. Its *group of invertible elements* is the group constituted by the elements

$$\mathrm{Cl}^*(V, Q) = \{A \in \mathrm{Cl}(V, Q), \exists A^{-1} : AA^{-1} = A^{-1}A = \mathbb{1}\} . \tag{45.3.1}$$

The definition is thus analogue to the definition of the general linear group in example 15.1.3, which is given by all invertible elements of a matrix algebra. For the general linear group we already know that it is a Lie group. One may thus expect the same to hold for the group we constructed here. This is indeed the case. Here we restrict ourselves to the canonical cases.

Theorem 45.3.1. *$\mathrm{Cl}^*_{k,l}(\mathbb{R})$ is a real Lie group of real dimension 2^{k+l}, while $\mathrm{Cl}^*_n(\mathbb{C})$ is a complex Lie group of complex dimension 2^n.*

Proof. ▶ . . . ◀ ■

In the following, we will construct various subgroups of $\mathrm{Cl}^*(V, Q)$. In order to construct these groups, it is helpful to note that the group of invertible elements acts (from the left) on the Clifford algebra in two canonical ways. We start with the following, which is straightforward.

Definition 45.3.2 (Adjoint representation of a Clifford algebra)**.** Let $\mathrm{Cl}(V, Q)$ be a Clifford algebra and $\mathrm{Cl}^*(V, Q)$ the group of invertible elements. The *adjoint representation* is defined as the group action

$$\begin{array}{rcccc} \mathrm{Ad} & : & \mathrm{Cl}^*(V, Q) \times \mathrm{Cl}(V, Q) & \to & \mathrm{Cl}(V, Q) \\ & & (u, v) & \mapsto & \mathrm{Ad}_u(v) = uvu^{-1} \end{array} . \tag{45.3.2}$$

The adjoint representation is a common operation in the theory of algebras. However, in the case of Clifford algebras, it turns out that there is another, related representation, which makes use of the canonical automorphism we provided in definition 45.2.2. This allows us to define the following.

Definition 45.3.3 (Twisted adjoint representation of a Clifford algebra)**.** Let $\mathrm{Cl}(V, Q)$ be a Clifford algebra and $\mathrm{Cl}^*(V, Q)$ the group of invertible elements. The *twisted adjoint representation* is defined as the group action

$$\begin{array}{rcccc} \widetilde{\mathrm{Ad}} & : & \mathrm{Cl}^*(V, Q) \times \mathrm{Cl}(V, Q) & \to & \mathrm{Cl}(V, Q) \\ & & (u, v) & \mapsto & \widetilde{\mathrm{Ad}}_u(v) = \tilde{u}vu^{-1} \end{array} . \tag{45.3.3}$$

Note that here the element u multiplied from the left is replaced by the element $\tilde{u}$ obtained using the canonical automorphism. We now aim to derive actions on the vector space V from these two representations. Note that for $v \in \mathrm{Cl}^1(V, Q) \cong V$, it is not necessarily satisfied that also $\mathrm{Ad}_u(v) \in \mathrm{Cl}^1(V,Q)$ or $\widetilde{\mathrm{Ad}}_u(v) \in \mathrm{Cl}^1(V,Q)$ for arbitrary u. However, there are elements u which actually satisfy these conditions, and they form groups. This brings us to the following definition.

Definition 45.3.4 (Reduced Clifford group)**.** Let $\mathrm{Cl}(V,Q)$ be a Clifford algebra. Its *reduced Clifford group* is the group $\mathrm{P}(V,Q)$ of invertible elements $u \in \mathrm{Cl}^*(V,Q)$ which satisfy $uvu^{-1} \in \mathrm{Cl}^1(V,Q)$ for all $v \in \mathrm{Cl}^1(V,Q)$.

Here we introduced the name *reduced Clifford group* for reasons which will become clear below. More common, more widely used and also more important for our following constructions is the following.

Definition 45.3.5 (Clifford group)**.** Let $\mathrm{Cl}(V,Q)$ be a Clifford algebra. Its *Clifford group* (or *Clifford-Lipschitz group*) is the group $\tilde{\mathrm{P}}(V,Q)$ of invertible elements $u \in \mathrm{Cl}^*(V,Q)$ which satisfy $\tilde{u}vu^{-1} \in \mathrm{Cl}^1(V,Q)$ for all $v \in \mathrm{Cl}^1(V,Q)$.

A more commonly found notation is $\Gamma(V,Q)$. In order to study the properties of these groups, and answer the question why we are interested in particular in the *twisted* adjoint representation $\widetilde{\mathrm{Ad}}$ and use it to define the Clifford group $\tilde{\mathrm{P}}(V,Q)$, it is helpful to take a look at some special elements of this group, namely those that lie in $\mathrm{Cl}^1(V,Q) \cong V$. Obviously, an element $u \in V$ is invertible in $\mathrm{Cl}(V,Q)$ if and only if $Q(v) \neq 0$, and we have $u^{-1} = u/Q(u)$, since

$$uu^{-1} = u^{-1}u = \frac{uu}{Q(u)} = \frac{Q(u)\mathbb{1}}{Q(u)} = \mathbb{1}\,. \tag{45.3.4}$$

Further, note that $\tilde{u} = -u$, since $u \in \mathrm{Cl}^-(V,Q)$ following theorem 45.2.3. Hence, for its action on $v \in \mathrm{Cl}^1(V,Q) \cong V$ we can write

$$\widetilde{\mathrm{Ad}}_u(v) = \tilde{u}vu^{-1} = -\frac{uvu}{Q(u)} = u\frac{uv - 2\langle u,v\rangle\mathbb{1}}{\langle u,u\rangle} = v - 2u\frac{\langle u,v\rangle}{\langle u,u\rangle}\,, \tag{45.3.5}$$

and so we get the *reflection along* u: we have $\widetilde{\mathrm{Ad}}_u(v) = v$ if u and v are orthogonal, $\langle u,v\rangle = 0$, and $\widetilde{\mathrm{Ad}}_u(v) = -v$ if they are collinear, $v = \lambda u$ for some $\lambda \in \mathbb{R}$. Note that this in particular has the following consequence.

Theorem 45.3.2. *The twisted adjoint representation of the Clifford-Lipschitz group $\tilde{\mathrm{P}}(V,Q)$ on the vector space V preserves the inner product,*

$$\langle\widetilde{\mathrm{Ad}}_u(v), \widetilde{\mathrm{Ad}}_u(w)\rangle = \langle v,w\rangle \tag{45.3.6}$$

for all $u \in \tilde{\mathrm{P}}(V,Q)$ and $v,w \in V$.

Proof. By direct calculation, we have

$$\begin{aligned}\langle\widetilde{\mathrm{Ad}}_u(v), \widetilde{\mathrm{Ad}}_u(w)\rangle &= \left\langle v - 2u\frac{\langle u,v\rangle}{\langle u,u\rangle}, w - 2u\frac{\langle u,w\rangle}{\langle u,u\rangle}, \right\rangle \\ &= \langle v,w\rangle - 2\frac{\langle u,v\rangle\langle u,w\rangle + \langle u,w\rangle\langle v,u\rangle}{\langle u,u\rangle} + 4\frac{\langle u,u\rangle\langle u,v\rangle\langle u,w\rangle}{\langle u,u\rangle^2} \\ &= \langle v,w\rangle\end{aligned} \tag{45.3.7}$$

for $u \in \tilde{\mathrm{P}}(V,Q)$ and $v,w \in V$. ∎

Theorem 45.3.3. *The Clifford group is given by*

$$\tilde{\mathrm{P}}(V,Q) = \{u_1 \cdots u_r, \blacktriangleright \dots \blacktriangleleft\} \tag{45.3.8}$$

Proof. ▶. . . ◀ ∎

Note that the "length" of u does not matter for the twisted adjoint representation, so we have $\widetilde{\mathrm{Ad}}_{\lambda u} = \widetilde{\mathrm{Ad}}_u$ for $\lambda \in \mathbb{R}^*$. Hence, we may restrict ourselves to such u which have unit length. This yields the following group.

Definition 45.3.6 (Pin group)**.** Let $\mathrm{Cl}(V,Q)$ be a Clifford algebra. Its *pin group* $\mathrm{Pin}(V,Q)$ is the group generated by the elements $v \in \mathrm{Cl}^1(V,Q) \cong V$ with $|Q(v)| = 1$.

Definition 45.3.7 (Spin group)**.** Let $\mathrm{Cl}(V,Q)$ be a Clifford algebra. Its *spin group* $\mathrm{Spin}(V,Q)$ is the group $\mathrm{Spin}(V,Q) = \mathrm{Pin}(V,Q) \cap \mathrm{Cl}^+(V,Q)$.

Definition 45.3.8 (Orthogonal pin group)**.** Let $\mathrm{Cl}(V,Q)$ be a Clifford algebra. Its *orthogonal pin group* $\check{\mathrm{Pin}}(V,Q)$ is the group containing those elements $v \in \mathrm{Pin}(V,Q)$ which satisfy $v\check{v} = \mathbb{1}$.

Definition 45.3.9 (Unitary pin group)**.** Let $\mathrm{Cl}(V,Q)$ be a Clifford algebra. Its *unitary pin group* $\hat{\mathrm{Pin}}(V,Q)$ is the group containing those elements $v \in \mathrm{Pin}(V,Q)$ which satisfy $v\hat{v} = \mathbb{1}$.

Definition 45.3.10 (Reduced spin group)**.** Let $\mathrm{Cl}(V,Q)$ be a Clifford algebra. Its *reduced spin group* $\mathrm{Spin}^0(V,Q)$ is the group $\mathrm{Spin}^0(V,Q) = \check{\mathrm{Pin}}(V,Q) \cap \hat{\mathrm{Pin}}(V,Q)$.

Theorem 45.3.4. *The reduced spin group is alternatively given by*

$$\mathrm{Spin}^0(V,Q) = \mathrm{Spin}(V,Q) \cap \check{\mathrm{Pin}}(V,Q) = \mathrm{Spin}(V,Q) \cap \hat{\mathrm{Pin}}(V,Q)\,. \tag{45.3.9}$$

Proof. ▶. . . ◀ ∎

Theorem 45.3.5. *If V is a real vector space of dimension n and Q a quadratic form with definite signature, then the corresponding pin and spin groups satisfy*

$$\check{\mathrm{Pin}}_{n,0}(\mathbb{R}) = \mathrm{Pin}_{n,0}(\mathbb{R})\,, \quad \hat{\mathrm{Pin}}_{n,0}(\mathbb{R}) = \mathrm{Spin}^0_{n,0}(\mathbb{R}) = \mathrm{Spin}_{n,0}(\mathbb{R})\,, \tag{45.3.10}$$

as well as

$$\hat{\mathrm{Pin}}_{0,n}(\mathbb{R}) = \mathrm{Pin}_{0,n}(\mathbb{R})\,, \quad \check{\mathrm{Pin}}_{0,n}(\mathbb{R}) = \mathrm{Spin}^0_{0,n}(\mathbb{R}) = \mathrm{Spin}_{0,n}(\mathbb{R})\,. \tag{45.3.11}$$

Proof. ▶. . . ◀ ∎

45.4 Spin structures

45.5 Spin bundles

45.6 Spinor bundles

Chapter 46

Non-commutative geometry

Chapter 47

Supermanifolds

Part III

Physical applications

Chapter 48

Differential equations

48.1 First-order ordinary differential equations of multiple variables

48.2 Second-order ordinary differential equations of multiple variables

48.3 Higher-order ordinary differential equations of multiple variables

Chapter 49

Lagrange theory on finite jet bundles

49.1 Lagrangians and action functionals

We now come to a physical application of the formalism introduced in the previous section. The physical system we consider here is called a *Lagrangian system*. It is modeled by a fiber bundle $\pi : E \to M$, where physical solutions of the system are a subset of the space of sections $\Gamma(E)$. This set of solutions is obtained from an *action principle*. In order to clarify these terms, we start with a few definitions.

Definition 49.1.1 (Lagrangian)**.** Let $\pi : E \to M$ be a fiber bundle with $\dim M = n$. A *Lagrangian* of order $r \in \mathbb{N}$ on E is a horizontal n-form L on the r-th jet bundle $J^r(E)$.

Recall that the elements of $J^r(E)$ are equivalence classes of sections $\sigma \in \Gamma(E)$ obtained by choosing a point $p \in M$ and evaluating the section to $\sigma(p)$ and its partial derivatives of order up to r. A Lagrangian thus depends on p, $\sigma(p)$ and its derivatives at p, i.e., on the *local* behavior of the section σ. We know that we can obtain a *global* property if we integrate a differential form. This will be done in the next definition.

Definition 49.1.2 (Action functional)**.** Let $\pi : E \to M$ be a fiber bundle and $L \in \Omega^{n,0}(J^r(E))$ a Lagrangian of order r on E. The *action functional* of L over an open domain $U \subset M$ is the function

$$\begin{array}{rcccc} S & : & \Gamma_{|U}(E) & \to & \mathbb{R} \\ & & \sigma & \mapsto & \displaystyle\int_U (j^r\sigma)^*(L) \end{array} . \tag{49.1.1}$$

Here the jet prolongation $j^r\sigma : U \to J^r(E)$ is used to pull back the n-form $L \in \Omega^n(J^r(E))$ to an n-form $(j^r\sigma)^*(L) \in \Omega^n(U)$. Since we have $\dim M = \dim U = n$, this can be integrated over the domain U. To illustrate this further, we give an example from classical mechanics.

Example 49.1.1 (First order Lagrangian of a point mass on a metric manifold with potential). Let $M = \mathbb{R}$ and Q a manifold of dimension n. Let $E = \mathbb{R} \times Q$ be the trivial fiber bundle with projection $\pi : \mathbb{R} \times Q \to \mathbb{R}$ onto the first factor. Sections of this bundle are uniquely expressed by maps $\gamma \in C^\infty(\mathbb{R}, Q)$, i.e., by curves on Q. We use the one-dimensional Euclidean coordinate t on $\mathbb{R}$ and arbitrary coordinates (q^a) on Q, so that we have coordinates (t, q^a) on $\mathbb{R} \times Q$. From these coordinates we derive the coordinates $(t, q^a_{(0)}, q^a_{(1)})$ on $J^1(E) \cong \mathbb{R} \times TQ$.

To construct a particular Lagrangian, let further $g \in \Gamma(T^0_2 Q)$ be a non-degenerate, positive definite, symmetric tensor field of type $(0, 2)$ (the *metric*) and $V \in C^\infty(Q, \mathbb{R})$ (the *potential*). To illustrate this, we do this in five steps, each of which we explain in our geometric language:

- We take an element of $J^1(E) \cong \mathbb{R} \times TQ$ and project it onto the second factor. This yields a tangent vector $q_{(1)} = q^a_{(1)} \partial_a \in T_{q_{(0)}} Q$, where $q_{(0)} \in Q$ is the result of using the bundle map of TQ on $q_{(1)}$. For convenience, we write $q = q_{(0)}, \dot{q} = q_{(1)}$ and also the coordinates $q^a = q^a_{(0)}, \dot{q}^a = q^a_{(1)}$.
- We take the metric g, which is a section of the tensor bundle $T^0_2 Q$, and maps q to $g(q) \in T^0_{2\,q} Q$. This is a covariant tensor, so we can contract it with two copies of the vector $\dot{q} \in T_q Q$ and obtain a real number. In coordinates we thus get $g_{ab}(q) \dot{q}^a \dot{q}^b \in \mathbb{R}$. Doing this for all elements of $J^1(E)$ gives us a real function on $J^1(E)$.
- We take the potential V, which is a real function on Q, and evaluate it at q, so we obtain another real number $V(q)$. Doing this for all elements of $J^1(E)$ gives us another real function on $J^1(E)$.
- We take the canonical one-form $\omega = \mathrm{d}t \in \Omega^1(\mathbb{R})$ on $\mathbb{R}$ and pull it back via the projection $\pi_1 : J^1(E) \to \mathbb{R}$. This yields us a horizontal one-form $\pi_1^*(\omega) = \mathrm{d}t \in \Omega^1(J^1(E))$.
- We combine the two real functions and the one-form constructed above to the first order Lagrangian

$$L(t, q, \dot{q}) = \left(\frac{1}{2} g_{ab}(q) \dot{q}^a \dot{q}^b - V(q) \right) \mathrm{d}t \in \Omega^1(J^1(E)) \,. \tag{49.1.2}$$

Finally, we construct the action functional. For this purpose we consider a section, described by a curve $\gamma : \mathbb{R} \to Q$, which in our chosen coordinates is described by functions $\gamma^a(t)$. The pullback along this section then simply amounts to replacing the coordinates q^a and $\dot{q}^a$ in the Lagrangian by $\gamma^a(t)$ and $\partial \gamma^a(t)/\partial t$. This yields a one-form on $M = \mathbb{R}$, which can be integrated to the action

$$S[q] = \int_{\mathbb{R}} \left(\frac{1}{2} g_{ab}(\gamma(t)) \frac{\partial \gamma^a(t)}{\partial t} \frac{\partial \gamma^b(t)}{\partial t} - V(\gamma(t)) \right) \mathrm{d}t \tag{49.1.3}$$

One now easily recognizes the action of a point mass, with all function arguments explicitly written out in order to clarify that this is now truly an object on M. Of course one may ask why we use this particular Lagrangian - for now the answer is simply: "Because it yields us the correct physics in the end." But we still need to arrive at the reason for this.

We discuss another example from field theory.

Example 49.1.2 (First order Lagrangian of a massive scalar field on a metric manifold). Let M be a manifold of dimension n and $E = M \times \mathbb{R}$ the trivial line bundle with projection $\pi : M \times \mathbb{R} \to M$ onto the first factor. Sections of this bundle are uniquely expressed by maps $\varphi \in C^\infty(M, \mathbb{R})$, i.e., by real functions on M. We use arbitrary coordinates (x^a) on M

and the one-dimensional Euclidean coordinate ϕ on $\mathbb{R}$, so that we have coordinates (x^a, ϕ) on $M \times \mathbb{R}$. From these coordinates we derive the coordinates

$$(x^a, \phi, \phi_{,a}) = (x^a, \phi_{(0,\ldots,0)}, \phi_{(1,0,\ldots,0)}, \ldots, \phi_{(0,\ldots,0,1)}) \tag{49.1.4}$$

on $J^1(E) \cong T^*M \times \mathbb{R}$.

To construct a particular Lagrangian, let further $g \in \Gamma(T_2^0 M)$ be a non-degenerate, symmetric tensor field of type $(0, 2)$ (the *metric*) and $V \in C^\infty(\mathbb{R}, \mathbb{R})$ (the *potential*). To illustrate this, we do this in five steps, each of which we explain in our geometric language:

- From an element $(x^a, \phi, \phi_{,a})$ we obtain elements $\phi \in \mathbb{R}$, $\phi_{,a}\mathrm{d}x^a \in T_x^*M$ and $x \in M$ by applying suitable projections as in the previous example.
- Since the metric is non-degenerate, it possesses an inverse $g^{-1} \in \Gamma(T_0^2 M)$, which is also non-degenerate and symmetric. If we evaluate it at $x \in M$, we get an element $g^{-1}(x) \in T_{0\,x}^2 M$. Contracting this element with two copies of $\phi_{,a}\mathrm{d}x^a$ yields a real number $g^{ab}(x)\phi_{,a}\phi_{,b}$.
- The potential $V \in C^\infty(\mathbb{R}, \mathbb{R})$ can be applied to $\phi \in \mathbb{R}$, which yields another real number $V(\phi) \in \mathbb{R}$.
- The metric induces a volume form $\sqrt{|\det g(x)|}\mathrm{d}^n x$ on M. The pullback of this volume form along $\pi_1 : J^1(E) \to M$ is a horizontal n-form on $J^1(E)$.
- From the objects constructed above we compose the Lagrangian

$$\left(\frac{1}{2}g^{ab}(x)\phi_{,a}\phi_{,b} - V(\phi)\right)\sqrt{|\det g(x)|}\mathrm{d}^n x\,. \tag{49.1.5}$$

To obtain the action, one finally considers a section, which is described in coordinates by a function $\varphi(x)$, and replaces the coordinates φ and $\varphi_{,a}$ by $\varphi(x)$ and $\partial\varphi(x)/\partial x^a$. The resulting one-form on M then yields the action

$$S[\varphi] = \int_M \left(\frac{1}{2}g^{ab}(x)\frac{\partial\varphi(x)}{\partial x^a}\frac{\partial\varphi(x)}{\partial x^b} - V(\varphi(x))\right)\sqrt{|\det g(x)|}\mathrm{d}^n x\,. \tag{49.1.6}$$

Also here we have explicitly written out every dependence on the point x to illustrate that we are indeed integrating over a n-form on M. Note that in this example we have treated only ϕ as a dynamical field and assumed a fixed background metric g. Both of these objects are sections of vector bundles, and normally one would consider the sum of these vector bundles as the starting point of the construction to make both fields dynamical.

We see that we can formulate these two classical examples in terms of differential geometric objects (sections of bundles) without using coordinates. The coordinates are used here only to illustrate the process and to provide explicit formulas. However, the Lagrangian formulation presented here is independent of the choice of coordinates.

49.2 Action principle and variation

We finally come to the question how to derive equations of motion, and thus the space of solutions of the Lagrangian system introduced in the last section. The *principle of least action* states that solutions of a Lagrangian system are those local sections $\sigma \in \Gamma_|(E)$ for which the action assumes a local minimum in the space of sections. However, here we will restrict ourselves to considering *extremals* of the action. This will now be clarified.

Definition 49.2.1 (Extremal of the action)**.** Let $\pi : E \to M$ be a fiber bundle with action functional S on an open domain $U \subset M$. A local section $\sigma \in \Gamma_{|U}(E)$ is called an *extremal of the action* if for all smooth families $\tilde{\sigma}_\bullet : \mathbb{R} \to \Gamma_{|U}(E)$ of local sections with $\tilde{\sigma}_0 = \sigma$ the function

$$\begin{array}{rcccc} S[\tilde{\sigma}_\bullet] & : & \mathbb{R} & \to & \mathbb{R} \\ & & \epsilon & \mapsto & S[\tilde{\sigma}_\epsilon] \end{array} \tag{49.2.1}$$

has vanishing derivative at $\epsilon = 0$.

Here we call the family $\tilde{\sigma}_\bullet : \mathbb{R} \to \Gamma_{|U}(E)$ of sections smooth if and only if the map $\tilde{\sigma}_\bullet(\bullet) : \mathbb{R} \times U \to E$ is smooth. The definition 49.2.1 means that the action is *stationary*, which is expressed by the equation

$$\left.\frac{\mathrm{d}S[\tilde{\sigma}_\epsilon]}{\mathrm{d}\epsilon}\right|_{\epsilon=0} = 0 \tag{49.2.2}$$

for all smooth families of sections with $\tilde{\sigma}_0 = \sigma$. This allows us to consider the function composition

$$\mathbb{R} \xrightarrow{\tilde{\sigma}_\bullet} \Gamma_{|U}(E) \xrightarrow{S} \mathbb{R} \tag{49.2.3}$$

and to apply some kind of "chain rule". This will be done in several steps in the following sections.

49.3 Variation of sections and their jet prolongations

Recall that in the previous sections we considered families $\epsilon \mapsto \tilde{\sigma}_\epsilon \in \Gamma_{|U}(E)$ of local sections of a fiber bundle $\pi : E \to M$ with a fixed domain $U \subset M$, and we denoted $\sigma = \tilde{\sigma}_0$. For every $p \in U$, this yields a curve

$$\begin{array}{rcccc} \gamma_p & : & \mathbb{R} & \to & E \\ & & \epsilon & \mapsto & \tilde{\sigma}_\epsilon(p) \end{array}, \tag{49.3.1}$$

where we simply evaluate the section $\tilde{\sigma}_\epsilon$ at the point p. Every such curve clearly has

$$\gamma_p(0) = \tilde{\sigma}_0(p) = \sigma(p)\,. \tag{49.3.2}$$

Further, one has $\pi(\gamma_p(\epsilon)) = p$ for all ϵ, so that the tangent vectors $\dot{\gamma}_p(0) \in T_{\sigma(p)}E$ satisfy

$$\pi_*(\dot{\gamma}_p(0)) = 0\,. \tag{49.3.3}$$

Hence, they are in fact vertical tangent vectors, $\dot{\gamma}_p(0) \in V_{\sigma(p)}E$. Denoting the vertical tangent bundle of E by $\nu : VE \to E$, we see that there exists a map

$$\begin{array}{rcccc} \xi & : & M & \to & VE \\ & & p & \mapsto & \dot{\gamma}_p(0) \end{array}, \tag{49.3.4}$$

which satisfies $\nu \circ \xi = \sigma$ (and which is thus equivalent to a section of the pullback bundle $\sigma^*\nu : \sigma^*VE \to M$, following theorem 2.9.1). Also one now easily checks that for every $p \in M$ the action of $\xi(p)$ on a function $f \in C^\infty(E, \mathbb{R})$ is given by

$$\xi(p)(f) = \dot{\gamma}_p(0)(f) = (f \circ \gamma_p)'(0) = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon} f(\tilde{\sigma}_\epsilon(p))\right|_{\epsilon=0}, \tag{49.3.5}$$

which directly follows from its definition.

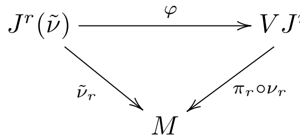
For the jet prolongation $j^r\sigma : U \to J^r(E)$, which appears in the action functional, we could now proceed on full analogy. By fixing a point $p \in U$ we obtain a curve $\epsilon \mapsto j^r_p\tilde{\sigma}_e \in J^r(E)$, whose tangent vector at $\epsilon = 0$, which we denote $\xi_r(p)$, is vertical with respect to π_r, and hence satisfies $(\pi_r)_*(\xi_r(p)) = 0$. The corresponding map $\xi_r : M \to VJ^r(E)$ then satisfies

$$\xi_r(p)(f) = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon} f(j^r_p \tilde{\sigma}_\epsilon)\right|_{\epsilon=0}, \tag{49.3.6}$$

for every $p \in U$ and $f \in C^\infty(J^r(E), \mathbb{R})$. Denoting the vertical tangent bundle of $J^r(\pi)$ by $\nu_r : VJ^r(\pi) \to J^r(\pi)$, we obviously have $\nu_r \circ \xi_r = j^r\sigma$.

Since the jet prolongation $j^r\sigma$ is uniquely defined by σ, and the same holds for $j^r\tilde{\sigma}_\epsilon$ and $\tilde{\sigma}_\epsilon$, one may already expect that also ξ_r is completely determined by ξ. To see that this is indeed the case, consider first the fiber bundle $\tilde{\nu} = \pi \circ \nu : VE \to M$. Note that this bundle is *not* a vector bundle, since the vector space structure on VE is defined only on the fibers over E - the fibers over M are not vector spaces, but unions of vector spaces, and one cannot add elements which lie in different vector spaces within this union. Clearly ξ is a local section of this bundle with domain U. As with any fiber bundle, we may now construct the jet bundle $\tilde{\nu}_r : J^r(\tilde{\nu}) \to M$, and prolong ξ to a local section $j^r\xi : U \to J^r(\tilde{\nu})$ of this bundle. Now the following fact turns out to be useful.

Theorem 49.3.1. *There exists a canonical fiber bundle isomorphism φ covering the identity such that the diagram*



(49.3.7)

commutes.

Instead of a full proof, we will only sketch the construction of φ. Let $\xi : M \to VE$ be a section of $\tilde{\nu} : VE \to M$ with $\sigma = \nu \circ \xi : M \to E$. Given $p \in M$, we thus have a jet $j^r_p\xi \in J^r(\tilde{\nu})$. Our aim is to construct an element $\varphi(j^r_p\xi) \in VJ^r(\pi)$. Since the latter is a (vertical) vector on $J^r(\pi)$, and thus a derivation, it can be applied to arbitrary (smooth) functions $f \in C^\infty(J^r(\pi), \mathbb{R})$. We now consider a particular class of such functions, which we construct as follows. Let $0 \leq k \leq r$, $\gamma \in C^\infty(\mathbb{R}, M)$ with $\gamma(0) = p$ and $u \in C^\infty(E, \mathbb{R})$, and define

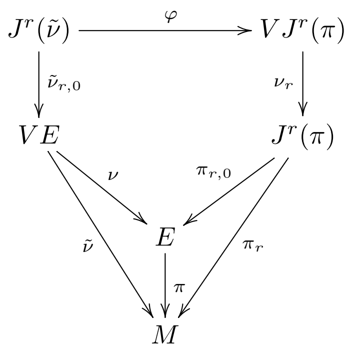
$$\begin{array}{cccc} F^k_{\gamma,u} & : & \Gamma_p(\pi) & \to & \mathbb{R} \\ & & \tau & \mapsto & (u \circ \tau \circ \gamma)^{(k)}(0) \end{array}. \tag{49.3.8}$$

By definition 21.6.1 of the r-jet $j^r_p\tau$, this function yields the same value for all local sections that have the same r-jet. Hence, it defines a function $\tilde{F}^k_{\gamma,u} \in C^\infty(J^r(E), \mathbb{R})$. We then define the action of $\varphi(j^r_p\xi)$ as

$$\varphi(j^r_p\xi)(\tilde{F}^k_{\gamma,u}) = \left.\frac{\mathrm{d}}{\mathrm{d}t} \langle \xi(\gamma(t)), \mathrm{d}u(\sigma(\gamma(t))) \rangle\right|_{t=0}. \tag{49.3.9}$$

Note that $\xi(\gamma(t)) \in V_{\sigma(\gamma(t))}E \subset T_{\sigma(\gamma(t))}E$ and $\mathrm{d}u(\sigma(\gamma(t))) \in T^*_{\sigma(\gamma(t))}E$, so that this is indeed well-defined.

We illustrate the construction using the following commutative diagram:



(49.3.10)

Using appropriate charts, we have coordinates (x^a) on M and adapted coordinates (x^a, y^μ) on E. These can then be used to construct coordinates (x^a, y^μ, v^μ) on the vertical tangent bundle VE, by writing a vertical tangent vector in the coordinate basis as

$$v^\mu \frac{\partial}{\partial y^\mu} \,. \tag{49.3.11}$$

On $J^r(\pi)$, one constructs coordinates (x^a, y^μ_Λ) with $0 \leq |\Lambda| \leq r$. Writing a vertical tangent vector to $J^r(\pi)$ as

$$v^\mu_\Lambda \frac{\partial}{\partial y^\mu_\Lambda} \,, \tag{49.3.12}$$

one obtains coordinates $(x^a, y^\mu_\Lambda, v^\mu_\Lambda)$ on $VJ^r(\pi)$. Finally, from the coordinates (x^a, y^μ, v^μ) on the vertical tangent bundle VE one also obtains coordinates $(x^a, y^\mu_\Lambda, v^\mu_\Lambda)$ on $J^r(\tilde{\nu})$. In these coordinates the isomorphism $\varphi : J^r(\tilde{\nu}) \to VJ^r(\pi)$ maps each tuple of coordinates to itself.

Using this bundle isomorphism, one can finally write $\xi_r : M \to VJ^r(\pi)$ as $\xi_r = \varphi \circ j^r\xi$ using the jet prolongation $j^r\xi : M \to J^r(\tilde{\nu})$.

49.4 Variation of forms on jet bundles

In order to calculate the variation of the action we need to know how $(j^r\tilde{\sigma}_\epsilon)^*(L)$ varies with ϵ. This is given by the following theorem.

Theorem 49.4.1. *Let $\tilde{\sigma}_\epsilon : M \to E$ be a smooth family of sections of the fiber bundle $\pi : E \to M$ and $\omega \in \Omega^{k,0}(J^r)(E))$ a horizontal k-form on the r-jet bundle $J^r(E)$. Then the pullback of ω along $j^r\tilde{\sigma}_\epsilon$ satisfies*

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}(j^r\tilde{\sigma}_\epsilon)^*(\omega)\bigg|_{\epsilon=0} = (j^r\sigma)^*\left(\iota_{\xi_r}(\mathrm{d}\omega)\right) \,, \tag{49.4.1}$$

where ξ_r is constructed from the r-jet prolongation of ξ and $\sigma = \tilde{\sigma}_0$.

We will not prove this here, but a few clarifications are in order. First, recall that $\xi_r : M \to VJ^r(E) \subset TJ^r(E)$ takes an element $p \in M$ and assigns to it a (vertical) tangent vector $\xi_r(p)$ at $j^r_p\sigma \in J^r(E)$. Equivalently, it can be seen as a particular section of the pullback bundle of $TJ^r(E) \to J^r(E)$ to M along $j^r\sigma$, following theorem 2.9.1. A section of $\Lambda^{k+1}T^*J^r(E) \to J^r(E)$ can be pulled back along $j^r\sigma$ to a section of the corresponding pullback bundle over M. Hence, one obtains sections of bundles $TJ^r(E) \to M$ and $\Lambda^{k+1}T^*J^r(E) \to M$. Since the latter can be seen as the bundle of alternating multilinear functions on $TJ^r(E)$, there exists a well-defined notion of an interior product of these two sections, which yields a section of $\Lambda^kT^*J^r(E) \to M$. This can finally be pulled back to a section of $\Lambda^kT^*M \to M$, which is a k-form on M. This is the object constructed on the right hand side of (49.4.1).

We also illustrate it using the coordinates we introduced in the previous section. Recall that any horizontal k-form can be written in the form

$$\omega(x, y_\Lambda) = \omega_{a_1\cdots a_k}(x, y_\Lambda)\mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k} \,. \tag{49.4.2}$$

Its pullback to M along the r-jet of a section $\tilde{\sigma}_\epsilon$ is obtained by replacing the coordinate arguments y^μ_Λ by the partial derivatives $\partial_\Lambda y^\mu_\epsilon(x)$ of $\tilde{\sigma}_\epsilon$, so that one obtains

$$(j^r\tilde{\sigma}_\epsilon)^*(\omega)(x) = \omega_{a_1\cdots a_k}(x, \partial_\Lambda y_\epsilon(x))\mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k} \,. \tag{49.4.3}$$

Taking the derivative with respect to ϵ and using the chain rule yields the left hand side

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\epsilon}(j^r\tilde{\sigma}_\epsilon)^*(\omega)\bigg|_{\epsilon=0} &= \left[\frac{\mathrm{d}}{\mathrm{d}\epsilon}\partial_\Lambda y^\mu_\epsilon(x)\bigg|_{\epsilon=0} \bar{\partial}^\Lambda_\mu \omega_{a_1\cdots a_k}(x, \partial_\Lambda y(x))\right]\mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k} \\ &= \left[v^\mu_\Lambda(x)\bar{\partial}^\Lambda_\mu \omega_{a_1\cdots a_k}(x, \partial_\Lambda y(x))\right]\mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k} \,,\end{aligned} \tag{49.4.4}$$

where we used the coordinate expression $v^{\mu}_{\Lambda}(x)\bar{\partial}^{\Lambda}_{\mu}$ for the r-jet $j^r\xi$ of ξ. To compare with the right hand side, we calculate the exterior derivative

$$\mathrm{d}\omega(x, y_\Lambda) = \partial_b\omega_{a_1\cdots a_k}(x, y_\Lambda)\mathrm{d}x^b \wedge \mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k} + \bar{\partial}^{\Lambda}_{\mu}\omega_{a_1\cdots a_k}(x, y_\Lambda)\mathrm{d}y^{\mu}_{\Lambda} \wedge \mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_k}\,. \tag{49.4.5}$$

After inserting the vertical vector field ξ_r, which satisfies $\mathrm{d}y^{\mu}_{\Lambda}(\xi_r) = v^{\mu}_{\Lambda}(x)$ and $\mathrm{d}x^a(\xi_r) = 0$, and taking the pullback via $j^r\sigma$ we finally arrive at the same coordinate expression as for the left hand side.

If $\omega = L$ is a Lagrangian, we thus find that the action is stationary at the section $\sigma : M \to E$ if and only if

$$0 = \delta S = \left.\frac{\mathrm{d}S[\tilde{\sigma}_\epsilon]}{\mathrm{d}\epsilon}\right|_{\epsilon=0} = \int_M (j^r\sigma)^*\left(\iota_{\xi_r}(\mathrm{d}L)\right) \tag{49.4.6}$$

for every $\xi : M \to VE$ with $\nu \circ \xi = \omega$, i.e., if and only if

$$(j^r\sigma)^*\left(\iota_{\xi_r}(\mathrm{d}L)\right) \tag{49.4.7}$$

is exact for every $\xi : M \to VE$ with $\nu \circ \xi = \omega$. This already brings us closer to our goal. However, this expression is still rather cumbersome, as it requires calculating ξ_r and hence the r-jet $j^r\xi$ for every possible ξ, and checking whether the result is an exact form. To get rid of this calculation, we need another step.

49.5 Integration by parts

We will now further simplify the condition for a stationary action. In this section we discuss the question under which circumstances the pullback $(j^r\sigma)^*(\omega)$ of a horizontal k-form $\omega \in \Omega^{k,0}(J^r(E))$ is exact. The answer to this question is given by the following statement.

Theorem 49.5.1. *The pullback $(j^r\sigma)^*(\omega)$ of a d_H-exact horizontal k-form $\omega \in \Omega^{k,0}(J^r(E))$ is exact.*

We will not prove this here, but illustrate it using coordinates. Let ω be the k-form given by

$$\omega(x, y_\Lambda) = \omega_{\alpha_1\cdots\alpha_k}(x, y_\Lambda)\mathrm{d}x^{\alpha_1} \wedge \ldots \wedge \mathrm{d}x^{\alpha_k}\,. \tag{49.5.1}$$

For its pullback along the r-jet of a section σ we write

$$(j^r\sigma)^*(\omega)(x) = \omega_{\alpha_1\cdots\alpha_k}(x, \partial_\Lambda y(x))\mathrm{d}x^{\alpha_1} \wedge \ldots \wedge \mathrm{d}x^{\alpha_k}\,. \tag{49.5.2}$$

In the case that $\omega = \mathrm{d}_H\eta$ is d_H-exact, we have

$$\mathrm{d}_H\eta(x, y_\Lambda) = \left[\frac{\partial\eta_{\alpha_1\cdots\alpha_{k-1}}}{\partial x^\beta} + \sum_\Lambda y^a_{(\lambda_1,\ldots,\lambda_\beta+1,\ldots,\lambda_n)}\frac{\partial\eta_{\alpha_1\cdots\alpha_{k-1}}}{\partial y^a_\Lambda}\right]\mathrm{d}x^\beta \wedge \mathrm{d}x^{\alpha_1} \wedge \ldots \wedge \mathrm{d}x^{\alpha_{k-1}}\,, \tag{49.5.3}$$

whose pullback along $j^r\sigma$ is just the total derivative

$$(j^r\sigma)^*(\mathrm{d}_H\eta)(x) = \mathrm{d}(j^{r-1}\sigma)^*(\eta)(x)\,. \tag{49.5.4}$$

This yields the statement above for our chosen coordinates - of course we would have to show it in an coordinate independent fashion if we wanted a proof.

Returning to our original problem, we may thus add an arbitrary d_H-exact form $\mathrm{d}_H\eta \in \Omega^{n,0}(J^r(E))$ to $\iota_{\xi_r}(\mathrm{d}L)$ without changing the exactness of $(j^r\sigma)^*\left(\iota_{\xi_r}(\mathrm{d}L)\right)$. Here also the following statement will help.

Theorem 49.5.2. *For any horizontal k-form η and section $\xi : M \to VE$ holds*

$$\mathrm{d}_H(\iota_{\xi_{r-1}}\eta) = -\iota_{\xi_r}(\mathrm{d}_H\eta)\,. \tag{49.5.5}$$

The proof is straightforward. This in particular means that if η is d_H-exact, then also $\iota_{\xi_{r-1}}\eta$ is d_H-exact. For our problem thus follows that we may add an arbitrary d_H-exact form $\mathrm{d}_H\eta \in \Omega^{n,1}(J^r(E))$ to $\mathrm{d}L$. We define the following operator, which will yield us this form.

Definition 49.5.1 (Internal Euler operator)**.** Let $\pi : E \to M$ be a fiber bundle with $\dim M = n$ and $\Omega^{n,s}(J^r(E))$ with $s \geq 1$ the space of forms of type (n, s) on the r-jet bundle $J^r(E)$. The *internal Euler operator* is the unique function $\varrho : \Omega^{n,s}(J^r(E)) \to \Omega^{n,s}(J^{2r}(E))$ such that:

- ϱ is a projector (up to lifts to higher jet bundles): $\varrho^2 = \pi^*_{4r,2r} \circ \varrho$.
- For $\omega \in \Omega^{n,s}(J^r(E))$, the difference $\pi^*_{2r,r}\omega - \varrho(\omega)$ is d_H-exact, i.e., there exists $\eta \in \Omega^{n-1,s}(J^{2r-1}(E))$ such that $\mathrm{d}_H\eta = \pi^*_{2r,r}\omega - \varrho(\omega)$.
- ϱ vanishes on d_H-exact forms: $\varrho \circ \mathrm{d}_H = 0$.
- $\iota_X \circ \varrho = 0$ for all vector fields X on $J^{2r}(E)$ with $\pi_{2r,0*} \circ X = 0$.

We will not prove the existence and uniqueness of the internal Euler operator here, and we will not construct it explicitly. Instead, we will only provide the coordinate expression, which is given by

$$\begin{array}{rcccl} \varrho & : & \Omega^{n,s}(J^r(E)) & \to & \Omega^{n,s}(J^{2r}(E)) \\ & & \omega & \mapsto & \dfrac{1}{s}\displaystyle\sum_\Lambda (-1)^{|\Lambda|}\theta^a_{(0,\ldots,0)} \wedge \mathrm{D}_\Lambda\left(\iota_{\bar{\partial}^\Lambda_a}\omega\right) \end{array}\,. \tag{49.5.6}$$

Here we used the total derivative operator

$$\mathrm{D}_\Lambda = (\mathrm{D}_1)^{\lambda_1}\cdots(\mathrm{D}_n)^{\lambda_n}\,, \tag{49.5.7}$$

which acts on functions $f \in \Omega^0(J^r(E))$ as

$$\mathrm{D}_\alpha f = \partial_\alpha f + \sum_\Lambda y^a_{(\lambda_1,\ldots,\lambda_\alpha+1,\ldots,\lambda_n)}\bar{\partial}^\Lambda_a f\,. \tag{49.5.8}$$

To construct its action on higher degree forms, one uses the rules

$$\mathrm{D}_\alpha(\omega \wedge \eta) = \mathrm{D}_\alpha(\omega) \wedge \eta + \omega \wedge \mathrm{D}_\alpha(\eta)\,, \quad \mathrm{D}_\alpha(\mathrm{d}\omega) = \mathrm{d}(\mathrm{D}_\alpha\omega)\,. \tag{49.5.9}$$

From these follows in particular the action on the coordinate one-forms as

$$\mathrm{D}_\alpha \mathrm{d}x^\beta = 0\,, \quad \mathrm{D}_\alpha \mathrm{d}y^a_\Lambda = \mathrm{d}y^a_{(\lambda_1,\ldots,\lambda_\alpha+1,\ldots,\lambda_n)}\,. \tag{49.5.10}$$

Note that the total derivative is *not* an exterior derivative - it does not change the degree of a form, does not square to zero and is not an antiderivation.

Since ϱ is a projector, $\varrho(\Omega^{n,s}(J^r(E)))$ is an invariant subspace which we denote $\mathcal{F}^s(J^{2r}(E))$, which we can describe in coordinates as follows. Since a vector field X on $J^{2r}(E)$ with $\pi_{2r,0*} \circ X = 0$ has the coordinate expression

$$X = \sum_{|\Lambda|\geq 1} X^\Lambda_a \bar{\partial}^a_\Lambda\,, \tag{49.5.11}$$

it follows from the last condition in the definition of the internal Euler operator that the elements of $\mathcal{F}^s(J^{2r}(E))$ are of the form

$$\omega = \omega_{a_1\cdots a_s}\theta^{a_1}_{(0,\ldots,0)} \wedge \ldots \wedge \theta^{a_s}_{(0,\ldots,0)} \wedge \mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^n \,. \tag{49.5.12}$$

For a section $\xi : M \to VE$, which we expressed by the coordinate functions $y^a(x)$ and $v^a(x)$, we thus find that $\iota_{\xi_{2r}}\omega$ does not depend on the derivatives of the coordinate functions $v^a(x)$. We can thus simplify the task of finding sections $\sigma : M \to E$ for which the action is stationary by replacing $\mathrm{d}L$ with $\varrho(\mathrm{d}L)$.

49.6 Euler operator and Euler-Lagrange equations

We now finally use the results we derived so far and put them together. For this purpose we first introduce another helpful shorthand notation.

Definition 49.6.1 (Euler operator)**.** The *Euler operator* is the function $\mathcal{E} = \varrho \circ \mathrm{d} : \Omega^{n,0}(J^r(E)) \to \mathcal{F}^1(J^{2r}(E))$.

Writing a Lagrangian $L \in \Omega^{n,0}(J^r(E))$ in coordinates as $L = \mathcal{L}\mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^n$, we can write the Euler operator as

$$\mathcal{E}L = \mathcal{E}_\mu\mathcal{L}\,\theta^\mu_{(0,\ldots,0)} \wedge \mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^n \,, \tag{49.6.1}$$

where

$$\mathcal{E}_\mu\mathcal{L} = \sum_\Lambda (-1)^{|\Lambda|}\mathrm{D}_\Lambda\left(\bar{\partial}^\Lambda_\mu\mathcal{L}\right) \,. \tag{49.6.2}$$

To proceed with our problem of stationary actions, we come to another very helpful statement.

Theorem 49.6.1. *Let $\xi : M \to VE$ be a section and $\sigma = \nu \circ \xi : M \to E$. The pullback $(j^r\sigma)^*\left(\iota_{\xi_r}(\mathcal{E}L)\right)$ is exact if and only if it vanishes.*

This greatly simplifies our task. Instead of determining whether a differential form is exact, we need to check whether it vanishes. But we can simplify our task even more by the help of the following statement.

Theorem 49.6.2. *Let $\sigma : M \to E$ be a section. The pullback $(j^r\sigma)^*\left(\iota_{\xi_r}(\mathcal{E}L)\right)$ vanishes for all $\xi : M \to VE$ with $\sigma = \nu \circ \xi$ if and only if $(\mathcal{E}L) \circ \sigma = 0$.*

We have now found an amazingly simple condition. Given a Lagrangian $L \in \Omega^{n,0}(J^r(E))$ we can now determine the sections $\sigma : M \to E$ as the solutions of the following equation which in coordinates turns into a differential equation for σ:

Definition 49.6.2 (Euler-Lagrange equation)**.** Let $L \in \Omega^{n,0}(J^r(E))$ be a Lagrangian and $\sigma \in \Gamma_|(E)$ a local section. The *Euler-Lagrange equation* of L applied to σ is the equation

$$(\mathcal{E}L) \circ j^\infty\sigma = 0 \,. \tag{49.6.3}$$

We apply this definition to the following example.

Example 49.6.1. Recall that in example 49.1.2 we considered the Lagrangian

$$L = \mathcal{L}\mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^n \tag{49.6.4}$$

with

$$\mathcal{L} = \left(\frac{1}{2}g^{ab}(x)\phi_{,a}\phi_{,b} - V(\phi)\right)\sqrt{|\det g(x)|}\,. \tag{49.6.5}$$

In this case the fiber of the bundle $\pi : E \to M$ is one-dimensional, and so there is only one value for the index μ in the Euler-Lagrange equations (49.6.1) and (49.6.2). The vertical coordinate derivatives are given by

$$\bar{\partial}^{(0,\ldots,0)}\mathcal{L} = -V'(\phi)\sqrt{|\det g(x)|} \tag{49.6.6}$$

and

$$\bar{\partial}^{(0,\ldots,\lambda_a=1,\ldots,0)}\mathcal{L} = g^{ab}(x)\phi_{,b}\sqrt{|\det g(x)|}\,. \tag{49.6.7}$$

Taking the total derivative of the latter yields

$$\mathrm{D}_a\left[g^{ab}(x)\phi_{,b}\sqrt{|\det g(x)|}\right] = \Box\phi\sqrt{|\det g(x)|}\,. \tag{49.6.8}$$

In total, we therefore find

$$\mathcal{E}\mathcal{L} = -\left[\Box\phi + V'(\phi)\right]\sqrt{|\det g(x)|}\,. \tag{49.6.9}$$

In particular for the case $V(\phi) = \frac{1}{2}m^2\phi^2$, and hence $V'(\phi) = m^2\phi$, where $m \in \mathbb{R}$ denotes the mass, we obtain the field equation for the Klein-Gordon field.

49.7 Lepage forms

We now introduce another class of differential forms on jet bundles, which are useful for the description of Lagrangian systems and in particular their symmetries, which we will study in section 51. We define this class as follows.

Definition 49.7.1 (Lepage form)**.** A n-form $\rho \in \Omega^n(J^r(E))$ on the r-jet bundle $J^r(E)$ of a fiber bundle $\pi : E \to M$ over a n-dimensional manifold M is called a *Lepage form* if and only if the 1-contact component $\mathrm{p}_1\mathrm{d}\rho \in \Omega^{n,1}(J^{r+1}(E))$ is $\pi_{r+1,0}$-horizontal.

Recall that a form $\omega \in \Omega^{k+1}(J^{r+1}(E))$ is $\pi_{r+1,0}$-horizontal if and only if $\iota_X\omega = 0$ for all $\pi_{r+1,0}$-vertical vector fields $X \in \mathrm{Vect}(J^{r+1}(E))$, i.e., vector fields satisfying $\pi_{r+1,0*} \circ X = 0$. In jet bundle coordinates (x^a, y^μ_Λ), such a form ω is thus a linear combination of exterior products of the basis one-forms $\mathrm{d}x^a$ and $\mathrm{d}y^\mu$ only, with no occurrence of the derivative basis elements $\mathrm{d}y^\mu_\Lambda$ with $|\Lambda| \geq 1$. Alternatively, one may replace the basis one-forms $\mathrm{d}y^\mu$ with the contact basis θ^μ. For a Lepage form, the 1-contact component $\mathrm{p}_1\mathrm{d}\rho$ thus has the form

$$\mathrm{p}_1\mathrm{d}\rho = E_{\mu a_1\cdots a_n}\theta^\mu \wedge \mathrm{d}x^{a_1} \wedge \ldots \wedge \mathrm{d}x^{a_n}\,. \tag{49.7.1}$$

There exists a close relationship between Lepage forms and Lagrangians. To investigate this relationship, we start with the following definition.

Definition 49.7.2 (Lepage equivalent)**.** Let $\pi : E \to M$ be a fiber bundle and $L \in \Omega^{n,0}(J^r(E))$ a Lagrangian of order r on E. A *Lepage equivalent* of L is a Lepage form $\rho \in \Omega^n(J^s(E))$ such that $\mathrm{p}_0\rho = \pi^*_{s+1,r}L$.

One may ask whether every Lagrangian admits (at least) one Lepage equivalent. This is indeed the case, and we can even restrict its properties, in particular its jet bundle and contact order. This we can formulate as follows.

Theorem 49.7.1. *Let $\pi : E \to M$ be a fiber bundle and $L \in \Omega^{n,0}(J^r(E))$ a Lagrangian of order r on E. Then there exists an integer $s \leq 2r - 1$ and a Lepage equivalent $\rho \in \Omega^n(J^s(E))$ of L of contact order smaller or equal than one, i.e., $\mathrm{p}_i\rho = 0$ for $i \geq 2$.*

Proof. Recall that the Euler operator has the property

$$\mathcal{E}L - \pi^*_{2r,r}\mathrm{d}L = \varrho(\mathrm{d}L) - \pi^*_{2r,r}\mathrm{d}L = \mathrm{d}_H\eta\,, \tag{49.7.2}$$

for some $\eta \in \Omega^{n-1,1}(J^{2r-1}(E))$, since ϱ acts by adding a d_H-exact form. Then we define

$$\rho = \eta + \pi_{2r-1,r}L\,, \tag{49.7.3}$$

and find that it satisfies

$$\pi^*_{2r,2r-1}\mathrm{d}\rho = \mathrm{d}_H\eta + \mathrm{d}_V\eta + \pi_{2r,r}\mathrm{d}L\,. \tag{49.7.4}$$

Note that, by construction, $\mathrm{d}_H\eta \in \Omega^{n,1}(J^{2r}(E))$ and $\mathrm{d}_V\eta \in \Omega^{n-1,2}(J^{2r}(E))$, while $\pi_{2r,r}\mathrm{d}L \in \Omega^{n,1}(J^{2r}(E))$. Projecting to the first contact order thus gives

$$\mathrm{p}_1\mathrm{d}\rho = \mathrm{d}_H\eta + \pi_{2r,r}\mathrm{d}L = \mathcal{E}L\,. \tag{49.7.5}$$

By definition of the Euler-Lagrange operator, ▶...◀ ■

Once a Lepage equivalent of a Lagrangian is known, it can be used to simplify various tasks. So far our main goal was to find extremals of an action, by deriving the Euler-Lagrange equations. Now we see how this is achieved by using a Lepage equivalent.

Theorem 49.7.2. *Let $\pi : E \to M$ be a fiber bundle, $L \in \Omega^{n,0}(J^r(E))$ a Lagrangian of order r on E and $\rho \in \Omega^n(J^s(E))$ a Lepage equivalent of L. Then the Euler-Lagrange form of L is given by*

$$\mathcal{E}L = \pi^*_{2r,s+1}\mathrm{p}_1\mathrm{d}\rho\,. \tag{49.7.6}$$

Proof. ▶...◀ ■

Chapter 50

Variational bicomplex

50.1 Infinite jet space

We have seen in section 21.6 that for every fiber bundle $\pi : E \to M$ the jet spaces $J^r(E)$ for $r \in \mathbb{N}$ form an inverse sequence

$$M \xleftarrow{\pi} E \xleftarrow{\pi_{1,0}} J^1(E) \xleftarrow{\pi_{2,1}} J^2(E) \xleftarrow{\pi_{3,2}} \dots, \tag{50.1.1}$$

where the maps $\pi_{r,k} : J^r(E) \to J^k(E)$ are the projections of fiber bundles. For the purpose of this section, where we essentially follow [Sau89, ch. 7], we need to discuss what happens in the limit $r \to \infty$. We define this limit as follows.

Definition 50.1.1 (Infinite jet space)**.** Let $\pi : E \to M$ be a fiber bundle. Its *infinite jet space* is the projective limit

$$J^\infty(E) = \varprojlim J^r(E) = \left\{ (z_0, z_1, \dots) \in \bigtimes_{r=0}^{\infty} J^r(E) \,\middle|\, \forall k \leq r : \pi_{r,k}(z_r) = z_k \right\}. \tag{50.1.2}$$

An element of $J^\infty(E)$ is thus an infinite sequence of elements $z_r \in J^r(E)$ such that for all $k \leq r$ the condition $\pi_{r,k}(z_r) = z_k$ is satisfied. To understand the meaning of this, recall that a jet $z_r \in J^r(E)$ is an equivalence class of local sections of E such that their "partial derivatives up to order r" agree. The condition $\pi_{r,k}(z_r) = z_k$ here simply means that if $z_r = j^r_p\sigma$ for some point $p \in M$ and some local section $\sigma \in \Gamma_p(E)$, then $z_k = j^k_p\sigma$. In other words, any lower element z_k of this sequence is uniquely defined by any higher element z_r by throwing away any derivatives of order higher than k. Naively, we could thus just forget about almost all elements of the sequence and only look at the last one, which contains all derivatives - but of course, there is no such last element in an infinite sequence. So the only way to describe a section and "all of its infinitely many derivatives" is by an infinite sequence like the one above, and these sequences form the infinite jet space $J^\infty(E)$.

Given coordinates (x^a) on a trivializing neighborhood $U \in M$ and (y^μ) on the fiber F of the bundle $\pi : E \to M$, so that we have coordinates (x^a, y^μ) on $\pi^{-1}(U) \cong U \times F$ we have previously introduced coordinates (x^a, y^μ_Λ) with $0 \leq |\Lambda| \leq r$ on $\pi_r^{-1}(U) \subset J^r(E)$. We get (infinitely many) coordinates on $J^\infty(E)$ by dropping the upper bound and allowing all multiindices Λ with $|\Lambda| \in \mathbb{N}$.

Note that $J^\infty(E)$ is *not* a manifold in the sense we defined manifolds - it is not locally diffeomorphic to any finite-dimensional Euclidean space $\mathbb{R}^n$. It has *some* properties of a manifold, so that *some* operations on manifolds can be generalized to $J^\infty(E)$, but not all of them, so we have to be careful when working with this object. The following notions can nicely be generalized.

Definition 50.1.2 (∞-jet projection)**.** Let $\pi : E \to M$ be a fiber bundle. For $r \in \mathbb{N}$ we define the *∞-jet projection*

$$\begin{array}{cccc} \pi_{\infty,r} & : & J^\infty(E) & \to & J^r(E) \\ & & (z_0, z_1, \dots) & \mapsto & z_r \end{array} . \tag{50.1.3}$$

The function $\pi_{\infty,0} : J^\infty(E) \to E$ is called the *target projection*, while $\pi_\infty = \pi \circ \pi_{\infty,0} : J^\infty(E) \to M$ is called the *source projection*.

As it is also the case for finite jet bundles, these projections throw away all derivatives of higher order than a fixed $r \in \mathbb{N}$.

Definition 50.1.3 (∞-jet of a section)**.** Let $\pi : E \to M$ be a fiber bundle, $p \in M$ and $\sigma \in \Gamma_p(E)$ a local section whose domain contains p. We define the *∞-jet* $j^\infty_p\sigma$ of σ at p as the infinite sequence

$$(j^0_p\sigma, j^1_p\sigma, \dots) \in J^\infty(E) . \tag{50.1.4}$$

The ∞-jet is the object which captures "all derivatives" of a local section σ at some point $p \in M$. Also one easily checks that this is an element of the infinite jet space, since

$$\pi_{r,k}(j^r_p\sigma) = j^k_p\sigma \tag{50.1.5}$$

for all $k \le r$, by definition of a jet. If we let p run over all points of the domain of σ, we obtain the following notion.

Definition 50.1.4 (∞-jet prolongation)**.** Let $\pi : E \to M$ be a fiber bundle and $\sigma \in \Gamma_{|U}(E)$ a local section with domain $U \subset M$. Its *∞-jet prolongation* is the function

$$\begin{array}{cccc} j^\infty\sigma & : & U & \to & J^\infty(E) \\ & & p & \mapsto & j^\infty_p\sigma \end{array} . \tag{50.1.6}$$

As in the finite-dimensional case, taking the ∞-jet at each point p in the domain of σ yields its prolongation into $J^\infty(E)$.

50.2 Variational bicomplex

Another concept that can nicely be generalized to $J^\infty(E)$ is that of differential forms. Note that the pullbacks along the projection maps define a sequence

$$\Omega^k(M) \xrightarrow{\pi^*} \Omega^k(E) \xrightarrow{\pi^*_{1,0}} \Omega^k(J^1(E)) \xrightarrow{\pi^*_{2,1}} \Omega^k(J^2(E)) \xrightarrow{\pi^*_{3,2}} \dots \tag{50.2.1}$$

for all $k \in \mathbb{N}$. Here it makes sense to consider k-forms with arbitrarily high k, since the dimension of the manifolds $J^r(E)$ is growing with r, so there will be non-trivial k-forms for any k. We can use this sequence to define the following object.

Definition 50.2.1 (Pullback to $J^\infty(E)$)**.** Let $\pi : E \to M$ be a fiber bundle and $J^\infty(E)$ its infinite jet space. For $k \in \mathbb{N}$ we define the space $\Omega^k(J^\infty(E))$ as the direct limit

$$\Omega^k(J^\infty(E)) = \varinjlim \Omega^k(J^r(E)) = \biguplus_{r=0}^{\infty} \Omega^k(J^r(E)) \Big/ \sim\,, \tag{50.2.2}$$

where two k-forms $\omega \in \Omega^k(J^r(E))$ and $\chi \in \Omega^k(J^{r'}(E))$ are considered equivalent, $\omega \sim \chi$, if and only if there exists $r'' \geq \max(r, r')$ such that $\pi^*_{r'',r}(\omega) = \pi^*_{r'',r'}(\chi)$. The equivalence class of $\omega \in \Omega^k(J^r(E))$ is denoted $\pi^*_{\infty,r}(\omega) \in \Omega^k(J^\infty(E))$ and called the *pullback* of ω to $J^\infty(E)$.

Note that despite the notation, the elements of $\Omega^k(J^\infty(E))$ are *not* differential forms, and $J^\infty(E)$ is not a manifold, so we cannot immediately use any operations which we defined on differential forms. Instead, they are equivalence classes of differential forms on finite jet spaces. Note that since for $r' \geq r$ the pullbacks $\pi^*_{r',r} : \Omega^k(J^r(E)) \to \Omega^k(J^{r'}(E))$ are injective functions (which is a consequence of the fact that the maps $\pi_{r',r} : J^{r'}(E) \to J^r(E)$ are surjective), two k-forms $\omega \in \Omega^k(J^r(E))$ and $\chi \in \Omega^k(J^{r'}(E))$ are equivalent if and only if $\pi^*_{r',r}\omega = \chi$. In other words, we identify all elements of $\Omega^k(J^r(E))$ with their images in $\Omega^k(J^{r'}(E))$. We thus obtain a sequence of inclusions

$$\Omega^k(J^0(E)) \tilde{\subset} \Omega^k(J^1(E)) \tilde{\subset} \ldots \tilde{\subset} \Omega^k(J^\infty(E))\,, \tag{50.2.3}$$

where $\tilde{\subset}$ should be read as "the set formed by equivalence classes of elements contained in the set on the left is a subset of the set formed by equivalence classes of elements contained in the set on the right".

To illustrate this definition, we discuss how to write the elements of $\Omega^k(J^\infty(E))$ using the coordinates (x^a, y^μ_Λ) we introduced on (finite and infinite) jet bundles. Any k-form $\omega \in \Omega^k(J^r(E))$ on a finite jet bundle $J^r(E)$ can be written as a finite linear combination of k-fold wedge products of the coordinate one-forms dx^a, dy^μ_Λ, where $|\Lambda| \leq r$. The pullback $\pi^*_{r',r}(\omega) \in \Omega^k(J^{r'}(E))$ of ω is a k-form on $J^{r'}(E)$ which has the same coordinate representation. Hence, the equivalence relation we introduced above simply identifies k-forms if and only if their coordinate representations in the coordinates (x^a, y^μ_Λ) agree. We can thus formally write an element of $\Omega^k(J^\infty(E))$ as a finite linear combination of k-fold wedge products of the coordinate one-forms dx^a, dy^μ_Λ, where $|\Lambda| \in \mathbb{N}$. Note, however, that so far this is only a notation - we have not defined a wedge product of such equivalence classes yet. But actually we can do so.

Definition 50.2.2 (Exterior product)**.** Let $\pi : E \to M$ be a fiber bundle and $\omega \in \Omega^k(J^\infty(E))$, $\chi \in \Omega^l(J^\infty(E))$. By definition, we can find $r \in \mathbb{N}$, $\bar{\omega} \in \Omega^k(J^r(E))$ and $\bar{\chi} \in \Omega^l(J^r(E))$ such that $\omega = \pi^*_{\infty,r}(\bar{\omega})$ and $\chi = \pi^*_{\infty,r}(\bar{\chi})$. We define the *exterior product* (or wedge product)

$$\omega \wedge \chi = \pi^*_{\infty,r}(\bar{\omega} \wedge \bar{\chi}) \in \Omega^{k+l}(J^\infty(E))\,. \tag{50.2.4}$$

We need a few remarks on this definition. First, note that we can always pick representatives $\bar{\omega}, \bar{\chi}$ of the equivalence classes ω, χ which are differential forms on the same jet space $J^r(E)$. If we had picked one of them to be on a different jet space $J^{r'}(E)$ with $r' < r$, we could just obtain another representative on $J^r(E)$ by applying the pullback $\pi^*_{r,r'}$. Further, the wedge product above is well-defined, i.e., independent of the choice of the jet space $J^r(E)$ from which we take the representatives, since the pullback distributes over wedge products. Note that in coordinates the wedge product just looks as it always looks like for ordinary differential forms, so we can just calculate it as usual. The same applies to the exterior derivative, which we define as follows.

Definition 50.2.3 (Exterior derivative)**.** Let $\pi : E \to M$ be a fiber bundle and $\omega \in \Omega^k(J^\infty(E))$. By definition, we can find $r \in \mathbb{N}$ and $\bar{\omega} \in \Omega^k(J^r(E))$ such that $\omega = \pi^*_{\infty,r}(\bar{\omega})$. We define the *exterior derivative*

$$\mathrm{d}\omega = \pi^*_{\infty,r}(\mathrm{d}\bar{\omega}) \in \Omega^{k+1}(J^\infty(E)) \,. \tag{50.2.5}$$

Also this is well-defined, since pullbacks and the exterior derivative commute. Also the exterior derivative looks in coordinates just as in the finite case. Finally, it is also easy to prove that the exterior derivative and exterior product satisfy all their nice properties which they also have for ordinary differential forms on finite-dimensional manifolds. We can therefore just use them as we would naturally do. In particular, the exterior derivative satisfies $\mathrm{d}^2 = 0$, so that we have an infinite sequence

$$\Omega^0(J^\infty(E)) \xrightarrow{\mathrm{d}} \Omega^1(J^\infty(E)) \xrightarrow{\mathrm{d}} \dots \,, \tag{50.2.6}$$

where the image of each function lies inside the kernel of the next one. This structure is called a *complex*. We will further refine this structure, and for this purpose need to decompose it further, generalizing what we have done in section 21.9. This leads us to the following definition.

Definition 50.2.4 (Horizontal form)**.** Let $\pi : E \to M$ be a fiber bundle. An element $\omega \in \Omega^k(J^\infty(E))$ is called *horizontal* if it is the pullback $\omega = \pi^*_{\infty,r}(\bar{\omega})$ of a horizontal k-form $\bar{\omega} \in \Omega^k(J^r(E))$, i.e., such that $\bar{\omega}$ vanishes on the kernel $\ker \pi_{r*}$ of $\pi_{r*} : TJ^r(E) \to TM$. The subspace of horizontal elements of $\Omega^k(J^\infty(E))$ is denoted $\Omega^{k,0}(J^\infty(E))$.

Here $\bar{\omega}$ is horizontal in the sense of definition 21.9.1. Recall that the kernel of π_{r*} is defined as the set of tangent vectors $v \in TJ^r(E)$ for which $\pi_{r*}(v) = 0$. These tangent vectors are tangent to the fibers $\pi_r^{-1}(p) \cong J^r_p(E)$ for $p \in M$. In coordinates (x^a, y^μ_Λ) on $J^r(E)$ the space of vertical vectors is spanned by the vector fields $\bar{\partial}^\Lambda_\mu = \partial/\partial y^\mu_\Lambda$. A k-form $\bar{\omega} \in \Omega^k(J^r(E))$ vanishes on these vectors if its coordinate representation contains only wedge products of $\mathrm{d}x^a$, but no $\mathrm{d}y^\mu_\Lambda$. The same holds for the coordinate representation of $\omega \in \Omega^k(J^\infty(E))$.

We also define a suitable counterpart, by generalizing definition 21.9.2.

Definition 50.2.5 (Contact form)**.** Let $\pi : E \to M$ be a fiber bundle. An element $\omega \in \Omega^k(J^\infty(E))$ is called a *contact form* if its pullback $(j^\infty\sigma)^*(\omega) \in \Omega^k(M)$ vanishes for every local section σ of $\pi : E \to M$. The subspace of contact forms of $\Omega^k(J^\infty(E))$ is denoted $\Omega^{0,k}(J^\infty(E))$.

The meaning of the pullback $(j^\infty\sigma)^*(\omega)$ for $\omega \in \Omega^k(J^\infty(E))$ should be almost clear. We can pick a representative $\bar{\omega} \in \Omega^k(J^r(E))$, and take its pullback $(j^r\sigma)^*(\bar{\omega}) \in \Omega^k(M)$ along the map $j^r\sigma : M \to J^r(E)$. This pullback is independent of the choice of the representative, and so defines a unique pullback $(j^\infty\sigma)^* : \Omega^k(J^\infty(E)) \to \Omega^k(M)$.

Using coordinates (x^a, y^μ_Λ) on $J^\infty(E)$ it is easy to write down a basis for the space $\Omega^{0,1}(J^\infty(E))$ of contact one-forms, analog to definition 21.9.3. We define the *basic contact forms* θ^μ_Λ as follows.

Definition 50.2.6 (Basic contact one-form)**.** Let $\pi : E \to M$ be a fiber bundle, $U \subset M$ and (x^a), (x^a, y^μ) and (x^a, y^μ_Λ) with $|\Lambda| \in \mathbb{N}$ local coordinates on U, $\pi^{-1}(U)$ and $\pi_\infty^{-1}(U)$, respectively. The *basic contact one-forms* with respect to these coordinates are the one-forms

$$\theta^\mu_\Lambda = \mathrm{d}y^\mu_\Lambda - y^\mu_{(\lambda_1+1,\lambda_2,\ldots,\lambda_n)}\mathrm{d}x^1 - y^\mu_{(\lambda_1,\lambda_2+1,\ldots,\lambda_n)}\mathrm{d}x^2 - \ldots - y^\mu_{(\lambda_1,\lambda_2,\ldots,\lambda_n+1)}\mathrm{d}x^n\,. \tag{50.2.7}$$

As in the finite-dimensional case, any contact one-form θ can be written as a *finite* sum in the form $\theta = f^\Lambda_a\theta^a_\Lambda$. Note, however, that this basis has infinitely many elements, and that there is no upper bound on $|\Lambda|$. This means that for *any* $\mathrm{d}y^\mu_\Lambda$ one can define a suitable contact one-form θ^μ_Λ, in contrast to the finite-dimensional case, where this is possible only for $|\Lambda| < r$. Hence, one can define a basis on the space $\Omega^1(J^\infty(E))$ as follows.

Definition 50.2.7 (Contact basis)**.** Let $\pi : E \to M$ be a fiber bundle, $U \subset M$ and (x^a), (x^a, y^μ) and (x^a, y^μ_Λ) with $|\Lambda| \in \mathbb{N}$ local coordinates on U, $\pi^{-1}(U)$ and $\pi_\infty^{-1}(U)$, respectively. The *contact basis* with respect to these coordinates is the basis of $\Omega^1(J^\infty(E))$ given by

$$(\mathrm{d}x^a, \theta^\mu_\Lambda)\,. \tag{50.2.8}$$

One can see from the structure of the contact basis that every one-form on the infinite jet bundle uniquely decomposes into horizontal and contact parts. Hence, $\Omega^1(J^\infty(E)) = \Omega^{1,0}(J^\infty(E)) \oplus \Omega^{0,1}(J^\infty(E))$. We now aim to generalize this to higher k-forms. This is indeed possible, due to the following property, as we will see below. Note first that the following properties are adapted directly from the finite-dimensional statements 21.9.1 and 21.9.2.

Theorem 50.2.1. *The exterior product of two horizontal forms on the infinite jet space is again horizontal.*

Theorem 50.2.2. *The contact forms on the infinite jet space form an ideal (the* contact ideal*) of the exterior algebra, i.e., the exterior product of an arbitrary form and a contact form is again a contact form.*

This is not difficult to prove - it follows immediately from the corresponding finite-dimensional cases and the fact that the pullback distributes over wedge products. Again like in the finite-dimensional case, we can define the following.

Definition 50.2.8 (l-contact form)**.** Let $\pi : E \to M$ be a fiber bundle. A $(k+l)$-form $\omega \in \Omega^{k+l}(J^\infty(E))$ on the infinite jet space is called l*-contact* if it is a linear combination of exterior products of k horizontal one-forms and l contact one-forms. The space of all such forms is denoted $\Omega^{k,l}(J^\infty(E))$.

We now come to an important difference compared to the finite-dimensional case. It turns out that using horizontal and contact forms, we can generate all of $\Omega^k(J^\infty(E))$ as a consequence of the following property.

Theorem 50.2.3. *For each $k \in \mathbb{N}$, the space $\Omega^k(J^\infty(E))$ splits into a direct sum*

$$\Omega^k(J^\infty(E)) = \bigoplus_{i=0}^{k}\Omega^{k-i,i}(J^\infty(E))\,, \tag{50.2.9}$$

so that every $\omega \in \Omega^k(J^\infty(E))$ *decomposes uniquely as*

$$\omega = \sum_{i=0}^{k} \mathrm{p}_i \omega \,, \tag{50.2.10}$$

where $\mathrm{p}_i\omega \in \Omega^{k-i,i}(J^\infty(E))$.

Recall from theorem 21.9.3 that in the finite-dimensional case, one does not decompose ω itself, but its lift to the next higher jet bundle, so that also the operators p_i were defined to include this lift. Here the situation becomes simpler, and the operators p_i become projectors onto vector subspaces. In coordinates, every such space $\Omega^{k,l}(J^\infty(E))$ is spanned by wedge products of the form

$$\mathrm{d}x^{\alpha_1} \wedge \ldots \wedge \mathrm{d}x^{\alpha_k} \wedge \theta^{a_1}_{\Lambda_1} \wedge \ldots \wedge \theta^{a_l}_{\Lambda_l} \,. \tag{50.2.11}$$

It is now easy to see how the exterior product and exterior derivative behave under this splitting.

Theorem 50.2.4. *Let* $\omega \in \Omega^{k,l}(J^\infty(E))$ *and* $\chi \in \Omega^{k',l'}(J^\infty(E))$*. Then the following hold:*

1. $\omega \wedge \chi \in \Omega^{k+k',l+l'}(J^\infty(E))$,
2. $\mathrm{d}\omega \in \Omega^{k+1,l}(J^\infty(E)) \oplus \Omega^{k,l+1}(J^\infty(E))$.

This property of the exterior product is immediately clear. For the exterior derivative it means that $\mathrm{d}\omega$ can be uniquely written as the sum of two terms, one of them belonging to $\Omega^{k+1,l}(J^\infty(E))$, the other one to $\Omega^{k,l+1}(J^\infty(E))$. Note again the difference with theorem 21.9.4, where we did not decompose $\mathrm{d}\omega$, but again its lift to the next higher jet bundle, which is not necessary here. This now allows us to decompose the exterior derivative in the following way.

Definition 50.2.9 (Horizontal and vertical differentials). Let $\pi : E \to M$ be a fiber bundle. For $k, l \in \mathbb{N}$ the *horizontal* (or total) differential

$$\begin{array}{rcccc} \mathrm{d}_H & : & \Omega^{k,l}(J^\infty(E)) & \to & \Omega^{k+1,l}(J^\infty(E)) \\ & & \omega & \mapsto & \mathrm{p}_l \mathrm{d}\omega \end{array} \tag{50.2.12}$$

and *vertical* differential

$$\begin{array}{rcccc} \mathrm{d}_V & : & \Omega^{k,l}(J^\infty(E)) & \to & \Omega^{k,l+1}(J^\infty(E)) \\ & & \omega & \mapsto & \mathrm{p}_{l+1} \mathrm{d}\omega \end{array} \tag{50.2.13}$$

are the unique functions such that $\mathrm{d}_H\omega + \mathrm{d}_V\omega = \mathrm{d}\omega$ for all $\omega \in \Omega^{k,l}(J^\infty(E))$.

Note that in contrast to definition 21.9.6, we could omit the pullback in the last line. This is also the case for the following properties, which generalize theorem 21.9.5.

Theorem 50.2.5. *For each* $\omega \in \Omega^{k,l}(J^r(E))$ *and* $\chi \in \Omega^{k',l'}(J^r(E))$ *the horizontal and vertical differentials* d_H *and* d_V *satisfy:*

1. d_H *and* d_V *are antiderivations:*

$$\mathrm{d}_H(\omega \wedge \chi) = \mathrm{d}_H\omega \wedge \chi + (-1)^{k+l}\omega \wedge \mathrm{d}_H\chi \,, \tag{50.2.14a}$$

$$\mathrm{d}_V(\omega \wedge \chi) = \mathrm{d}_V\omega \wedge \chi + (-1)^{k+l}\omega \wedge \mathrm{d}_V\chi \,. \tag{50.2.14b}$$

2. $\mathrm{d}_H^2 = 0$, $\mathrm{d}_V^2 = 0$ *and* $\mathrm{d}_H\mathrm{d}_V = -\mathrm{d}_V\mathrm{d}_H$.

With these properties we can now construct the coordinate expressions for d_H and d_V, in analogy to the finite-dimensional case studied in section 21.9. For $f \in \Omega^{0,0}(J^\infty(E))$ we have the vertical differential given by

$$\mathrm{d}_V f = \frac{\partial f}{\partial y^\mu_\Lambda} \theta^\mu_\Lambda \in \Omega^{0,1}(J^\infty(E))\,, \tag{50.2.15}$$

where, in contrast to (21.9.19), the summation over Λ now goes over the infinite set $|\Lambda| \in \mathbb{N}$, although only finitely many terms are non-vanishing. For the horizontal differential then follows

$$\mathrm{d}_H f = \mathrm{d}f - \mathrm{d}_V f = \mathrm{D}_a f \, \mathrm{d}x^a \in \Omega^{1,0}(J^\infty(E))\,, \tag{50.2.16}$$

where the total derivative

$$\mathrm{D}_a f = \frac{\partial f}{\partial x^a} + \sum_{|\Lambda|=0}^{\infty} y^\mu_{(\lambda_1,\ldots,\lambda_\alpha+1,\ldots,\lambda_n)} \frac{\partial f}{\partial y^\mu_\Lambda} = \frac{\partial f}{\partial x^a} + \sum_{|\Lambda|=0}^{\infty} y^\mu_{\Lambda a} \frac{\partial f}{\partial y^\mu_\Lambda} \tag{50.2.17}$$

now also turns into an infinite sum, with finitely many non-vanishing terms. For the horizontal and vertical coordinate differentials we have

$$\mathrm{d}_H(\mathrm{d}x^a) = 0\,, \quad \mathrm{d}_V(\mathrm{d}x^a) = 0\,, \quad \mathrm{d}_H(\mathrm{d}y^\mu_\Lambda) = 0\,, \quad \mathrm{d}_V(\mathrm{d}y^\mu_\Lambda) = 0\,, \tag{50.2.18}$$

which again follows from the fact that they are closed, i.e., their exterior derivatives vanish. Finally, the basic contact forms satisfy

$$\mathrm{d}_H \theta^\mu_\Lambda = \mathrm{d}x^1 \wedge \theta^\mu_{(\lambda_1+1,\lambda_2,\ldots,\lambda_n)} + \mathrm{d}x^2 \wedge \theta^\mu_{(\lambda_1,\lambda_2+1,\ldots,\lambda_n)} + \ldots + \mathrm{d}x^n \wedge \theta^\mu_{(\lambda_1,\lambda_2,\ldots,\lambda_n+1)} \tag{50.2.19}$$

as well as

$$\mathrm{d}_V \theta^\mu_\Lambda = 0\,. \tag{50.2.20}$$

Since any differential form on $J^\infty(E)$ can be constructed as a linear combination of wedge products of the forms above, as in the finite-dimensional case, we can thus explicitly calculate the vertical and horizontal differentials for all differential forms. In this case, however, we can even go one step further. We thus return to the sequence induced by the exterior derivative $\mathrm{d} : \Omega^k(J^\infty(E)) \to \Omega^{k+1}(J^\infty(E))$. Using the horizontal and vertical differentials we can construct a similar structure, which is not a complex, but a *bicomplex*, which we define as follows.

Definition 50.2.10 (Variational bicomplex)**.** Let $\pi : E \to M$ be a fiber bundle with $\dim M = n$. The *variational bicomplex* is the structure

$$\begin{array}{ccccccc}
\vdots & & \vdots & & \vdots & & \vdots \\
\uparrow & & \uparrow & & \uparrow & & \uparrow \\
\Omega^{0,1}(J^\infty(E)) & \xrightarrow{\mathrm{d}_H} & \Omega^{1,1}(J^\infty(E)) & \dashrightarrow & \Omega^{n-1,1}(J^\infty(E)) & \xrightarrow{\mathrm{d}_H} & \Omega^{n,1}(J^\infty(E)) \\
\mathrm{d}_V \uparrow & & \mathrm{d}_V \uparrow & & \mathrm{d}_V \uparrow & & \mathrm{d}_V \uparrow \\
\Omega^{0,0}(J^\infty(E)) & \xrightarrow{\mathrm{d}_H} & \Omega^{1,0}(J^\infty(E)) & \dashrightarrow & \Omega^{n-1,0}(J^\infty(E)) & \xrightarrow{\mathrm{d}_H} & \Omega^{n,0}(J^\infty(E))
\end{array} \tag{50.2.21}$$

The variational bicomplex offers another description of physical systems in the Lagrangian formulation, as we will see next.

50.3 Vector fields on the infinite jet space

Another concept we must generalize is that of vector fields on jet spaces. Recall from section 49.3 that we described the variation of a local section σ of the bundle $\pi : E \to M$ by vertical vector fields, and their prolongation to jet spaces. Since we are dealing with an infinite jet space now, we must clarify how to describe (vertical) tangent vectors to this structure. We use the following definition.

Definition 50.3.1 (Tangent bundle of $J^\infty(E)$)**.** Let $\pi : E \to M$ be a fiber bundle and $J^\infty(E)$ its infinite jet space. The *tangent bundle* of $J^\infty(E)$ is the projective limit

$$TJ^\infty(E) = \varprojlim TJ^r(E) = \left\{ (v_0, v_1, \ldots) \in \bigtimes_{r=0}^{\infty} TJ^r(E) \middle| \forall k \leq r : \pi_{r,k*}(v_r) = v_k \right\} . \tag{50.3.1}$$

Analogously, we define:

Definition 50.3.2 (Vertical tangent bundle of $J^\infty(E)$)**.** Let $\pi : E \to M$ be a fiber bundle and $J^\infty(E)$ its infinite jet space. The *vertical tangent bundle* of $J^\infty(E)$ is the projective limit

$$VJ^\infty(E) = \varprojlim VJ^r(E) = \left\{ (v_0, v_1, \ldots) \in \bigtimes_{r=0}^{\infty} VJ^r(E) \middle| \forall k \leq r : \pi_{r,k*}(v_r) = v_k \right\} . \tag{50.3.2}$$

In other words, an element of $VJ^\infty(E)$ is an infinite sequence of elements $v_r \in V^r(E)$ such that for all $k \leq r$ the condition $\pi_{r,k*}(v_r) = v_k$ is satisfied.

We also introduce suitable coordinate bases on $TJ^\infty(E)$ and $VJ^\infty(E)$. Recall that on $J^\infty(E)$ we used coordinates (x^a, y^μ_Λ) derived from the coordinates (x^a) on M and (x^a, y^μ) on E. An element of a tangent space $T_qJ^\infty(E)$ with $q \in J^\infty(E)$ can be written in the form

$$u^a \frac{\partial}{\partial x^a} + v^\mu_\Lambda \frac{\partial}{\partial y^\mu_\Lambda} = u^a \partial_a + v^\mu_\Lambda \bar{\partial}^\Lambda_\mu . \tag{50.3.3}$$

This yields us coordinates (u^a, v^μ_Λ) on $T_qJ^\infty(E)$, and thus coordinates $(x^a, y^\mu_\Lambda, u^a, v^\mu_\Lambda)$ on $TJ^\infty(E)$, where the first half specifies the point q and the second half the tangent vector at q. On the vertical tangent bundle $VJ^\infty(E)$ one thus has coordinates $(x^a, y^\mu_\Lambda, v^\mu_\Lambda)$. These coordinates have the same form as in the finite-dimensional case in section 49.3, but in this case the order $|\Lambda|$ may be arbitrary.

Note that there exist well-defined functions

$$\begin{array}{cccc} \tau_\infty & : & TJ^\infty(E) & \to & J^\infty(E) \\ & & (v_0, v_1, \ldots) & \mapsto & (\tau_0(v_0), \tau_1(v_1), \ldots) \end{array} \tag{50.3.4}$$

and

$$\begin{array}{cccc} \nu_\infty & : & VJ^\infty(E) & \to & J^\infty(E) \\ & & (v_0, v_1, \ldots) & \mapsto & (\nu_0(v_0), \nu_1(v_1), \ldots) \end{array} , \tag{50.3.5}$$

where $\tau_r : TJ^r(E) \to J^r(E)$ and $\nu_r : VJ^r(E) \to J^r(E)$ are the projections of the (vertical) tangent bundle to the finite-dimensional jet spaces. One may thus naively regard the (vertical) tangent bundle as an ordinary fiber bundle over the infinite jet space, and consider its section.

However, this naive construction leads to technical difficulties, and so we use a more formal, abstract definition, which is adapted from [GMS09, sec. 1.7].

Definition 50.3.3 (Vector field on $J^\infty(E)$)**.** Let $\pi : E \to M$ be a fiber bundle. The space $\mathrm{Vect}(J^\infty(E))$ of *vector fields* on $J^\infty(E)$ is the dual $\Omega^0(J^\infty(E))$ module of $\Omega^1(J^\infty(E))$.

This needs a few clarifications, which are most easily expressed by considering the coordinate expressions of the mentioned spaces. First note that $\Omega^0(J^\infty(E))$ consists of those functions $f \in \Omega^0(J^\infty(E))$, which depend on only finitely many coordinates, since every such function must have a representative $\bar{f} \in C^\infty(J^r(E), \mathbb{R})$ on a finite-dimensional jet space. Similarly, also a one-form $\omega \in \Omega^1(J^\infty(E))$ has a representative $\bar{\omega} \in \Omega^1(J^r(E))$, and hence is of the form

$$\omega = \omega_a \mathrm{d}x^a + \bar{\omega}_\mu^\Lambda \mathrm{d}y_\Lambda^\mu\,, \tag{50.3.6}$$

with an arbitrary, but finite number of non-vanishing terms, and coefficients depending on an arbitrary, but finite number of coordinates. These one-forms form a $\Omega^0(J^\infty(E))$ module, since $f\omega \in \Omega^1(J^\infty(E))$ for $f \in \Omega^0(J^\infty(E))$ and $\omega \in \Omega^1(J^\infty(E))$. Following definition 50.3.3, a vector field on $J^\infty(E)$ is an element of the dual module, hence a module homomorphism $\xi : \Omega^1(J^\infty(E)) \to \Omega^0(J^\infty(E))$. In other words, there exists a pairing, which we suggestively write as an interior product $\iota_\xi\omega \in \Omega^0(J^\infty(E))$. Using coordinates allows us to identify ξ as the formal sum

$$\xi = \xi^a \partial_a + \bar{\xi}_\Lambda^\mu \bar{\partial}_\mu^\Lambda\,, \tag{50.3.7}$$

with $(\partial_a, \bar{\partial}_\mu^\Lambda)$ identified as the dual basis of $(\mathrm{d}x^a, \mathrm{d}y_\Lambda^\mu)$, so that the interior product reads

$$\iota_\xi\omega = \xi^a\omega_a + \bar{\xi}_\Lambda^\mu \bar{\omega}_\mu^\Lambda \in \Omega^0(J^\infty(E))\,. \tag{50.3.8}$$

Note that ω has only finitely many non-vanishing components, and so ξ may have *infinitely* many components, while still yielding a finite sum $\iota_\xi\omega$. However, every component of ξ) may depend only on *finitely* many coordinates, in order for $\iota_\xi\omega$ to depend on finitely many coordinates as well.

For later use, we follow [GMS09, sec. 2.2] to define a particular class of vector fields on the infinite jet space, which is closely related to the variational bicomplex introduced in the preceding section.

Definition 50.3.4 (Contact vector field)**.** A vector field $\xi \in \mathrm{Vect}(J^\infty(E))$ is called a *contact vector field* if it preserves the contact ideal, i.e., if

$$\mathcal{L}_\xi\theta = \iota_\xi \mathrm{d}\theta + \mathrm{d}\iota_\xi\theta \tag{50.3.9}$$

is a contact form for every contact form θ.

There exist a number of important constructions which yield vector fields on the infinite jet space. To define these, it is helpful to first generalize the concept of a vector field on the manifolds M and E.

Definition 50.3.5 (Generalized vector field)**.** Let $\pi : E \to M$ be a fiber bundle. A *generalized vector field*...

1. ...on M is an equivalence class of sections $X \in \Gamma(\pi_r^*TM)$ of the pullback bundle π_r^*TM,

2. ...on E is an equivalence class of sections $X \in \Gamma(\pi_{r,0}^*TE)$ of the pullback bundle $\pi_{r,0}^*TE$,

where two sections X, Y of jet order $r \leq r'$ are regarded equivalent if and only if $Y = X \circ \pi_{r',r}$.

Written in coordinates, a generalized vector field has the usual form $X^a\partial_a$ for a generalized vector field on M, and $X^a\partial_a + \bar{X}^\mu\bar{\partial}_\mu$ for a vector field on E, but with component functions which depend on an arbitrary, but finite number of jet bundle coordinates (x^a, y^μ). By a slight abuse of notation, we denote the corresponding spaces of generalized vector fields by $\Gamma(\pi_\infty^*TM)$ and $\Gamma(\pi_{\infty,0}^*TE)$, respectively.

We may now define a number of useful operations on generalized vector fields.

Definition 50.3.6 (Prolongation)**.** Let $\pi : E \to M$ be a fiber bundle and $X \in \Gamma(\pi_{\infty,0}^*TE)$ be a generalized vector field on E. Its *prolongation* is the unique map $\operatorname{pr} X : J^\infty(E) \to TJ^\infty(E)$ which satisfies:

1. it projects to X on E: $X = \pi_{\infty,0*} \circ \operatorname{pr} X$,

2. $\operatorname{pr} X$ is a contact vector field.

It is instructive to derive a coordinate expression for the prolongation. Let

$$X(x, y_\Lambda) = X^a(x, y_\Lambda)\partial_a + \bar{X}^\mu(x, y_\Lambda)\bar{\partial}_\mu \tag{50.3.10}$$

be a generalized vector field on E and $Y = \operatorname{pr} X$ with

$$Y(x, y_\Lambda) = Y^a(x, y_\Lambda)\partial_a + \bar{Y}_\Lambda^\mu(x, y_\Lambda)\bar{\partial}_\mu^\Lambda\,. \tag{50.3.11}$$

From the first condition we find

$$X = \pi_{\infty,0*} \circ Y = Y^a\partial_a + \bar{Y}^\mu\bar{\partial}_\mu\,, \tag{50.3.12}$$

and hence $Y^a = X^a$ and $\bar{Y}^\mu = \bar{X}^\mu$. Applying the second condition to a basic contact one-form θ_Λ^μ we find

$$\begin{aligned}
\mathcal{L}_Y\theta_\Lambda^\mu &= \iota_Y d\theta_\Lambda^\mu + d\iota_Y\theta_\Lambda^\mu\\
&= Y^a dy_{\Lambda a}^\mu - \bar{Y}_{\Lambda a}^\mu dx^a + d(\bar{Y}_\Lambda^\mu - y_{\Lambda a}^\mu Y^a)\\
&= d\bar{Y}_\Lambda^\mu - y_{\Lambda a}^\mu dY^a - \bar{Y}_{\Lambda a}^\mu dx^a\\
&= (\bar{\partial}_\nu^\Omega\bar{Y}_\Lambda^\mu - y_{\Lambda a}^\mu\bar{\partial}_\nu^\Omega Y^a)\theta_\Omega^\nu + (D_a\bar{Y}_\Lambda^\mu - y_{\Lambda b}^\mu D_a Y^b - \bar{Y}_{\Lambda a}^\mu)dx^a\,,
\end{aligned} \tag{50.3.13}$$

where in the last step we used the split $d = d_V + d_H$ of the exterior derivative, together with the explicit coordinate expressions of the vertical and horizontal derivatives. We see that the result is a contact form if and only if the horizontal part vanishes, hence

$$\bar{Y}_{\Lambda a}^\mu = D_a\bar{Y}_\Lambda^\mu - y_{\Lambda b}^\mu D_a Y^b\,. \tag{50.3.14}$$

We have thus obtained a recursive formula for the components of $Y = \operatorname{pr} X$, from which one may guess the explicit formula [Olv86, sec. 5.1]

$$\bar{Y}_\Lambda^\mu = D_\Lambda(\bar{X}^\mu - X^a y_a^\mu) + X^a y_{\Lambda a}^\mu\,. \tag{50.3.15}$$

This formula can be proven by induction. First, for $|\Lambda| = 1$ we have

$$\bar{Y}^{\mu}_{a} = \mathrm{D}_a\bar{X}^{\mu} - y^{\mu}_{\Lambda b}\mathrm{D}_a X^b = \mathrm{D}_a(\bar{X}^{\mu} - X^b y^{\mu}_b) + X^b y^{\mu}_{ab}\,, \tag{50.3.16}$$

which proves the first step. Then, by induction, we show that

$$\bar{Y}^{\mu}_{\Lambda a} = \mathrm{D}_a\left[\mathrm{D}_{\Lambda}(\bar{X}^{\mu} - X^b y^{\mu}_b) + X^b y^{\mu}_{\Lambda b}\right] - y^{\mu}_{\Lambda b}\mathrm{D}_a X^b = \mathrm{D}_{\Lambda a}(\bar{X}^{\mu} - X^b y^{\mu}_b) + X^b y^{\mu}_{\Lambda a b}\,, \tag{50.3.17}$$

which proves all higher orders. Noting that also the case $|\Lambda| = 0$ can be written in the same form

$$\bar{Y}^{\mu} = (\bar{X}^{\mu} - X^a y^{\mu}_a) + X^a y^{\mu}_a = \bar{X}^{\mu}\,, \tag{50.3.18}$$

we may write the prolongation as

$$\mathrm{pr}\,X = X^a\partial_a + \sum_{|\Lambda|=0}^{\infty}\left[\mathrm{D}_{\Lambda}(\bar{X}^{\mu} - X^a y^{\mu}_a) + X^a y^{\mu}_{\Lambda a}\right]\bar{\partial}^{\Lambda}_{\mu}\,, \tag{50.3.19}$$

where we explicitly wrote the sum over Λ to clarify its range.

Another important construction is the following.

Definition 50.3.7 (Total vector field)**.** Let $\pi : E \to M$ be a fiber bundle and $X \in \Gamma(\pi^*_{\infty}TM)$ be a generalized vector field on M. Its *total vector field* is the unique map $\mathrm{tot}\,X : J^{\infty}(E) \to TJ^{\infty}(E)$ which satisfies:

1. it projects to X on M: $X = \pi_{\infty *} \circ \mathrm{tot}\,X$,
2. it annihilates the contact one-forms, i.e., $\iota_{\mathrm{tot}\,X}\theta = 0$ for every contact one-form $\theta \in \Omega^{0,1}(J^{\infty}(E))$.

We can easily derive a coordinate expression for $\mathrm{tot}\,X$. Let $X(x, y_{\Lambda}) = X^a(x, y_{\Lambda})\partial_a$ be a generalized vector field on M and $Y = \mathrm{tot}\,X$ with

$$Y(x, y_{\Lambda}) = Y^a(x, y_{\Lambda})\partial_a + Y^{\mu}_{\Lambda}(x, y_{\Lambda})\bar{\partial}^{\Lambda}_{\mu}\,. \tag{50.3.20}$$

From the first condition we find

$$X = \pi_{\infty *} \circ Y = Y^a\partial_a\,, \tag{50.3.21}$$

and hence $Y^a = X^a$. Further, consider a basic contact one-form θ^{μ}_{Λ} and calculate

$$0 = \iota_Y\theta^{\mu}_{\Lambda} = Y^{\mu}_{\Lambda} - y^{\mu}_{\Lambda a}X^a\,, \tag{50.3.22}$$

so that $Y^{\mu}_{\Lambda} = y^{\mu}_{\Lambda a}X^a$. In summary, we thus have

$$\mathrm{tot}\,X = X^a\left(\partial_a + y^{\mu}_{\Lambda a}\bar{\partial}^{\Lambda}_{\mu}\right) = X^a\mathrm{D}_a\,. \tag{50.3.23}$$

Hence, the total vector field formalizes the total derivative in the direction of X. Observe that the coefficient $X^a y^{\mu}_{\Lambda a}$ of the $|\Lambda|$'th order basis element $\bar{\partial}^{\Lambda}_{\mu}$ contains the coordinate $y^{\mu}_{\Lambda a}$ of order $|\Lambda| + 1$. The total vector field therefore cannot be defined on any finite jet bundle.

50.4 Euler-Lagrange complex

With the preliminaries discussed in the previous sections, we can now formulate the Lagrange theory, which we discussed on finite jet bundles in chapter 49, in terms of the variational bicomplex. For this purpose, we need to adapt some of the notions which we have previously introduced on finite jet bundles, and generalize them to the infinite jet bundle. We start with some basic objects we introduced earlier, and which we now define as follows.

Definition 50.4.1 (Lagrangian)**.** Let $\pi : E \to M$ be a fiber bundle with $\dim M = n$. A *Lagrangian* on E is a horizontal n-form L on the infinite jet bundle $J^\infty(E)$.

Note that in contrast to the original definition 49.1.1, we omitted to specify the order r of the Lagrangian, as it will turn out to be less important if one is working with the variational bicomplex. Nevertheless, also in this formulation a Lagrangian *does* have a finite order. To see this, recall from definition 50.2.1 that we defined differential forms on the infinite jet space $J^\infty(E)$ as equivalence classes of differential forms on finite jet spaces, which are identified via pullback along jet bundle projections. Hence, there exist representatives $\bar{L} \in \Omega^n(J^r(E))$ on finite jet spaces. The lowest r for which such a representative exists may thus be regarded as the order of the Lagrangian.

Also the following definition is now straightforward, in analogy to definition 49.1.2.

Definition 50.4.2 (Action functional)**.** Let $\pi : E \to M$ be a fiber bundle and $L \in \Omega^{n,0}(J^\infty(E))$ a Lagrangian on E. The *action functional* of L over an open domain $U \subset M$ is the function

$$\begin{array}{rccc} S & : & \Gamma_{|U}(E) & \to & \mathbb{R} \\ & & \sigma & \mapsto & \displaystyle\int_U (j^\infty\sigma)^*(L) \end{array} . \tag{50.4.1}$$

It is clear why we define a Lagrangian as a *horizontal* n-form: by pullback along $j^\infty\sigma$, we want to obtain a n-form on M, which we can integrate on U. Further, by virtue of the decomposition shown in theorem 50.2.3, an arbitrary n-form on the infinite jet space uniquely decomposes into horizontal and contact parts, and the latter does not contribute to the action functional, since its pullback along $j^\infty\sigma$ vanishes. Hence, only the horizontal part is relevant for the action functional. Also note that we recover the original definition 49.1.2 if we replace the Lagrangian by a representative $\bar{L}$ on a finite-dimensional jet bundle.

We then come to study extremals of the action. Since the action functional given in definition 50.4.2 assigns a real number to every local section σ, as in the finite-dimensional case, there is no need to alter the definition 49.2.1, since it does not depend on how this assignment $\Gamma_{|U}(E) \to \mathbb{R}$ is defined. However, we will need to make some adaptations when we break the variation of the action into variations of sections and their jet prolongations, as well as variations of the Lagrangian, since these are now considered as objects on an infinite jet space, and so we must check how to generalize the corresponding concepts from the finite-dimensional case. Starting from a family $\tilde{\sigma}_\bullet : \mathbb{R} \to \Gamma_{|U}(E), \epsilon \mapsto \tilde{\sigma}_\epsilon$ with $\tilde{\sigma}_0 = \sigma \in \Gamma_{|U}(E)$ of local sections and $p \in U$ we may construct the infinite sequence

$$\xi_\infty(p) = \frac{\mathrm{d}}{\mathrm{d}\epsilon} j^\infty\tilde{\sigma}_\epsilon(p)\bigg|_{\epsilon=0} = \left(\frac{\mathrm{d}}{\mathrm{d}\epsilon} j^0\tilde{\sigma}_\epsilon(p)\bigg|_{\epsilon=0}, \frac{\mathrm{d}}{\mathrm{d}\epsilon} j^1\tilde{\sigma}_\epsilon(p)\bigg|_{\epsilon=0}, \ldots\right) \in V_{j_p^\infty\sigma} J^\infty(E) . \tag{50.4.2}$$

One easily checks that the members of this sequence are indeed vertical tangent vectors and satisfy the condition $\pi_{r,k*}(v_r) = v_k$ for all $k \leq r$. The assignment $\xi_\infty : U \to VJ^\infty(E), p \mapsto \xi_\infty(p)$, which can be regarded as an infinite jet prolongation of the object ξ introduced in section 49.3, will be used to describe variations of the section σ given by the family $\tilde{\sigma}_\bullet$.

We can now use the prolongation ξ_∞ constructed above to calculate the variation of a differential form on the infinite jet bundle. For this purpose, we must adapt theorem 49.4.1 as follows.

Theorem 50.4.1. *Let $\tilde{\sigma}_\epsilon : M \to E$ be a smooth family of sections of the fiber bundle $\pi : E \to M$ and $\omega \in \Omega^{k,0}(J^\infty)(E))$ a horizontal k-form on the infinite jet bundle $J^\infty(E)$. Then the pullback* *of* ω *along* $j^\infty \tilde{\sigma}_\epsilon$ *satisfies*

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}(j^\infty \tilde{\sigma}_\epsilon)^*(\omega)\bigg|_{\epsilon=0} = (j^\infty \sigma)^* \left(\iota_{\xi_\infty}(\mathrm{d}_V \omega)\right) , \tag{50.4.3}$$

where ξ_∞ *is constructed from the infinite jet prolongation of* ξ *and* $\sigma = \tilde{\sigma}_0$ *via the formula* (50.4.2).

Note that in contrast to the finite-dimensional case 49.4.1, we used the *vertical* derivative on the right hand side of equation (50.4.3). This is justified since we may write $\mathrm{d}\omega = \mathrm{d}_V\omega + \mathrm{d}_H\omega$ and ξ_∞, by construction, is vertical, so that $\iota_{\xi_\infty}(\mathrm{d}_H\omega) = 0$. We have not done this in the finite-dimensional case, since there the vertical derivative d_V also involves a pullback to a different jet bundle. Here it is rather a matter of convenience, as we will illustrate using coordinates.

►...◄

Definition 50.4.3 (Internal Euler operator)**.** Let $\pi : E \to M$ be a fiber bundle with $\dim(M) = n$ and $\Omega^{n,s}(J^\infty(E))$ with $s \geq 1$ the space of forms of type (n, s) on the infinite jet bundle $J^\infty(E)$. The *internal Euler operator* is the unique function $\varrho : \Omega^{n,s}(J^\infty(E)) \to \Omega^{n,s}(J^\infty(E))$ such that:

- ϱ is a projector: $\varrho^2 = \varrho$.
- For $\omega \in \Omega^{n,s}(J^\infty(E))$, the difference $\omega - \varrho(\omega)$ is d_H-exact, i.e., there exists $\eta \in \Omega^{n-1,s}(J^\infty(E))$ such that $\mathrm{d}_H\eta = \omega - \varrho(\omega)$.
- ϱ vanishes on d_H-exact forms: $\varrho \circ \mathrm{d}_H = 0$.
- $\iota_X \circ \varrho = 0$ for all vector fields X on $J^\infty(E)$ with $\pi_{\infty,0*} \circ X = 0$.

Definition 50.4.4 (Euler operator)**.** The *Euler operator* is the function $\mathcal{E} = \varrho \circ \mathrm{d}_V : \Omega^{n,0}(J^\infty(E)) \to \mathcal{F}^1(J^\infty(E))$.

Definition 50.4.5 (Augmented vertical derivative)**.** For $s \geq 1$, the *augmented vertical derivative* is the function $\delta_V = \varrho \circ \mathrm{d}_V : \mathcal{F}^s(J^\infty(E)) \to \mathcal{F}^{s+1}(J^\infty(E))$.

With this definition we can now extend the variational bicomplex introduced in section 50.2 as follows.

Definition 50.4.6 (Augmented variational bicomplex)**.** Let $\pi : E \to M$ be a fiber bundle with $\dim M = n$. The *augmented variational bicomplex* is the structure

$$\begin{array}{ccccccccc}
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
\uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow \\
\Omega^{0,2} & \xrightarrow{\mathrm{d}_H} & \Omega^{1,2} & \dashrightarrow & \Omega^{n-1,2} & \xrightarrow{\mathrm{d}_H} & \Omega^{n,2} & \xrightarrow{\varrho} & \mathcal{F}^2 \\
\uparrow{\scriptstyle \mathrm{d}_V} & & \uparrow{\scriptstyle \mathrm{d}_V} & & \uparrow{\scriptstyle \mathrm{d}_V} & & \uparrow{\scriptstyle \mathrm{d}_V} & & \uparrow{\scriptstyle \delta_V} \\
\Omega^{0,1} & \xrightarrow{\mathrm{d}_H} & \Omega^{1,1} & \dashrightarrow & \Omega^{n-1,1} & \xrightarrow{\mathrm{d}_H} & \Omega^{n,1} & \xrightarrow{\varrho} & \mathcal{F}^1 \\
\uparrow{\scriptstyle \mathrm{d}_V} & & \uparrow{\scriptstyle \mathrm{d}_V} & & \uparrow{\scriptstyle \mathrm{d}_V} & & \uparrow{\scriptstyle \mathrm{d}_V} & \nearrow{\scriptstyle \mathcal{E}} & \\
\Omega^{0,0} & \xrightarrow{\mathrm{d}_H} & \Omega^{1,0} & \dashrightarrow & \Omega^{n-1,0} & \xrightarrow{\mathrm{d}_H} & \Omega^{n,0} & &
\end{array} \tag{50.4.4}$$

where we omitted the part $J^\infty(E)$ in the notation.

Definition 50.4.7 (Euler-Lagrange complex)**.** Let $\pi : E \to M$ be a fiber bundle with $\dim M = n$. The *Euler-Lagrange complex* is the edge

$$\Omega^{0,0} \xrightarrow{\mathrm{d}_H} \cdots \xrightarrow{\mathrm{d}_H} \Omega^{n,0} \xrightarrow{\mathcal{E}} \mathcal{F}^1 \xrightarrow{\delta_V} \cdots \tag{50.4.5}$$

of the augmented variational bicomplex.

Chapter 51

Noether's theorems

51.1 Symmetries of Lagrangian systems

We have previously discussed Lagrangian systems in the language of the variational bicomplex, which is defined on the infinite jet space of a fiber bundle. We will now discuss a particular notion of symmetry of a Lagrangian in this formalism. Since Lagrangians are defined as differential forms on the infinite jet space $J^r(E)$ of a fiber bundle $\pi : E \to M$, we will describe symmetries in terms of (complete) vector fields on $J^r(E)$, whose flow leaves the Euler-Lagrange equations invariant.

However, we cannot consider arbitrary vector fields on $J^r(E)$. To see this, note that the flow of a vector field on $J^r(E)$ is a one-parameter group of diffeomorphisms of $J^r(E)$. Recall that the elements of $J^r(E)$ are jets of sections of $\pi : E \to M$. We are in particular interested in those diffeomorphisms of $J^r(E)$ which are generated by diffeomorphisms of the space $\Gamma(E)$ of sections. In other words, we are looking for diffeomorphisms ϕ whose action on a jet $j^r_x\sigma$ of a section $\sigma \in \Gamma(E)$ at a point $x \in M$ is given by $\phi(j^r_x\sigma) = j^r_x\varphi(\sigma)$ for some diffeomorphism φ of $\Gamma(E)$. This in particular means that ϕ should preserve the subspace $\pi_r^{-1}(x) = J^r_x(E) \subset J^r(E)$ for every $x \in M$, from which follows that the generating vector field must be vertical. We will now construct these vector fields, starting with the following definition.

Definition 51.1.1 (Evolutionary vector field)**.** Let $\pi : E \to M$ be a fiber bundle and $\nu : VE \to E$ the vertical tangent bundle of E. An *evolutionary vector field* is a map $X \in C^\infty(J^r(E), VE)$ such that $\nu \circ X = \pi_{r,0}$.

First of all, note that an evolutionary vector field, despite its name, is not a vector field. In the literature one often finds the term *generalized vector field* for a map taking jets of sections to tangent vectors on E. Formally, it can be written as a "vector field on E with coefficients in $J^r(E)$.

To further understand the meaning of the definition, consider a section $\sigma \in \Gamma(E)$. At a point $x \in M$ this gives us the image $\sigma(x) \in E$ and the r-jet $j^r_x\sigma$, where $\pi_{r,0}(j^r_x\sigma) = \sigma(x)$. An evolutionary vector field X assigns to the jet $j^r_x\sigma$ a vertical tangent vector $X(j^r_x\sigma) \in V_{\sigma(x)}E$. This tangent vector will describe how much the value $\sigma(x)$ of the section σ changes under a certain type of flow.

An evolutionary vector field thus describes how much a section will change at each point. This tells us also how a section as a whole will change under this flow, and thus also how its jets will change. In other words, we can obtain a vertical vector field on $J^r(E)$, which we define as follows.

Definition 51.1.2 (Prolongation)**.** Let $\pi : E \to M$ be a fiber bundle and $X \in C^\infty(J^r(E), VE)$ an evolutionary vector field. Its *prolongation* is the unique vertical vector field $\operatorname{pr} X$ on $J^r(E)$ such that $X = \pi_{r,0*} \circ \operatorname{pr} X$ and $\mathcal{L}_{\operatorname{pr} X}\theta$ is a contact form for every contact form θ.

To clarify this definition, recall that a contact forms is defined as a differential form on $J^r(E)$ whose pullback along the r-jet $j^r\sigma$ of any section $\sigma \in \Gamma(E)$ vanishes. Since we wish that the flow of $\operatorname{pr} X$ maps the r-jets of sections again to r-jets of sections, it also maps contact forms to contact forms. However, it is easier to work with contact forms, which is why we used them in the definition above.

As a further illustration, we write the prolongation in terms of coordinates. Let (x^α) be coordinates on M and (x^α, y^a) coordinates on E corresponding to a local trivialization. In these coordinates an evolutionary vector field X can be written in the form $X = X^a\bar{\partial}_a$, where the coefficients X^a depend on the jet coordinates (x^α, y^a_Λ). One can show that the prolongation of X is then given by

$$\operatorname{pr} X = \sum_\Lambda \mathrm{D}_\Lambda X^a \bar{\partial}^\Lambda_a \,. \tag{51.1.1}$$

Here D_Λ denotes the total derivative (49.5.7). The reason for this formula is intuitively clear: if the flow of X describes the transformation of a section σ, then we need to take all derivatives of X to see how the flow of $\operatorname{pr} X$ transforms the jet of a section.

The prolongations of evolutionary vector fields have a few nice properties, which we summarize here.

Theorem 51.1.1. *The prolongation* $\operatorname{pr} X$ *of an evolutionary vector field* X *satisfies:*

$$\begin{aligned}
\mathcal{L}_{\operatorname{pr} X} &= \iota_{\operatorname{pr} X}\mathrm{d}_V + \mathrm{d}_V\iota_{\operatorname{pr} X} \,, && \text{(51.1.2a)}\\
0 &= \iota_{\operatorname{pr} X}\mathrm{d}_H + \mathrm{d}_H\iota_{\operatorname{pr} X} \,, && \text{(51.1.2b)}\\
0 &= \mathcal{L}_{\operatorname{pr} X}\mathrm{d}_H - \mathrm{d}_H\mathcal{L}_{\operatorname{pr} X} \,, && \text{(51.1.2c)}\\
0 &= \mathcal{L}_{\operatorname{pr} X}\mathrm{d}_V - \mathrm{d}_V\mathcal{L}_{\operatorname{pr} X} \,. && \text{(51.1.2d)}
\end{aligned}$$

We will not prove the first property here. Note that the second statement follows from the first one, since

$$\begin{aligned}
0 &= \mathcal{L}_{\operatorname{pr} X} - (\iota_{\operatorname{pr} X}\mathrm{d}_V + \mathrm{d}_V\iota_{\operatorname{pr} X})\\
&= \iota_{\operatorname{pr} X}\mathrm{d} + \mathrm{d}\iota_{\operatorname{pr} X} - \iota_{\operatorname{pr} X}\mathrm{d}_V - \mathrm{d}_V\iota_{\operatorname{pr} X}\\
&= \iota_{\operatorname{pr} X}\mathrm{d}_H + \mathrm{d}_H\iota_{\operatorname{pr} X} \,.
\end{aligned} \tag{51.1.3}$$

Now the third statement follows from the second one and the commutation relations of the horizontal and vertical derivative, since

$$\begin{aligned}
\mathcal{L}_{\operatorname{pr} X}\mathrm{d}_H &= (\mathrm{d}_V\iota_{\operatorname{pr} X} + \iota_{\operatorname{pr} X}\mathrm{d}_V)\mathrm{d}_H\\
&= -\mathrm{d}_V\mathrm{d}_H\iota_{\operatorname{pr} X} - \iota_{\operatorname{pr} X}\mathrm{d}_H\mathrm{d}_V\\
&= \mathrm{d}_H(\mathrm{d}_V\iota_{\operatorname{pr} X} + \iota_{\operatorname{pr} X}\mathrm{d}_V)\\
&= \mathrm{d}_H\mathcal{L}_{\operatorname{pr} X} \,.
\end{aligned} \tag{51.1.4}$$

The last statement finally follows from the third one and the fact that Lie derivative and exterior derivative commute.

Now we have found the class of vector fields on $J^r(E)$ which correspond to transformations of the space $\Gamma(E)$ of sections. We can now restrict ourselves to those vector fields from this class which leave the dynamics of the Lagrangian system, given by the Euler-Lagrange equations, invariant. We define them as follows.

Definition 51.1.3 (Symmetry)**.** Let $\pi : E \to M$ be a fiber bundle with $\dim M = n$ and $L \in \Omega^{n,0}(J^r(E))$ a Lagrangian. A *symmetry* of L is an evolutionary vector field X such that $\mathcal{L}_{\mathrm{pr}\,X} L \in \Omega^{n,0}(J^r(E))$ is d_H-exact.

The definition above states that the flow of $\mathrm{pr}\,X$ changes the Lagrangian L only by a d_H-exact form. This means that the pullback $(j^r\sigma)^*(L)$ of L to M along the r-jet of a section changes only by an exact form on M. This in turn means that the action functional

$$S[\sigma] = \int_M (j^r\sigma)^*(L) \tag{51.1.5}$$

is invariant. It also follows that $\mathcal{L}_{\mathrm{pr}\,X}\mathcal{E}L = 0$, i.e., the Euler-Lagrange equations are invariant. Also it is helpful to note that since $\mathrm{pr}\,X$ is vertical by construction, while L is horizontal by definition, we have

$$\mathcal{L}_{\mathrm{pr}\,X} L = \underbrace{\mathrm{d}_V \iota_{\mathrm{pr}\,X} L}_{=0} + \iota_{\mathrm{pr}\,X}\mathrm{d}_V L = \iota_{\mathrm{pr}\,X}\mathrm{d}_V L\,. \tag{51.1.6}$$

51.2 Conserved currents

The task of finding the solutions to the Euler-Lagrange equations can often be simplified if the Lagrangian system contains something known as a *conserved current* in field theory, or a *constant of motion* in mechanics. Here we will use the term *conserved current* and the following definition.

Definition 51.2.1 (Conserved current)**.** Let $\pi : E \to M$ be a fiber bundle with $\dim M = n$ and $L \in \Omega^{n,0}(J^r(E))$ a Lagrangian. A *conserved current* of L is an element $\psi \in \Omega^{n-1,0}(J^r(E))$ such that $\mathrm{d}_H\psi = 0$ on the subspace of $J^r(E)$ where $\mathcal{E}L = 0$.

In order to understand the meaning of this, let $\sigma \in \Gamma(E)$ be a solution of the Euler-Lagrange equations, i.e., $\mathcal{E}L \circ j^r\sigma = 0$. Then $(j^r\sigma)^*(\psi)$ is a $n-1$-form on M, where $n = \dim M$, with

$$\mathrm{d}(j^r\sigma)^*(\psi) = (j^r\sigma)^*(\mathrm{d}_H\psi) = 0\,. \tag{51.2.1}$$

In other words, for each solution σ, the pullback $(j^r\sigma)^*(\psi)$ is closed. This resembles the standard notion of a conserved current.

51.3 Noether's first theorem

With the preliminary definitions made in the previous sections we can now come to the central topic of this lecture, which is Noether's theorem. In the formalism we use here, it is formulated as follows.

Theorem 51.3.1 (Noether's first theorem)**.** *Let X be a symmetry of a Lagrangian $L \in \Omega^{n,0}(J^r(E))$ on a fiber bundle $\pi : E \to M$ with $\dim M = n$. Then*

$$\psi = \iota_{\mathrm{pr}\,X}\eta - \omega \in \Omega^{n-1,0}(J^r(E))\,, \tag{51.3.1}$$

where $\mathrm{d}_H\omega = \iota_{\mathrm{pr}\,X}\mathrm{d}_V L$ and $\mathrm{d}_H\eta = \mathcal{E}L - \mathrm{d}_V L$, is a conserved current.

Proof. By definition of a symmetry, $\iota_{\mathrm{pr}\,X}\mathrm{d}_V L$ is d_H-exact, i.e., there exists $\omega \in \Omega^{n-1,0}(J^r(E))$ such that $\mathrm{d}_H\omega = \iota_{\mathrm{pr}\,X}\mathrm{d}_V L$. Further, by the definition of the internal Euler operator ϱ, the difference $\mathcal{E}L - \mathrm{d}_V L = \varrho(\mathrm{d}_V L) - \mathrm{d}_V L$ is also d_H-exact, i.e., there exists $\eta \in \Omega^{n-1,1}(J^r(E))$ such that $\mathrm{d}_H\eta = \mathcal{E}L - \mathrm{d}_V L$.

Using the fact that $\mathrm{pr}\,X$ is the prolongation of an evolutionary vector field, we can now evaluate $\mathrm{d}_H\psi$ and find

$$\begin{aligned} \mathrm{d}_H\psi &= \mathrm{d}_H\iota_{\mathrm{pr}\,X}\eta - \mathrm{d}_H\omega \\ &= -\iota_{\mathrm{pr}\,X}\mathrm{d}_H\eta - \iota_{\mathrm{pr}\,X}\mathrm{d}_V L \\ &= -\iota_{\mathrm{pr}\,X}\mathcal{E}L\,, \end{aligned} \tag{51.3.2}$$

using theorem 51.1.1. This obviously vanishes where $\mathcal{E}L = 0$, so that ψ is a conserved current. ■

The theorem is as elegant and as simple as a theorem could be. We can make the current even more explicit if we have a Lepage equivalent of the Lagrangian. Then we have the following property.

Theorem 51.3.2 (Noether's first theorem)**.** *Let X be a symmetry of a Lagrangian $L \in \Omega^{n,0}(J^r(E))$ on a fiber bundle $\pi : E \to M$ with $\dim M = n$ and $\rho \in \Omega^n(J^s(E))$ a Lepage equivalent of L. Then*

$$\psi = \mathrm{p}_0\iota_{\mathrm{pr}\,X}\rho - \omega \in \Omega^{n-1,0}(J^{s+1}(E))\,, \tag{51.3.3}$$

where $\mathrm{d}_H\omega = \iota_{\mathrm{pr}\,X}\mathrm{d}_V L$, is a conserved current (up to pullback along jet bundle projections).

Proof. By definition of a symmetry of the Lagrangian, $\iota_{\mathrm{pr}\,X}\mathrm{d}_V L$ is d_H-exact, and so the desired form ω exists. Then, by direct calculation, we find

$$\begin{aligned} \mathrm{d}_H\psi &= \mathrm{d}_H\mathrm{p}_0\iota_{\mathrm{pr}\,X}\rho - \mathrm{d}_H\omega \\ &= \mathrm{p}_0\mathrm{d}\iota_{\mathrm{pr}\,X}\rho - \iota_{\mathrm{pr}\,X}\mathrm{d}_V L \\ &= \mathrm{p}_0(\mathcal{L}_{\mathrm{pr}\,X}\rho - \iota_{\mathrm{pr}\,X}\mathrm{d}\rho) - \iota_{\mathrm{pr}\,X}\mathrm{d}_V L \\ &= \iota_{\mathrm{pr}\,X}\mathrm{d}_V\mathrm{p}_0\rho - \iota_{\mathrm{pr}\,X}\mathrm{d}_V L - \mathrm{p}_0\iota_{\mathrm{pr}\,X}\mathrm{d}\rho \\ &= \iota_{\mathrm{pr}\,X}\mathrm{d}_V L - \iota_{\mathrm{pr}\,X}\mathrm{d}_V L - \iota_{\mathrm{pr}\,X}\mathrm{p}_1\mathrm{d}\rho \\ &= -\iota_{\mathrm{pr}\,X}\mathcal{E}L\,, \end{aligned} \tag{51.3.4}$$

where we used the following relations:

- The horizontal component of $\mathrm{d}\iota_{\mathrm{pr}\,X}\rho$ is given by

$$\mathrm{p}_0\mathrm{d}\iota_{\mathrm{pr}\,X}\rho = \mathrm{d}_H\mathrm{p}_0\iota_{\mathrm{pr}\,X}\rho\,, \tag{51.3.5}$$

  since the only possibility to obtain a horizontal component is by taking the horizontal derivative of an already horizontal part, without obtaining contact components.

- The Lie derivative can be split by Cartan's magic formula:

$$\mathcal{L}_{\mathrm{pr}\,X}\rho = \mathrm{d}\iota_{\mathrm{pr}\,X}\rho + \iota_{\mathrm{pr}\,X}\mathrm{d}\rho\,. \tag{51.3.6}$$

- The Lagrangian is the horizontal part of the Lepage form: $L = \mathrm{p}_0\rho$.

- The Euler-Lagrange equations are given by $\mathcal{E}L = \mathrm{p}_1\mathrm{d}\rho$.

- The horizontal projector p_0 and Lie derivative $\mathcal{L}_{\mathrm{pr}\,X}$ with respect to a vertical vector field of ρ yield

$$\mathrm{p}_0\mathcal{L}_{\mathrm{pr}\,X}\rho = \underbrace{\mathrm{p}_0\mathrm{d}_V\iota_{\mathrm{pr}\,X}\rho}_{=0} + \mathrm{p}_0\iota_{\mathrm{pr}\,X}\mathrm{d}_V\rho = \iota_{\mathrm{pr}\,X}\mathrm{p}_1\mathrm{d}_V\rho = \iota_{\mathrm{pr}\,X}\mathrm{d}_V\mathrm{p}_0\rho = \iota_{\mathrm{pr}\,X}\mathrm{d}_V L\,. \tag{51.3.7}$$

The result vanishes when the Euler-Lagrange equations are imposed. ■

We will now apply the theorem to a few examples.

Example 51.3.1 (Momentum conservation). Let $M = \mathbb{R}$ and $E = \mathbb{R} \times Q$ with a manifold Q, so that the bundle $\pi : E \to M$ is a trivial bundle with $\pi = \mathrm{pr}_{\mathbb{R}}$ being the projection onto the first factor. Using the coordinate t on $\mathbb{R}$ and coordinates (q^a) on Q, we have coordinates (t, q^a) on E and $(t, q^a = q^a_0, \dot{q}^a = q^a_1, \ddot{q}^a = q^a_2, \ldots)$ on $J^r(E)$. This systems can be used to model, for example, the motion of a point mass on a manifold Q, with t measuring time and q^a the position of the point mass.

We now consider a Lagrangian of the form $L = \mathcal{L}(\dot{q})\mathrm{d}t \in \Omega^{1,0}(J^r(E))$, where $\mathcal{L}$ depends only on the velocity $\dot{q}^a$, but not on the position q^a. Taking the vertical derivative we obtain

$$\mathrm{d}_V L = \frac{\partial \mathcal{L}}{\partial \dot{q}^a} \dot{\theta}^a \wedge \mathrm{d}t \,. \tag{51.3.8}$$

Further applying the internal Euler operator ϱ we obtain

$$\mathcal{E}L = \varrho \mathrm{d}_V L = -\mathrm{D}_t \frac{\partial \mathcal{L}}{\partial \dot{q}^a} \theta^a \wedge \mathrm{d}t = -\frac{\partial^2 \mathcal{L}}{\partial \dot{q}^a \partial \dot{q}^b} \ddot{q}^b \theta^a \wedge \mathrm{d}t \,, \tag{51.3.9}$$

where D_t is the total time derivative. The second derivative of $\mathcal{L}$ appearing here is also called the *Lagrange metric*, and is usually assumed to be non-degenerate, so that the Euler-Lagrange equations imply $\ddot{q}^a = 0$. From the expressions above it is easy to see that

$$\mathcal{E}L - \mathrm{d}_V L = -\left(\mathrm{D}_t \frac{\partial \mathcal{L}}{\partial \dot{q}^a} \theta^a + \frac{\partial \mathcal{L}}{\partial \dot{q}^a} \dot{\theta}^a\right) \wedge \mathrm{d}t = \mathrm{d}_H \left(\frac{\partial \mathcal{L}}{\partial \dot{q}^a} \theta^a\right) = \mathrm{d}_H \eta \tag{51.3.10}$$

is indeed d_H-exact, by the definition of the internal Euler operator ϱ.

We now consider the evolutionary vector field $X = \xi^a \bar{\partial}_a$ on $J^r(E)$ with constant ξ^a. Its prolongation is simply the vector field itself, $\mathrm{pr}\, X = X$. One easily checks that it is a symmetry of the Lagrangian, since

$$\iota_{\mathrm{pr}\, X} \mathrm{d}_V L = \iota_{\xi^a \bar{\partial}_a} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}^a} \dot{\theta}^a \wedge \mathrm{d}t\right) = 0 \,. \tag{51.3.11}$$

This ultimately follows from the fact that $\mathcal{L}$ does not depend on the position q^a, so that $\mathrm{d}_V L$ does not contain the contact form θ^a, which would give a non-vanishing contribution with $\bar{\partial}_a$. We thus simply have $\omega = 0$. This yields us the conserved current

$$\psi = \iota_{\mathrm{pr}\, X} \eta - \omega = \xi^a \frac{\partial \mathcal{L}}{\partial \dot{q}^a} = \xi^a p_a \,. \tag{51.3.12}$$

The components p_a defined above are called *canonical momenta*. One can see that this is indeed a conserved current, since

$$\mathrm{d}_H \psi = \xi^a \mathrm{D}_t p_a \mathrm{d}t = \xi^a \frac{\partial^2 \mathcal{L}}{\partial \dot{q}^a \partial \dot{q}^b} \ddot{q}^b \mathrm{d}t \tag{51.3.13}$$

vanishes on solutions of the Euler-Lagrange equations.

Example 51.3.2 (Total energy conservation). We consider the same fiber bundle as in the previous example, but allow the Lagrangian $L = \mathcal{L}(q, \dot{q})\mathrm{d}t \in \Omega^{1,0}(J^r(E))$ to depend also on the position q^a. The vertical derivative is then given by

$$\mathrm{d}_V L = \left(\frac{\partial \mathcal{L}}{\partial q^a} \theta^a + \frac{\partial \mathcal{L}}{\partial \dot{q}^a} \dot{\theta}^a\right) \wedge \mathrm{d}t \,. \tag{51.3.14}$$

Application of the internal Euler operator then yields

$$\mathcal{E}L = \varrho \mathrm{d}_V L = \left(\frac{\partial \mathcal{L}}{\partial q^a} - \mathrm{D}_t \frac{\partial \mathcal{L}}{\partial \dot{q}^a} \right) \theta^a \wedge \mathrm{d}t = \left(\frac{\partial \mathcal{L}}{\partial q^a} - \frac{\partial^2 \mathcal{L}}{\partial \dot{q}^a \partial q^b} \dot{q}^b - \frac{\partial^2 \mathcal{L}}{\partial \dot{q}^a \partial \dot{q}^b} \ddot{q}^b \right) \theta^a \wedge \mathrm{d}t \,. \tag{51.3.15}$$

From this we read off that

$$\mathcal{E}L - \mathrm{d}_V L = -\left(\mathrm{D}_t \frac{\partial \mathcal{L}}{\partial \dot{q}^a} \theta^a + \frac{\partial \mathcal{L}}{\partial \dot{q}^a} \dot{\theta}^a \right) \wedge \mathrm{d}t = \mathrm{d}_H \left(\frac{\partial \mathcal{L}}{\partial \dot{q}^a} \theta^a \right) = \mathrm{d}_H \eta \,, \tag{51.3.16}$$

which actually yields the same expression for η as in the previous example.

We now consider the evolutionary vector field $X = \dot{q}^a \bar{\partial}_a$, whose prolongation is given by

$$\mathrm{pr}\, X = \sum_{\lambda=0}^{\infty} q^a_{\lambda+1} \bar{\partial}^\lambda_a \,. \tag{51.3.17}$$

This is a symmetry of the Lagrangian, since

$$\iota_{\mathrm{pr}\, X} \mathrm{d}_V L = \left(\frac{\partial \mathcal{L}}{\partial q^a} \dot{q}^a + \frac{\partial \mathcal{L}}{\partial \dot{q}^a} \ddot{q}^a \right) \mathrm{d}t = \mathrm{D}_t \mathcal{L} \mathrm{d}t = \mathrm{d}_H \mathcal{L} = \mathrm{d}_H \omega \tag{51.3.18}$$

is d_H-exact with $\omega = \mathcal{L}$. This gives us the conserved current

$$\psi = \iota_{\mathrm{pr}\, X} \eta - \omega = \frac{\partial \mathcal{L}}{\partial \dot{q}^a} \dot{q}^a - \mathcal{L} = p_a \dot{q}^a - \mathcal{L} = \mathcal{H} \,, \tag{51.3.19}$$

which is called the *Hamiltonian* and describes the total energy of the system. This is a conserved current, since

$$\begin{aligned} \mathrm{d}_H \psi = \mathrm{D}_t \mathcal{H} \mathrm{d}t &= \left(\frac{\partial^2 \mathcal{L}}{\partial \dot{q}^a \partial q^b} \dot{q}^b + \frac{\partial^2 \mathcal{L}}{\partial \dot{q}^a \partial \dot{q}^b} \ddot{q}^b \right) \dot{q}^a \mathrm{d}t + \frac{\partial \mathcal{L}}{\partial \dot{q}^a} \ddot{q}^a \mathrm{d}t - \frac{\partial \mathcal{L}}{\partial q^a} \dot{q}^a \mathrm{d}t - \frac{\partial \mathcal{L}}{\partial \dot{q}^a} \ddot{q}^a \mathrm{d}t \\ &= \left(\frac{\partial^2 \mathcal{L}}{\partial \dot{q}^a \partial q^b} \dot{q}^b + \frac{\partial^2 \mathcal{L}}{\partial \dot{q}^a \partial \dot{q}^b} \ddot{q}^b - \frac{\partial \mathcal{L}}{\partial q^a} \right) \dot{q}^a \mathrm{d}t \,, \end{aligned} \tag{51.3.20}$$

which vanishes when the Euler-Lagrange equations are imposed.

51.4 Noether's second theorem

Chapter 52

Gauge theory

52.1 Finite gauge transformations

In physics we often encounter theories which are invariant under "local symmetries", i.e., "local actions" of a group. This is the basic idea of gauge theory: one considers a Lagrangian system which is invariant under the action of a group, where the group element which acts on a fiber of a bundle depends on the base point of the fiber, hence the action is "local". Many examples come from field theory, such as electromagnetism, which is locally invariant under U(1), or the strong interaction between quarks, which is locally invariant under SU(3). We will now discuss how these theories can be described using the geometrical notions we introduced in the previous lectures. In particular, we make use of principal bundles, which are the most suitable setting for gauge theories. In this section we start by introducing a few important notions in the way they are used in the context of gauge theories. We start by providing a formal definition of such "point dependent" or "local" group action.

Definition 52.1.1 (Finite gauge transformation)**.** Let $\pi : P \to M$ be a principal G-bundle with Lie group G. A *(finite) gauge transformation* is a vertical principal bundle automorphism of P, i.e., a diffeomorphism $\varphi : P \to P$ such that $\pi\circ\varphi = \pi$ and $R_g\circ\varphi = \varphi\circ R_g$ for all $g \in G$. The gauge transformations form a group with respect to function composition $\circ$, which is called the *gauge group* and denoted $\mathcal{G}$.

A few remarks are in order. First, note that φ must preserve the fibers - it maps each fiber to itself. Further, it must commute with the right translation R_g by an element $g \in G$ of the structure group G. Note that if we would choose $\varphi = R_h$ for $h \in G$, we in general do *not* obtain an element of the gauge group, $\varphi \notin \mathcal{G}$, since only the first property $\pi \circ \varphi = \pi$ is satisfied, but in general

$$R_g \circ \varphi = R_g \circ R_h = R_{hg} \neq R_{gh} = R_h \circ R_g = \varphi \circ R_g \,, \tag{52.1.1}$$

unless the structure group G is abelian.

Note also that while the structure group G acts from the right on P, the gauge group acts *from the left* by defining $\varphi \cdot p = \varphi(p)$ for $(\varphi, p) \in \mathcal{G} \times P$. This is obviously a left action, since

$$(\varphi \circ \varphi') \cdot p = \varphi(\varphi'(p)) = \varphi \cdot (\varphi' \cdot p) \,. \tag{52.1.2}$$

Also it is clear that the left action by $\mathcal{G}$ and the right action by G commute, $(\varphi\cdot p)\cdot g = \varphi\cdot(p\cdot g)$.

Since the action of the gauge group commutes with the action of the structure group, one may expect that it preserves also other structures defined by the latter. This is indeed the case, as the following statement shows.

Theorem 52.1.1. *The fundamental vector fields $\tilde{X}$ for $X \in \mathfrak{g}$ are invariant under gauge transformations, $\tilde{X} \circ \varphi = \varphi_* \circ \tilde{X}$.*

Proof. First, from the fact that φ commutes with the right translation that

$$R^{\varphi(p)}(g) = R_g(\varphi(p)) = \varphi(R_g(p)) = \varphi(R^p(g)) = (\varphi \circ R^p)(g) \tag{52.1.3}$$

for all $p \in P$ and $g \in G$. This implies $R^{\varphi(p)} = \varphi \circ R^p$, so that we further find

$$(\tilde{X} \circ \varphi)(p) = (R^{\varphi(p)})_*(X(e)) = [\varphi_* \circ (R^p)_*](X(e)) = (\varphi_* \circ \tilde{X})(p)\,, \tag{52.1.4}$$

so that indeed $\tilde{X} \circ \varphi = \varphi_* \circ \tilde{X}$. ■

Defining the gauge group as a group of functions makes it rather cumbersome to list all elements of the group and to understand its geometric structure, and so we aim for a more practical and geometric description. A first step towards this description is the following.

Theorem 52.1.2. *There is a one-to-one correspondence between elements $\varphi \in \mathcal{G}$ of the gauge group and equivariant maps $\hat{\varphi} \in C^\infty_G(P, G)$, where G acts on itself from the right by conjugation, and the correspondence is established by the relation*

$$\varphi(p) = R_{\hat{\varphi}(p)}(p) = p \cdot \hat{\varphi}(p)\,. \tag{52.1.5}$$

Proof. We start with an element $\varphi \in \mathcal{G}$. By definition, φ preserves the fibers, and so for every $p \in P$ we have $\pi(\varphi(p)) = \pi(p)$. Since the structure group G acts freely and transitively on the fibers of P, there exists a unique group element $\hat{\varphi}(p) \in G$ such that $\varphi(p) = p \cdot \hat{\varphi}(p)$. Hence, we obtain a map $\hat{\varphi} : P \to G$. To see that this map is equivariant, we calculate

$$(p \cdot g) \cdot \hat{\varphi}(p \cdot g) = \varphi(p \cdot g) = \varphi(p) \cdot g = p \cdot (\hat{\varphi}(p)g) = (p \cdot g) \cdot (g^{-1}\hat{\varphi}(p)g)\,, \tag{52.1.6}$$

from which follows $\hat{\varphi}(p \cdot g) = g^{-1}\hat{\varphi}(p)g$. Since the right hand side of this equation is simply the action of G on itself from the right by conjugation, it follows that $\hat{\varphi}$ is indeed an equivariant map.

Conversely, we start with an equivariant map $\hat{\varphi} \in C^\infty_G(P, G)$ and use the relation (52.1.5) now in the opposite direction in order to define a map $\varphi : P \to P$. Since right translation by an element $\hat{\varphi}(p) \in G$ preserves the fibers, the same holds also for φ. To show that φ commutes with the right translation on P, one now calculates

$$\varphi(p) \cdot g = p \cdot (\hat{\varphi}(p)g) = (p \cdot g) \cdot (g^{-1}\hat{\varphi}(p)g) = (p \cdot g) \cdot \hat{\varphi}(p \cdot g) = \varphi(p \cdot g)\,, \tag{52.1.7}$$

this time using the equivariance of $\hat{\varphi}$ for proving the opposite direction. This shows that $\varphi \in \mathcal{G}$.

Finally, it is easy to check that these two constructions are inverses of each other. We omit this part of the proof here. ■

The correspondence $\mathcal{G} \cong C^\infty_G(P, G)$ also yields us a group structure on $C^\infty_G(P, G)$ as follows. Consider $\varphi, \varphi' \in \mathcal{G}$ and apply their composition

$$\begin{aligned}(\varphi \circ \varphi')(p) &= \varphi(\varphi'(p))\\ &= \varphi(p \cdot \hat{\varphi}'(p))\\ &= \varphi(p) \cdot \hat{\varphi}'(p)\\ &= (p \cdot \hat{\varphi}(p)) \cdot \hat{\varphi}'(p)\\ &= p \cdot (\hat{\varphi}(p)\hat{\varphi}'(p))\,,\end{aligned} \tag{52.1.8}$$

so that $\widehat{\varphi \circ \varphi'}(p) = \hat{\varphi}(p)\hat{\varphi}'(p)$. Hence, the induced group multiplication on $C^\infty_G(P, G)$ is given by pointwise multiplication, and the unit element is given by the map $p \mapsto e \in G$ for all $p \in P$.

The construction detailed already gives us more insight into the structure of the gauge group, and allows us to express its elements in an easier way. However, it is still rather cumbersome to work with functions on P and have an equivariance condition. To further simplify this notion, and arrive at the structure we aim for, we continue with the following definition.

Definition 52.1.2 (Group bundle)**.** Let $\pi : P \to M$ be a principal G-bundle with Lie group G and denote by $\alpha : G \times G \to G$ the left action $\alpha(g,h) = ghg^{-1}$ of G on itself by conjugation. The associated bundle $\pi_\alpha : P^G = P \times_\alpha G \to M$ is called the *group bundle* of P.

Note that although the fibers of both the principal bundle P and the group bundle P^G are diffeomorphic to G and both carry actions of G, there are a few crucial differences between these two bundles. First, note that the action of G on P^G is in general *not* transitive nor free, and so P^G is *not* a principal bundle. (In the case of an abelian group G, the action is even trivial.) Also the fibers of P^G are equipped with a group structure, which is *not* the case for a principal bundle P. This group structure can be seen as follows. For $p \in P$ and $g, g' \in G$, so that $[p,g], [p,g'] \in P^G$ lie in the same fiber, define

$$[p,g][p,g'] = [p,gg'] \,. \tag{52.1.9}$$

To see that this is well-defined, consider a representative belonging to a different $p' = p \cdot h$. By the definition of an associated bundle we then have

$$\begin{aligned}
[p',g][p',g'] &= [p\cdot h,g][p\cdot h,g'] \\
&= [p,\alpha(h,g)][p,\alpha(h,g')] \\
&= [p,hgh^{-1}][p,hg'h^{-1}] \\
&= [p,hgg'h^{-1}] \\
&= [p,\alpha(h,gg')] \\
&= [p\cdot h,gg'] \\
&= [p',gg'] \,,
\end{aligned} \tag{52.1.10}$$

which shows that the definition is independent of the choice of the representative. Further, we see that for each $p \in P$ the map $[p,\bullet] : G \to P^G, g \mapsto [p,g]$ is even a Lie group isomorphism; the unit element in each fiber is given by $[p,e]$. This finally brings us to the description of the gauge group we want to use. This is established by the following theorem.

Theorem 52.1.3. *There is a one-to-one correspondence between equivariant maps $\hat{\varphi} \in C^\infty_G(P,G)$, where G acts on itself from the right by conjugation, and sections $\tilde{\varphi} \in \Gamma(P^G)$ of the group bundle $\pi_\alpha : P^G \to M$.*

Proof. This is essentially a consequence of theorem 20.3.3. The only technicality we have to take into account is that P^G is, as usual, defined via a *left* action, while for the equivariant maps we considered a *right* action. However, this is simply solved by writing this right action as $\bar{\alpha}(g,h) = h^{-1}gh = \alpha(h^{-1},g)$. Now a map $\hat{\varphi} : P \to G$ is equivariant with respect to α if and only if it is equivariant with respect to $\bar{\alpha}$, since

$$\bar{\alpha}(\hat{\varphi}(p),g) = \alpha(g^{-1},\hat{\varphi}(p)) \,, \tag{52.1.11}$$

and hence $\hat{\varphi}(p \circ g) = \bar{\alpha}(\hat{\varphi}(p),g)$ if and only if $\hat{\varphi}(p \circ g) = \alpha(g^{-1},\hat{\varphi}(p))$, taking into account that we have the to take the inverse group element if we consider equivariance between right and left actions, following definition 15.5.1. Thus, we can apply theorem 20.3.3, which completes the proof and establishes the desired correspondence. ■

The description of the gauge group $\mathcal{G} \cong \Gamma(P^G)$ in terms of sections of the bundle $\pi_\alpha : P^G \to M$ is the one we will be using in this chapter. Using the fact that each fiber of P^G carries a group structure, it is easy to see that the group structure of the gauge group is simply given by pointwise multiplication $(\tilde{\varphi}\tilde{\varphi}')(x) = \tilde{\varphi}(x)\tilde{\varphi}'(x)$, and that the unit element is given by the section of P^G which assigns to each $x \in M$ the unit element in the fiber $P^G_x = \pi_\alpha^{-1}(x)$.

A section of P^G is now essentially what we described in the beginning of this section: it assigns to every $x \in M$ an element of a group P^G_x *isomorphic* to the structure group G. However, note that this isomorphism $G \to P^G_x$ is not canonical, but depends on the choice of an element $p \in P_x = \pi^{-1}(x)$. If we wanted to express a gauge transformation by a map $M \to G$, we would have to pick such an element p in each fiber - and hence specify a section of P. We cannot always do so globally, unless P is trivial, but we can do so locally by means of the following definition.

Definition 52.1.3 (Gauge). Let $\pi : P \to M$ be a principal G-bundle with Lie group G. A *gauge* on an open subset $U \subset M$ is a local section $\epsilon : U \to P$.

By choosing a gauge $\epsilon : U \to P$ on $U \subset M$ we can locally express a gauge transformation φ by a map $\varphi^\epsilon : U \to G$, which is defined such that

$$\tilde{\varphi}(x) = [\epsilon(x), \varphi^\epsilon(x)] \in P^G \tag{52.1.12}$$

for all $x \in U$, using the group bundle section $\tilde{\varphi} : M \to P^G$ formulation of the gauge transformation.

Recall that for a principal bundle a local section is equivalent to a local trivialization of the bundle, and so we can also use it to construct induced bundle coordinates. It is instructive to derive the corresponding coordinate expressions for a simple example, such as the structure group G being a matrix group.

Example 52.1.1 (Gauge transformations on principal matrix group bundles). Let $G \subset M_{n,n}$ be a matrix group and $\pi : P \to M$ a principal G-bundle. We will use the matrix components $(g^a{}_b)$ as coordinates on G (imposing suitable restrictions on them, in order to represent only those matrices that lie in G). In order to construct coordinates on P, we pick a gauge $\epsilon : U \to P$ on an open subset $U \subset M$. This induces a local trivialization $\phi : \pi^{-1}(U) \to U \times G$ such that $\phi(\epsilon(x) \cdot g) = (x, g)$ for all $(x, g) \in U \times G$. Introducing coordinates (x^μ) on U, we can write the coordinates on $\pi^{-1}(U)$ which are induced by the trivialization ϕ and the coordinates $(g^a{}_b)$ on G in the form $(x^\mu, p^a{}_b)$. These are the coordinates we will use. Note that in these coordinates the images $\epsilon(x)$ of the section ϵ are represented by (x^μ, δ^a_b).

The gauge ϵ also induces a local trivialization $\phi_\alpha : \pi_\alpha^{-1}(U) \to U \times G$ of the group bundle P^G, which is defined such that $\phi_\alpha([\epsilon(x), g]) = (x, g)$ for all $(x, g) \in U \times G$. We denote the coordinates induced by ϕ_α on $\pi_\alpha^{-1}(U)$ by $(x^\mu, q^a{}_b)$. Note that the group structure on the fibers of P^G is represented by $(q\tilde{q})^a{}_b = q^a{}_c\tilde{q}^c{}_b$, i.e., by the matrix multiplication inherited from G. This also implies that the unit elements in each fiber are represented by the coordinates (x^μ, δ^a_b).

Now consider a (global) section $\tilde{\varphi} : M \to P^G$ of the group bundle. Our aim is to derive the coordinate expression for the gauge transformation $\varphi : P \to P$ from the coordinate expression for $\tilde{\varphi}$, relative to the gauge ϵ. This will be done in several steps. First note that $\tilde{\varphi}$ can be expressed in our chosen coordinates such that it assigns to $(x^\mu) \in U$ an element with coordinates $(x^\mu, \tilde{\varphi}^a{}_b(x)) \in P^G$. These coordinates are easily understood by applying the trivialization ϕ_α, which yields

$$\phi_\alpha(\tilde{\varphi}(x)) = \phi_\alpha([\epsilon(x), \varphi^\epsilon(x)]) = (x, \varphi^\epsilon(x))\,, \tag{52.1.13}$$

which shows that $\tilde{\varphi}^a{}_b(x)$ are simply the coordinates of $\varphi^\epsilon(x)$ in G.

In the next step we derive a coordinate expression for the equivariant map $\hat{\varphi} \in C^\infty_G(P, G)$. This can be done by first evaluating $\hat{\varphi}$ along the section ϵ, which yields

$$[\epsilon(x), \hat{\varphi}(\epsilon(x))] = \tilde{\varphi}(x) = [\epsilon(x), \varphi^\epsilon(x)]\,, \tag{52.1.14}$$

so that $\hat{\varphi}(\epsilon(x)) = \varphi^\epsilon(x)$ for all $x \in U$. To evaluate $\hat{\varphi}$ at any point $p = \epsilon(x) \cdot g$, we use the equivariance of $\hat{\varphi}$ and find

$$\hat{\varphi}(\epsilon(x) \cdot g) = g^{-1}\varphi^\epsilon(x)g\,. \tag{52.1.15}$$

Now recall that the coordinates $(x^\mu, p^a{}_b)$ of the element $p = \epsilon(x) \cdot g$ are defined such that $(p^a{}_b)$ are the coordinates of g in G, i.e., its matrix components in case of a matrix group. Defining $((p^{-1})^a{}_b)$ to be the components of the inverse matrix, we thus find that $\hat{\varphi}$ can be expressed in coordinates as

$$\hat{\varphi}^a{}_b(x, p) = (p^{-1})^a{}_c\tilde{\varphi}^c{}_d(x)p^d{}_b\,. \tag{52.1.16}$$

Keep in mind that $((p^{-1})^a{}_b)$ only denotes the inverse of the *coordinate expression* of p in terms of an invertible matrix. The element $p \in P$ itself does not have any notion of an inverse, since the fibers of P have no group structure.

Finally, we will come to the coordinate expression of the gauge transformation $\varphi : P \to P$, which is defined by $\varphi(p) = p \cdot \hat{\varphi}(p)$. Writing the coordinates of $\varphi(p)$ as $(x^\mu, \varphi^a{}_b(x, p))$, it is easy to see that

$$\varphi^a{}_b(x, p) = p^a{}_c\hat{\varphi}^c{}_b(x, p) = \tilde{\varphi}^a{}_c(x)p^c{}_b\,, \tag{52.1.17}$$

since the matrices $p^a{}_b$ and $(p^{-1})^a{}_b$ on the left just cancel each other.

Hence, in the coordinates induced by choosing a gauge ϵ, the action of the gauge transformation is simply expressed by left multiplication with the matrix representation of φ^ϵ. This nicely reflects the fact that the gauge group $\mathcal{G}$ acts on P from the left. Also one sees immediately that gauge transformations commute with the right translation, since $(\tilde{\varphi}^a{}_c(x)p^c{}_d)g^d{}_b = \tilde{\varphi}^a{}_c(x)(p^c{}_dg^d{}_b)$. However, note that expressing gauge transformations by matrix multiplication with elements of G from the left is only a coordinate description induced by choosing a gauge - the structure group G itself does *not* act from the left on P.

We finally remark that given a gauge $\epsilon : U \to P$ and a gauge transformation $\varphi : P \to P$, one can construct a new gauge as $\epsilon' = \varphi \circ \epsilon$. Hence, a gauge transformation *transforms a gauge* - which justifies the name. We will see later what is the practical use of this notion.

52.2 Infinitesimal gauge transformations

Symmetries of Lagrangian systems are usually described in terms of evolutionary vector fields, which are the infinitesimal generators of the symmetry. Hence, we are also interested in an infinitesimal description of gauge transformations. As in the case of Lie group actions, we may obtain this description by considering one-parameter subgroups $\psi : \mathbb{R} \to \mathcal{G}, t \mapsto \psi_t$. Here $\psi_t : P \to P$ is a gauge transformation, and we further have the composition law

$$\psi_t \circ \psi_s = \psi_{t+s} \tag{52.2.1}$$

for $t, s \in \mathbb{R}$. For every $p \in P$ this fixes a curve $\gamma_p : \mathbb{R} \to P$ given by $\gamma_p(t) = \psi_t(p)$. This curve is vertical, which means that $\pi(\gamma_p(t)) = \pi(p)$. Taking its tangent vector at $t = 0$ we thus obtain a vertical tangent vector. This defines a map $\xi : P \to VP, p \mapsto \dot{\gamma}_p(0)$, and thus a vertical vector field on P. To understand how the tangent vectors $\xi(p)$ and $\xi(p \circ g)$ for some $g \in G$ are related to each other, recall that each ψ_t commutes with the right action of g on P, so that

$$\gamma_{R_g(p)}(t) = \psi_t(R_g(p)) = R_g(\psi_t(p)) = R_g(\gamma_p(t))\,. \tag{52.2.2}$$

Hence, their tangent vectors are related by

$$\xi(R_g(p)) = \dot{\gamma}_{R_g(p)}(0) = R_{g*}(\dot{\gamma}_p(0)) = R_{g*}(\xi(p))\,, \tag{52.2.3}$$

and so $\xi \circ R_g = R_{g*} \circ \xi$ for all $g \in G$. This means that ξ is an equivariant map, where the right action of G on VP is given by $(v, g) \mapsto R_{g*}(v)$. This leads us to the following definition.

Definition 52.2.1 (Infinitesimal gauge transformation)**.** Let $\pi : P \to M$ be a principal G-bundle with Lie group G. An *infinitesimal gauge transformation* is a G-equivariant vertical vector field on P, i.e., a vector field $\xi : P \to VP$ such that $\xi \circ R_g = R_{g*} \circ \xi$ for all $g \in G$.

Since finite gauge transformations form a group, one may already expect that infinitely gauge transformations form a Lie algebra. This is established by the following theorem.

Theorem 52.2.1. *The infinitesimal gauge transformations form a Lie algebra with respect to the Lie bracket of vector fields, which is called the* gauge algebra *and denoted* $\mathfrak{G}$*.*

Proof. The Lie bracket between any two vertical vector fields is again vertical, since the vertical distribution VP is integrable. Further, let $\xi, \xi' \in \mathfrak{G}$. Using the fact that the commutator of vector fields is identical to the Lie derivative, we can write

$$\begin{aligned}
[\xi, \xi'](R_g(p)) &= (\mathcal{L}_\xi \xi')(R_g(p)) \\
&= \lim_{t\to 0} \frac{\psi_{-t*}(\xi'(\psi_t(R_g(p)))) - \xi'(R_g(p))}{t} \\
&= \lim_{t\to 0} \frac{\psi_{-t*}(\xi'(R_g(\psi_t(p)))) - \xi'(R_g(p))}{t} \\
&= \lim_{t\to 0} \frac{\psi_{-t*}(R_{g*}(\xi'(\psi_t(p)))) - R_{g*}(\xi'(p))}{t} \\
&= \lim_{t\to 0} \frac{R_{g*}(\psi_{-t*}(\xi'(\psi_t(p)))) - R_{g*}(\xi'(p))}{t} \\
&= \lim_{t\to 0} \frac{R_{g*}(\psi_{-t*}(\xi'(\psi_t(p))) - \xi'(p))}{t} \\
&= R_{g*}\left(\lim_{t\to 0} \frac{\psi_{-t*}(\xi'(\psi_t(p))) - \xi'(p)}{t}\right) \\
&= R_{g*}((\mathcal{L}_\xi \xi')(p)) \\
&= R_{g*}([\xi, \xi'](p))\,,
\end{aligned} \tag{52.2.4}$$

so that also $[\xi, \xi']$ is equivariant. ■

Another useful relation involving the Lie bracket is the following.

Theorem 52.2.2. *The Lie bracket between any infinitesimal gauge transformation ξ and fundamental vector field $\tilde{X}$ for $X \in \mathfrak{g}$ vanishes,* $[\xi, \tilde{X}] = 0$.

Proof. We make use that the fundamental vector fields commute with finite gauge transformations, as shown in theorem 52.1.1, so that we find

$$\begin{aligned}
[\xi, \tilde{X}](p) &= (\mathcal{L}_\xi \tilde{X})(p) \\
&= \lim_{t\to 0} \frac{\psi_{-t*}(\tilde{X}(\psi_t(p))) - \tilde{X}(p)}{t} \\
&= \lim_{t\to 0} \frac{\psi_{-t*}(\psi_{t*}(\tilde{X}(p))) - \tilde{X}(p)}{t} \\
&= \lim_{t\to 0} \frac{\tilde{X}(p) - \tilde{X}(p)}{t} \\
&= 0\,,
\end{aligned} \tag{52.2.5}$$

so that the Lie bracket vanishes everywhere. ■

Also in the case of the gauge algebra $\mathfrak{G}$ we aim for a more practical description, which will finally provide us with a simple coordinate expression. We can proceed in a similar way as for the gauge group $\mathcal{G}$ and explicitly construct a number of one-to-one correspondences. Here we start with the following.

Theorem 52.2.3. *There is a one-to-one correspondence between infinitesimal gauge transformations $\xi \in \mathfrak{G}$ and equivariant maps $\hat{\xi} \in C^{\infty}_{G}(P, \mathfrak{g})$, where G acts on $\mathfrak{g}$ by the adjoint representation, and the correspondence is established by the fundamental vector fields on P.*

Proof. Recall that following theorem 20.1.3 for each $p \in P$ the fundamental vector fields define a vector space isomorphism between the Lie algebra $\mathfrak{g}$ and the vertical tangent space V_pP, which assigns to $X \in \mathfrak{g}$ the element $\tilde{X}(p) \in V_pP$. This allows us to define $\hat{\xi}(p) \in \mathfrak{g}$ as the unique element such that

$$\xi(p) = \widetilde{\hat{\xi}(p)}(p)\,. \tag{52.2.6}$$

To see that $\hat{\xi}$ is equivariant, we calculate

$$\begin{aligned}
\widetilde{\hat{\xi}(p \cdot g)}(p \cdot g) &= \xi(p \cdot g) \\
&= R_{g*}(\xi(p)) \\
&= R_{g*}\left(\widetilde{\hat{\xi}(p)}(p)\right) \\
&= R_{g*}\left(R_{g^{-1}*}\left(\widetilde{\mathrm{Ad}_{g^{-1}}(\hat{\xi}(p))}(p \cdot g)\right)\right) \\
&= \widetilde{\mathrm{Ad}_{g^{-1}}(\hat{\xi}(p))}(p \cdot g)\,,
\end{aligned} \tag{52.2.7}$$

using theorem 20.1.4. Comparing the left and right hand side, we see that

$$\hat{\xi}(p \cdot g) = \mathrm{Ad}_{g^{-1}}(\hat{\xi}(p))\,, \tag{52.2.8}$$

which shows that $\hat{\xi}$ is indeed equivariant. Conversely, given an equivariant map $\hat{\xi} \in C^{\infty}_{G}(P, \mathfrak{g})$ one defines ξ through the relation (52.2.6) and shows that this is an equivariant vector field by essentially reversing the steps above. ■

Now the next steps are clear. We continue with a straightforward definition.

Definition 52.2.2 (Algebra bundle). Let $\pi : P \to M$ be a principal G-bundle with Lie group G and denote by $\mathrm{Ad} : G \times \mathfrak{g} \to \mathfrak{g}$ the left action of G on its Lie algebra $\mathfrak{g}$ by the adjoint representation. The associated bundle $\pi_{\mathrm{Ad}} : P^{\mathfrak{g}} = P \times_{\mathrm{Ad}} \mathfrak{g} \to M$ is called the *algebra bundle* of P.

Also the final step is straightforward.

Theorem 52.2.4. *There is a one-to-one correspondence between equivariant maps $\hat{\xi} \in C^{\infty}_{G}(P, \mathfrak{g})$, where G acts on $\mathfrak{g}$ by the adjoint representation* Ad*, and sections $\tilde{\xi} \in \Gamma(P^{\mathfrak{g}})$ of the algebra bundle $\pi_{\mathrm{Ad}} : P^{\mathfrak{g}} \to M$.*

Proof. This follows immediately from theorem 20.3.3, as it is simply a particular example. ■

This is the description we want to use. As one might expect from the fact that finite gauge transformations can be described by assigning to each $x \in M$ an element of a group isomorphic to the structure group G, in the infinitesimal case one has a Lie algebra (the fibers of $P^{\mathfrak{g}}$) isomorphic to $\mathfrak{g}$ instead. Also in this case this isomorphism is not canonical, but determined only if one chooses a gauge, i.e., a local section $\epsilon : U \to P$ of P. This then allows expressing an infinitesimal gauge transformation ξ locally by a map $\xi^{\epsilon} : U \to \mathfrak{g}$ as

$$\tilde{\xi}(x) = [\epsilon(x), \xi^{\epsilon}(x)] \in P^{\mathfrak{g}} . \tag{52.2.9}$$

We illustrate the contents of this section by an explicit example.

Example 52.2.1 (Infinitesimal gauge transformations on principal matrix group bundles). We consider the same principal bundle and the same coordinates as introduced in example 52.1.1. In addition, we use the fact that the Lie algebra $\mathfrak{g}$ can be represented by the tangent space T_eG to introduce coordinates by using the coordinate basis induced from the coordinates $(g^a{}_b)$ on G. We denote these coordinates by $(h^a{}_b)$. For the matrix Lie groups listed in example 15.1.3, coordinates induced by their matrix components on their matrix Lie algebras simply agree with the matrix components listed in example 15.6.2. In these coordinates the adjoint representation Ad takes the form $(\mathrm{Ad}_g(X))^a{}_b = g^a{}_c h^c{}_d (g^{-1})^d{}_b$.

We now proceed as in example 52.1.1. First, note that the gauge ϵ induces a local trivialization $\phi_{\mathrm{Ad}} : \pi_{\mathrm{Ad}}^{-1}(U) \to U \times \mathfrak{g}$ of the algebra bundle $P^{\mathfrak{g}}$, such that $\phi_{\mathrm{Ad}}([\epsilon(x), X]) = (x, X)$ for all $(x, X) \in U \times \mathfrak{g}$. We write the coordinates induced by ϕ_{Ad} as $(x^{\mu}, u^a{}_b)$. A global section $\tilde{\xi} : M \to P^{\mathfrak{g}}$, restricted to U, can be expressed in these coordinates as an assignment of $(x^{\mu}, \tilde{\xi}^a{}_b(x))$ to every (x^{μ}) in U. Using the trivialization

$$\phi_{\mathrm{Ad}}(\tilde{\xi}(x)) = \phi_{\mathrm{Ad}}([\epsilon(x), \xi^{\epsilon}(x)]) = (x, \xi^{\epsilon}(x)) \tag{52.2.10}$$

one sees that $\tilde{\xi}^a{}_b(x)$ are simply the coordinates of $\xi^{\epsilon}(x)$ in $\mathfrak{g}$.

We continue by deriving a coordinate expression for the equivariant map $\hat{\xi} \in C^{\infty}_G(P, \mathfrak{g})$. For this purpose we first evaluate $\hat{\xi}$ along the section ϵ, which yields

$$[\epsilon(x), \hat{\xi}(\epsilon(x))] = \tilde{\xi}(x) = [\epsilon(x), \xi^{\epsilon}(x)] , \tag{52.2.11}$$

so that $\hat{\xi}(\epsilon(x)) = \xi^{\epsilon}(x)$ for all $x \in U$. To evaluate $\hat{\xi}$ at any point $p = \epsilon(x) \cdot g$, we use the equivariance of $\hat{\xi}$ and find

$$\hat{\xi}(\epsilon(x) \cdot g) = \mathrm{Ad}_{g^{-1}}(\xi^{\epsilon}(x)) . \tag{52.2.12}$$

Now we once again make use of the fact that the coordinates $(x^{\mu}, p^a{}_b)$ of the element $p = \epsilon(x) \cdot g$ are defined such that $(p^a{}_b)$ are the coordinates of g in G, i.e., its matrix components in case of a matrix group. Recall that we defined $((p^{-1})^a{}_b)$ to be the components of the inverse matrix. Together with a matrix expression of the adjoint representation, this allows us to express $\hat{\xi}$ in coordinates as

$$\hat{\xi}^a{}_b(x, p) = (p^{-1})^a{}_c \tilde{\xi}^c{}_d(x) p^d{}_b . \tag{52.2.13}$$

This expression looks very similar to the corresponding coordinate expression in the finite case shown in example 52.1.1. Note, however, that the objects defined by these expressions lie in different spaces, and that different restrictions apply to the matrix components of Lie group and Lie algebra elements.

We finally come to the coordinate expression of the infinitesimal gauge transformation $\xi : P \to VP$, which is defined by $\xi(p) = \widetilde{\hat{\xi}(p)}(p)$. Since in our chosen coordinates the right action of G on P is given by matrix multiplication $(p \cdot g)^a{}_b = p^a{}_c g^c{}_b$, we can apply essentially the same construction as in example 20.1.1 to write the fundamental vector field $\tilde{X}$ of $X \in \mathfrak{g}$ as

$$\tilde{X} = p^a{}_c X^c{}_b \frac{\partial}{\partial p^a{}_b} . \tag{52.2.14}$$

Writing the coordinates of $\xi(p)$ as $(x^\mu, \xi^a{}_b(x,p))$ with respect to the vertical tangent space basis induced by the coordinates $(p^a{}_b)$ on P, we thus find that

$$\xi(x,p) = \xi^a{}_b(x,p)\frac{\partial}{\partial p^a{}_b} = p^a{}_c\hat{\xi}^c{}_b(x,p)\frac{\partial}{\partial p^a{}_b} = \tilde{\xi}^a{}_c(x)p^c{}_b\frac{\partial}{\partial p^a{}_b}\,, \tag{52.2.15}$$

since the matrices $p^a{}_b$ and $(p^{-1})^a{}_b$ on the left just cancel each other. Again we find that the result formally looks identical to the case of finite gauge transformations. This simply reflects the fact that here we are considering the infinitesimal version of the construction from the former section.

Having this expression at hand, it is now easy to calculate

$$\begin{aligned}
[\xi, \tilde{X}] &= \left[\tilde{\xi}^a{}_e(x)p^e{}_b\frac{\partial}{\partial p^a{}_b}, p^c{}_f X^f{}_d\frac{\partial}{\partial p^c{}_d}\right] \\
&= \tilde{\xi}^a{}_e(x)p^e{}_b\frac{\partial}{\partial p^a{}_b}\left(p^c{}_f X^f{}_d\right)\frac{\partial}{\partial p^c{}_d} - p^c{}_f X^f{}_d\frac{\partial}{\partial p^c{}_d}\left(\tilde{\xi}^a{}_e(x)p^e{}_b\right)\frac{\partial}{\partial p^a{}_b} \\
&= \tilde{\xi}^a{}_e(x)p^e{}_b\delta^c_a X^b{}_d\frac{\partial}{\partial p^c{}_d} - p^c{}_f X^f{}_d\tilde{\xi}^a{}_c(x)\delta^d_b\frac{\partial}{\partial p^a{}_b} \\
&= \left[\tilde{\xi}^a{}_e(x)p^e{}_b X^b{}_d - p^c{}_f X^f{}_d\tilde{\xi}^a{}_c(x)\right]\frac{\partial}{\partial p^a{}_d} \\
&= 0\,,
\end{aligned} \tag{52.2.16}$$

in agreement with theorem 52.2.2.

52.3 Matter fields

Since we are interested in gauge symmetries of Lagrangian systems, we now need a fiber bundle on whose jet bundle the Lagrangian will be defined, and whose sections will be subject to the Euler-Lagrange equations. For a gauge theory this bundle will carry an action of the structure group G. More specifically, it will be an associated bundle to a principal G-bundle $\pi : P \to M$. We introduce the following terminology, which is used in the context of gauge theories.

Definition 52.3.1 (Matter field)**.** A *matter field* is a section $\Phi : M \to P \times_\rho F$ of a fiber bundle $\pi_\rho : P \times_\rho F \to M$ with fiber F associated to a principal G-bundle $\pi : P \to M$ with Lie group G.

Recall that the total space $P \times_\rho F$ is constituted by equivalence classes $[p, f]$, where $p \in P$, $f \in F$ and equivalence is defined by $(p, f) \sim (p \cdot g, \rho(g^{-1}, f))$ for some $g \in G$. We now pose the question how matter fields change if we perform an operation on the principal bundle which preserves its fibers and the right action of the structure group G. We define this operation as follows.

Definition 52.3.2 (Gauge transformation of matter fields)**.** Let $\pi : P \to M$ be a principal G-bundle with Lie group G and $\rho : G \times F \to F$ a left action on the typical fiber G. The gauge group action of $\mathcal{G}$ on the associated bundle $\pi_\rho : P \times_\rho F \to M$ is the left action $\varphi_\rho : P \times_\rho F \to P \times_\rho F$ defined by

$$\varphi_\rho([p, f]) = \varphi \cdot [p, f] = [\varphi(p), f] \tag{52.3.1}$$

for $\varphi \in \mathcal{G}$ and $[p, f] \in P \times_\rho F$.

This action is well defined, since

$$\begin{aligned} \varphi_\rho([p \cdot g, \rho(g^{-1}, f)]) &= [\varphi(p \cdot g), \rho(g^{-1}, f)] \\ &= [\varphi(p) \cdot g, \rho(g^{-1}, f)] \\ &= [\varphi(p), f] \\ &= \varphi_\rho([p, f]) . \end{aligned} \tag{52.3.2}$$

Hence, the result does not depend on the choice of the representative, due to the fact that φ commutes with the right action of the structure group G on P.

Since working with gauge transformations as maps $\varphi : P \to P$ might not be the most convenient description, it is useful to also discuss how we can understand gauge transformations using the other descriptions we provided. If we describe a gauge transformation as an equivariant map $\hat{\varphi} \in C^\infty_G(P, G)$, then we find the expression

$$\varphi_\rho([p, f]) = [\varphi(p), f] = [p \cdot \hat{\varphi}(p), f] = [p, \rho(\hat{\varphi}(p), f)] . \tag{52.3.3}$$

In order to relate this expression to the description in terms of sections $\tilde{\varphi} : M \to P^G$ of the group bundle, recall that $\tilde{\varphi}$ is defined such that

$$\tilde{\varphi}(x) = [p, \hat{\varphi}(p)] \in P^G \tag{52.3.4}$$

for any representative $p \in \pi^{-1}(x)$, and that the result is independent of the choice of the representative. Defining

$$\varphi_\rho([p, f]) = [p, \rho(\hat{\varphi}(p), f)] = [p, \hat{\varphi}(p)] \cdot [p, f] = \tilde{\varphi}(\pi(p)) \cdot [p, f] , \tag{52.3.5}$$

we therefore see that this is also independent of the choice of the representative.

The action of $\mathcal{G}$ on $P \times_\rho F$ is fiber preserving and thus induces an action on the space $\Gamma(P \times_\rho F)$ of sections, where

$$\varphi_\rho(\Phi) = \varphi_\rho \circ \Phi \tag{52.3.6}$$

for $\Phi \in \Gamma(P \times_\rho F)$. Further, we obtain an action on the jet bundles $J^r(P \times_\rho F)$, which is defined such that

$$\varphi_\rho(j^r_x \Phi) = j^r_x \varphi_\rho(\Phi) . \tag{52.3.7}$$

To see that this is well-defined, we have to check that $\varphi_\rho(j^r_x \Phi)$ is independent of the choice of the representative Φ, i.e., that $j^r_x \varphi_\rho(\Phi_1) = j^r_x \varphi_\rho(\Phi_2)$ for two sections Φ_1, Φ_2 satisfying $j^r_x \Phi_1 = j^r_x \Phi_2$. This is indeed the case, which can easily be proven using the fact that $\varphi_\rho : P \times_\rho F \to P \times_\rho F$ is a bundle isomorphism.

Working with bundles whose elements are equivalence classes of sections or orbits of a group action, as it is the case for associated bundles, can sometimes become rather cumbersome. In order to construct local coordinates on these spaces it is more convenient to construct particular local trivializations. Here we can make use of the fact that a local trivialization of an associated bundle can be obtained from a local section of the underlying principal fiber bundle - which is simply a gauge.

With the choice of a gauge on $U \subset M$ we can express matter fields in a simpler form, which is of course valid only locally, i.e., only on U, and depends on the choice of the gauge. For a matter field $\Phi : M \to P \times_\rho F$ we define the local expression $\Phi^\epsilon : U \to F$ relative to the gauge $\epsilon : U \to P$ such that

$$\Phi(x) = [\epsilon(x), \Phi^\epsilon(x)] . \tag{52.3.8}$$

We can also make use of the fiber diffeomorphism introduced in definition 20.3.2 to explicitly write

$$\begin{array}{cccc} \Phi^\epsilon & : & U & \to & F \\ & & x & \mapsto & [\epsilon(x)]^{-1}(\Phi(x)) \end{array} . \tag{52.3.9}$$

Given coordinates on U and F, we thus obtain a coordinate description for Φ.

We also pose the question how the gauge fixed description Φ^ϵ changes if we replace ϵ by a different gauge $\epsilon' = \varphi \circ \epsilon$ obtained by a gauge transformation φ. To see this, we calculate

$$\begin{aligned} [\epsilon'(x), (\varphi_\rho \circ \Phi)^{\epsilon'}(x)] &= (\varphi_\rho \circ \Phi)(x) \\ &= \varphi_\rho([\epsilon(x), \Phi^\epsilon(x)]) \\ &= [\varphi(\epsilon(x)), \Phi^\epsilon(x)] \\ &= [\epsilon'(x), \Phi^\epsilon(x)] . \end{aligned} \tag{52.3.10}$$

From this we find the relation

$$(\varphi_\rho \circ \Phi)^{\epsilon'} = \Phi^\epsilon , \tag{52.3.11}$$

so that a gauge transformation of Φ is compensated by expressing the result in a transformed gauge.

We illustrate the constructions shown above with an example.

Example 52.3.1 (Associated gauge transformation for a matrix Lie group). Let $\pi : P \to M$ be the same principal bundle as in the examples 52.1.1 and 52.2.1, where we have chosen $G \subset M_{n,n}$ to be a matrix Lie group. We further choose ρ to be the natural action of G on the space $F = \mathbb{R}^n$ by left multiplication. A gauge $\epsilon : U \to P$ on $U \subset M$ allows us to construct a local trivialization $\phi_\rho : \pi_\rho^{-1}(U) \to U \times F$ such that $\phi_\rho([\epsilon(x), f]) = (x, f)$ for all $(x, f) \in U \times F$. Using coordinates (x^μ) on U and the canonical coordinates (f^a) on $\mathbb{R}^n$ we thus obtain coordinates (x^μ, f^a) on $\pi_\rho^{-1}(U) \subset P \times_\rho F$.

For the transformed point $\varphi_\rho(p, f)$ after applying a gauge transformation φ we write the coordinates as (x^μ, f'^a), and we aim to construct an expression for f'^a in terms of f^a and the coordinate expressions for φ derived in example 52.1.1. Here it is most convenient to consider the equivariant map $\hat{\varphi} \in C^\infty_G(P, G)$. Evaluating it at $p = \epsilon(x)$ for the chosen gauge, we find

$$\varphi_\rho([\epsilon(x), f]) = [\epsilon(x), \rho(\hat{\varphi}(\epsilon(x)), f)] = [\epsilon(x), \rho(\varphi^\epsilon(x), f)] . \tag{52.3.12}$$

This implies for the coordinate expressions the relation

$$f'^a = \tilde{\varphi}^a{}_b(x) f^b , \tag{52.3.13}$$

so that the fiber coordinates are simply multiplied by the coordinate expression of the gauge transformation.

The expression derived above can also be applied if one considers a section $\Phi : M \to P \times_\rho F$ instead of a single element $[p, f] \in P \times_\rho F$. This section can be written in coordinates as

$$\Phi : (x^\mu) \mapsto (x^\mu, \Phi^a(x)) . \tag{52.3.14}$$

One easily checks that $\Phi^a(x)$ are simply the components of $\Phi^\epsilon(x) \in F = \mathbb{R}^n$, due to the construction of the coordinates on $P \times_\rho F$ by making use of the gauge ϵ.

Finally, we also take a look at infinitesimal gauge transformations. It follows from the structure of finite gauge transformations that they can be defined as follows.

Definition 52.3.3 (Infinitesimal gauge transformation of matter fields)**.** Let $\pi : P \to M$ be a principal G-bundle with Lie group G and $\rho : G \times F \to F$ a left action on the typical fiber G. The gauge algebra action of $\mathfrak{G}$ on the associated bundle $\pi_\rho : P \times_\rho F \to M$ is the vertical vector field $\xi_\rho : P \times_\rho F \to V(P \times_\rho F)$ defined by

$$\xi_\rho([p, f]) = [\bullet, f]_*(\xi(p)) \tag{52.3.15}$$

for $\xi \in \mathfrak{G}$ and $[p, f] \in P \times_\rho F$.

This definition requires a few explanations. First we have to check that ξ_ρ is well-defined, i.e., that it is independent of the choice of the representative p. For this purpose we calculate

$$\begin{aligned}\xi_\rho([p \cdot g, \rho(g^{-1}, f)]) &= [\bullet, \rho(g^{-1}, f)]_*(\xi(p \cdot g))\\ &= ([\bullet, \rho(g^{-1}, f)]_* \circ R_{g*})(\xi(p))\\ &= [\bullet \cdot g, \rho(g^{-1}, f)]_*(\xi(p))\\ &= [\bullet, f]_*(\xi(p))\\ &= \xi_\rho([p, f])\,.\end{aligned} \tag{52.3.16}$$

Now we still have to check that ξ_ρ is a vertical vector field. We thus calculate

$$\begin{aligned}(\pi_{\rho*} \circ \xi_\rho)([p, f]) &= (\pi_{\rho*} \circ [\bullet, f]_* \circ \xi)(p)\\ &= (\pi_* \circ \xi)(p)\\ &= 0\,,\end{aligned} \tag{52.3.17}$$

since $\pi_\rho \circ [\bullet, f] = \pi$, using the fact that $p \mapsto [p, f]$ is a bundle morphism covering the identity, and $\pi_*(\xi(p)) = 0$ since ξ is a vertical vector field.

The object ξ_ρ we have just constructed brings us closer to Lagrangian symmetries. Writing the vertical tangent bundle as $\nu_\rho : V(P \times_\rho F) \to P \times_\rho F$, one realizes that $\xi_\rho^r = \xi_\rho \circ \pi_{\rho\, r,0}$, where $\pi_{\rho\, r,0} : J^r(P \times_\rho F) \to P \times_\rho F$ is a jet bundle projection, satisfies

$$\nu_\rho \circ \xi_\rho^r = \nu_\rho \circ \xi_\rho \circ \pi_{\rho\, r,0} = \pi_{\rho\, r,0}\,, \tag{52.3.18}$$

and so it is an evolutionary vector field. This will allow us to study gauge invariant Lagrangians.

Example 52.3.2 (Infinitesimal associated gauge transformation for a matrix Lie group). We return once more to our example of matrix Lie groups. It follows immediately from the relation (52.3.13) that an infinitesimal gauge transformation acts on matter fields via the vertical vector field

$$\xi_\rho(x, f) = \tilde{\xi}^a{}_b(x) f^b \frac{\partial}{\partial f^a}\,. \tag{52.3.19}$$

From this coordinate expression we can now also calculate the prolongation of this vector field to any jet bundle. Here we restrict ourselves to the first jet bundle $J^1(P \times_\rho F)$. Denoting the induced coordinates on $J^1(P \times_\rho F)$ by $(x^\mu, f^a, f^a_{,\mu})$ we find

$$\operatorname{pr} \xi_\rho(x, f, \partial f) = \tilde{\xi}^a{}_b(x) f^b \frac{\partial}{\partial f^a} + \left(\partial_\mu \tilde{\xi}^a{}_b(x) f^b + \tilde{\xi}^a{}_b(x) f^b_{,\mu}\right) \frac{\partial}{\partial f^a_{,\mu}}\,. \tag{52.3.20}$$

Note the appearance of the derivatives $\partial_\mu \tilde{\xi}^a{}_b(x)$ of the components of ξ. These terms will appear in the infinitesimal transformation of any Lagrangian that depends on the derivative coordinates $f^a_{,\mu}$ (which would be necessary for a kinetic term), and will break the gauge invariance. Hence, we must introduce suitable terms to cancel them.

52.4 Gauge fields

The next ingredient we will need for a gauge theory, and which will finally allow us to construct a gauge invariant Lagrangian, is the notion of a gauge field. Essentially, a gauge field is a principal Ehresmann connection, which is a G-equivariant section A of the jet bundle $\pi_{1,0} : J^1(P) \to P$ over a principal bundle $\pi : P \to M$, following definition 27.1.1. Here we are in a similar situation as in the case of matter fields. Recall that matter fields, i.e., sections $\Phi \in \Gamma(P \times_\rho F)$, can also be understood as G-equivariant maps $\Phi \in C^\infty_G(P, F)$. The situation here is a bit different, since we do not consider arbitrary G-equivariant maps from P to $J^1(P)$, but only sections. However, it is indeed possible to consider gauge fields as sections of a bundle over M, which we construct as follows.

Definition 52.4.1 (Principal connection bundle)**.** Let $\pi : P \to M$ be a principal G-bundle with Lie group G and $J^1(P)$ the first jet space. The space $C = J^1(P)/G$ of G-orbits in $J^1(P)$ together with the canonical projection $\chi : C \to M$ defines a fiber bundle called the *principal connection bundle.*

To check that this construction is valid, first note that the right actions of G on both P and $J^1(P)$ are free, i.e., for each $p \in P$ the subgroup of G sending p to itself contains only the unit element of G, and analogously for $J^1(P)$. As a consequence, all group orbits in P and $J^1(P)$ are diffeomorphic to G. The orbits in P are simply the fibers of the bundle $\pi : P \to M$, so that the space P/G of orbits is canonically diffeomorphic to M. We denote by $C = J^1(P)/G$ the space of orbits in $J^1(P)$. Note that the projection $\pi_1 : J^1(P) \to M$ satisfies $\pi_1 \circ R_g = \pi_1$ for all $g \in G$, i.e., it sends all elements of an orbit to the same image in M. Thus, there is a unique projection $\chi : C \to M$. One easily checks that this defines a fiber bundle. Its sections should already be familiar, as the following theorem states.

Theorem 52.4.1. *There is a one-to-one correspondence between principal Ehresmann connections in a principal G-bundle $\pi : P \to M$ and sections of its principal connection bundle $\chi : C \to M$.*

Proof. Let $A : P \to J^1(P)$ be a principal Ehresmann connection. Since A is an equivariant map, it preserves the orbits, i.e., if p and p' belong to the same orbit in P, then $A(p)$ and $A(p')$ belong to the same orbit in $J^1(P)$. Thus, A defines a map $\Omega : M \to C$ sending orbits in P to orbits in $J^1(P)$, such that the diagram

$$\begin{array}{ccc} P & \xrightarrow{A} & J^1(P) \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle \bullet\cdot G} \\ M & \xrightarrow{\Omega} & C \end{array} \tag{52.4.1}$$

commutes, where the map on the right is simply the canonical projection onto the space of orbits. Further, A is a section of the bundle $\pi_{1,0} : J^1(P) \to P$, so that $\pi_{1,0} \circ A = \mathrm{id}_P$. Thus, for all $p \in P$ we have

$$\chi(\Omega(\pi(p))) = \chi(A(p) \cdot G) = \pi(\pi_{1,0}(A(p))) = \pi(p)\,, \tag{52.4.2}$$

which follows from the fact that $\pi_{1,0}$ is G-equivariant and thus $\pi_1 = \pi \circ \pi_{1,0} = \chi \circ (\bullet \cdot G)$. This shows that Ω is a section of the principal connection bundle $\chi : C \to M$.

Conversely, let $\Omega : M \to C$ be a section of the principal connection bundle. For $p \in P$, define $A(p)$ as the unique jet in $J^1(P)$ such that $\pi_{1,0}(A(p)) = p$ and $A(p) \cdot G = \chi(x)$. One easily checks that the jet $A(p)$ defined this way always exists, that it is unique and that the resulting map $A : P \to J^1(P)$ is a principal Ehresmann connection. ■

With this statement at hand, we can now come to the following definition.

Definition 52.4.2 (Gauge field)**.** Let $\pi : P \to M$ be principal G-bundle with Lie group G. A *gauge field* is a section Ω of the principal connection bundle $\chi : C \to M$.

We finally discuss the question how gauge transformations act on sections of the principal connection bundle. The easiest way to construct such an action is to use G-equivariant connection forms $\theta : TP \to VP$. In this case we can simply define $\bar{\varphi}(\theta) = \varphi_* \circ \theta \circ \varphi_*^{-1}$. To see that this is again a G-equivariant connection form, first note that by construction, $\bar{\varphi}(\theta)$ is a vector bundle homomorphism covering the identity on P. Further, recall from theorem 52.1.1 that the fundamental vector fields are invariant under the action of a gauge transformation, $\varphi_* \circ \tilde{X} = \tilde{X} \circ \varphi$. From this in particular follows that the vertical tangent bundle VP is invariant under φ, i.e., $\varphi_*(v) \in VP$ for all $v \in VP$. We thus have

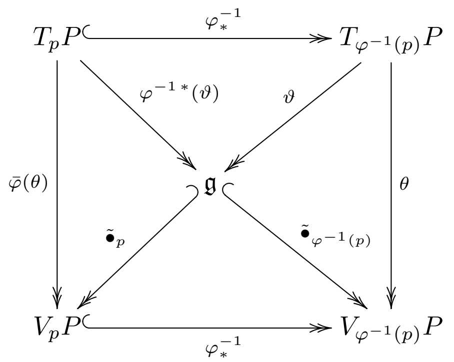
$$\bar{\varphi}(\theta)(v) = (\varphi_* \circ \theta \circ \varphi_*^{-1})(v) = (\varphi_* \circ \varphi_*^{-1})(v) = v\,, \tag{52.4.3}$$

so that $\bar{\varphi}(\theta)$ restricts to the identity on VP. Finally, for all $g \in G$ and $w \in TP$ we find

$$\bar{\varphi}(\theta)(R_{g*}(w)) = (\varphi_* \circ \theta \circ \varphi_*^{-1} \circ R_{g*})(w) = (R_{g*} \circ \varphi_* \circ \theta \circ \varphi_*^{-1})(w) = R_{g*}(\bar{\varphi}(\theta)(w))\,, \tag{52.4.4}$$

where we used the fact that all maps appearing above are G-equivariant, so that we can permute R_{g*} to the left. This shows that also $\bar{\varphi}(\theta)$ is G-equivariant.

From the invariance of the fundamental vector fields shown in theorem 52.1.1 we can see that the corresponding principal G-connection $\vartheta \in \Omega^1(P, \mathfrak{g})$ transforms in an even simpler way in form of the pullback $\varphi^{-1\,*}(\vartheta)$. This can be seen most easily using the following commutative diagram.



(52.4.5)

To see that $\varphi^{-1\,*}(\vartheta)$ is the principal G-connection of the connection form $\bar{\varphi}(\theta)$ and vice versa, we must show that the left triangle commutes. This is the case if and only if all other building blocks of this diagram commute. The lower triangle commutes due to the gauge invariance of the fundamental vector fields shown in theorem 52.1.1. The right triangle commutes by the relation between the connection form θ and the principal G-connection ϑ. The upper triangle commutes by definition of the pullback. Finally, the surrounding square commutes by definition of $\bar{\varphi}(\theta)$.

With this preliminary discussion we can now describe gauge transformations of principal Ehresmann connections, and thus of gauge fields. Let $A : P \to J^1(P)$ be a principal Ehresmann connection, which assigns to $p \in P$ with $\pi(p) = x$ a jet $A(p) = j^1_x\sigma_p \in J^1(P)$, and θ the corresponding connection form. We define $\bar{\varphi}(A)$ as the principal Ehresmann connection corresponding to $\bar{\varphi}(\theta)$. Then we have $\bar{\varphi}(A)(p) = j^1_x(\varphi \circ \sigma_{\varphi^{-1}(p)})$. To check this, we calculate

$$\begin{aligned} \bar{\varphi}(\theta)_p(w) &= w - \varphi_*(\sigma_{\varphi^{-1}(p)*}(\pi_*(w))) \\ &= \varphi_*\left[\varphi_*^{-1}(w) - \sigma_{\varphi^{-1}(p)*}(\pi_*(\varphi_*^{-1}(w)))\right] \\ &= \varphi_*(\theta_{\varphi^{-1}(p)}(\varphi_*^{-1}(w)))\,, \end{aligned} \tag{52.4.6}$$

which shows that our formula is correct. Finally, taking the quotient by the group action of G yields the action of the gauge group on the space $\Gamma(C)$ of sections of the connection bundle $\chi : C \to M$.

We also discuss how we can express a gauge field Ω if we have fixed a gauge, i.e., a local section $\epsilon : U \to P$ of our principal bundle. The easiest way is to view the gauge field in terms of the corresponding principal G-connection ϑ on P, and thus in particular as a $\mathfrak{g}$-valued one-form $\vartheta \in \Omega^1(P, \mathfrak{g})$ on P. Given a gauge ϵ we can thus define

$$\Omega^\epsilon = \epsilon^*(\vartheta) \in \Omega^1(M, \mathfrak{g}) . \tag{52.4.7}$$

Thus, the connection pulls back to a $\mathfrak{g}$-valued one-form on M. This is the description most often encountered in field theory.

Finally, we discuss how this description changes if we use a different gauge $\epsilon' = \varphi \circ \epsilon$. Let Ω' be the gauge field described by the connection form $\varphi^*(\vartheta)$. From the properties of the pullback follows immediately

$$\Omega'^{\epsilon'} = (\epsilon'^* \circ \varphi^{-1\,*})(\vartheta) = (\varphi^{-1} \circ \epsilon')^*(\vartheta) = \epsilon^*(\vartheta) = \Omega^\epsilon . \tag{52.4.8}$$

Hence, we find that also for gauge fields the local expressions in different gauges are related via the corresponding gauge transformations, in full analogy to the case for matter fields.

We will illustrate the constructions shown in this section using the example of a matrix Lie group G.

Example 52.4.1 (Gauge fields for a matrix Lie group). We consider the same principal bundle $\pi : P \to M$ with structure group G and coordinates derived from a local section $\epsilon : U \to P$ on $U \subset M$ as introduced in example 52.1.1. We write the corresponding coordinates on the first jet bundle $J^1(P)$ as $(x^\mu, p^a{}_b, p^a{}_{b,\mu})$. In order to construct coordinates on the principal connection bundle $C = J^1(P)/G$, we must determine the orbits of the right translation on $J^1(P)$. Recall that in coordinates the right translation reads

$$R : ((x^\mu, p^a{}_b, p^a{}_{b,\mu}), g^a{}_b) \mapsto (x^\mu, p^a{}_c g^c{}_b, p^a{}_{c,\mu} g^c{}_b) . \tag{52.4.9}$$

Obviously the coordinates (x^μ) are constant on every orbit, since they are not affected by the right translation. Note that also the expression $p^a{}_{c,\mu}(p^{-1})^c{}_b$ is invariant under the group action. Hence, we can parametrize the orbit space C (locally) with coordinates

$$(x^\mu, c^a{}_{b\mu}) = \{(x^\mu, g^a{}_b, c^a{}_{c,\mu} g^c{}_b), (g^a{}_b) \in G\} \in C . \tag{52.4.10}$$

Recall that a principal connection $\omega : P \to J^1(P)$ is expressed in coordinates through the connection coefficients $\Gamma^a{}_{b\mu}$ using the formula (27.1.23). By comparison with our coordinates on C, we see that ω maps the fiber over (x^μ) to the orbit with coordinate $c^a{}_{b\mu} = -\Gamma^a{}_{b\mu}(x)$. Hence, the gauge field Ω is expressed in these coordinates as

$$\Omega : (x^\mu) \mapsto (x^\mu, -\Gamma^a{}_{b\mu}(x)) . \tag{52.4.11}$$

In other words, the connection coefficients are simply the coordinates on the principal connection bundle, up to a sign.

If we have chosen a gauge $\epsilon : U \to P$ on an open set $U \subset M$, we can also derive the gauge fixed expression Ω^ϵ for the connection. This derivation becomes most simple in the case that the trivialization ϕ used in constructing the coordinates is given such that $\phi(\epsilon(x)) = (x, e)$. In this case we have $\epsilon : (x^\mu) \mapsto (x^\mu, \delta^a_b)$ for all $x \in U$, so that we find, using the formula (27.1.18),

$$[\Omega^\epsilon(x)]^a{}_b = [\epsilon^*(\vartheta)(x)]^a{}_b = \Gamma^a{}_{b\mu}(x)\mathrm{d}x^\mu . \tag{52.4.12}$$

This is the most common way to express a gauge field. Again we see that the coordinate expression is just given by the connection coefficients.

We finally show what happens if we use a different gauge $\epsilon' = \varphi \circ \epsilon$ obtained from a gauge transformation φ. For this purpose we write this new gauge in coordinates as $\epsilon' : (x^\mu) \mapsto (x^\mu, \epsilon'^a{}_b(x))$. By using once again the formula (27.1.18) in terms of the connection coefficients we then find

$$[\Omega^{\epsilon'}(x)]^a{}_b = (\epsilon'^{-1}(x))^a{}_c \left[\frac{\partial \epsilon'^c{}_b}{\partial x^\mu} + \epsilon'^d{}_b(x)\Gamma^c{}_{d\mu}(x)\right] \mathrm{d}x^\mu \,. \tag{52.4.13}$$

This shows the transformation behavior of a gauge field under gauge transformations. Note in particular that it transforms *inhomogeneously*, as one can see from the term involving derivatives of ϵ', which reminds that connections are sections of an *affine bundle*.

52.5 Gauge invariance of Lagrangian systems

Finally, we pose the question how to treat theories involving gauge and matter fields using the Lagrangian formalism we introduced in a previous lecture. Here we restrict ourselves to theories in which there is only one Lie group G and one principal G-bundle $\pi : P \to M$ and summarize all matter fields within a single associated bundle $\pi_\rho : P \times_\rho F \to M$. Thus, a field configuration is given by a gauge field $\Omega : M \to C$ and a matter field $\Phi : M \to P \times_\rho F$. We can combine both into a section (Ω, Φ) of the Cartesian product bundle

$$E = C \times_M (P \times_\rho F) = \bigcup_{x \in M} C_x \times (P_x \times_\rho F) \,, \tag{52.5.1}$$

whose fibers are the Cartesian products of the fibers of C and $P \times_\rho F$, and which canonically inherits a bundle projection $\Pi : E \to M$. The gauge group acts on both Ω and Φ, and thus also on the pair (Ω, Φ). This defines an action of the gauge group on sections of E, and thus also on the jet bundles $J^r(E)$. This allows us to define the following notion.

Definition 52.5.1 (Gauge invariant Lagrangian)**.** A Lagrangian $L \in \Omega^{n,0}(J^r(E))$ is called *gauge invariant* if it is invariant under the action of the gauge group on $J^r(E)$.

52.6 Conserved gauge currents

52.7 Spontaneous symmetry breaking

Chapter 53

Hamiltonian mechanics

53.1 Hamiltonian systems

Hamilton theory appears in physics in various different flavors, using similar geometric structures, but with certain differences.

53.2 Canonical coordinates

53.3 Canonical transformations

53.4 Action-angle variables

53.5 Legendre transformation

53.6 Constrained systems

Chapter 54

Canonical Hamiltonian field theory

Chapter 55

Covariant Hamiltonian field theory

Chapter 56

Hamilton-Jacobi theory

Chapter 57

Dynamical systems

57.1 Autonomous continuous dynamical systems

57.2 Autonomous discrete dynamical systems

57.3 Non-autonomous continuous dynamical systems

57.4 Non-autonomous discrete dynamical systems

57.5 Fixed points

57.6 Singularities

57.7 Stability

57.8 Poincaré sections

Chapter 58

Perturbation theory

Chapter 59

Geometric quantization

59.1 Prequantization

Chapter 60

BRST quantization

Bibliography

[AM78] ABRAHAM, Ralph ; MARSDEN, Jerrold E.: *Foundations of Mechanics*. Reading : Benjamin-Cummings, 1978. – ISBN 0–8053–0102–X

[And92] ANDERSON, Ian M.: Introduction to the variational bicomplex. Version: 1992. http://dx.doi.org/10.1090/conm/132/1188434. In: *Mathematical aspects of classical field theory (Seattle, WA, 1991)* Bd. 132. Amer. Math. Soc., Providence, RI, 1992. – DOI 10.1090/conm/132/1188434, S. 51–73

[Ant03] ANTONELLI, Peter L.: *Handbook of Finsler Geometry*. Dordrecht, Boston, London : Kluwer, 2003. – ISBN 1–4020–1557–7

[Bau14] BAUM, Helga: *Eichfeldtheorie*. Berlin, Heidelberg : Springer, 2014. – ISBN 978–3–642–38538–4

[BCS91] BAO, David ; CHERN, Shiing-Shen ; SHEN, Zhongmin: *An Introduction to Riemann-Finsler Geometry*. New York : Springer, 1991. – ISBN 0–387–98948–X

[Ber01] BERNDT, Rolf: *An Introduction to Symplectic Geometry*. Providence : American Mathematical Society, 2001. – ISBN 0–8218–2056–7

[Ber02] BERGER, Marcel: *A Panoramic View of Riemannian Geometry*. Berlin, Heidelberg : Springer, 2002. – ISBN 978–35–4065–317–2

[BM07] BUCATARU, Ioan ; MIRON, Radu: *Finsler-Lagrange Geometry*. Bucharest : Editura Academiei Romane, 2007

[Bur85] BURKE, William L.: *Applied differential geometry*. Cambridge University Press, 1985. – ISBN 0–521–26317–4

[CCL99] CHERN, S. S. ; CHEN, W. H. ; LAM, K. S.: *Lectures on Differential Geometry*. Singapore : World Scientific, 1999. – ISBN 981–023–494–5

[Che11] CHEN, Bang-Yen: *Pseudo-Riemannian geometry, δ-invariants and applications*. Singapore : World Scientific, 2011. – ISBN 987–981–4329–63–7

[CS05] CHERN, Shiing-Shen ; SHEN, Zhongmin: *Riemann-Finsler Geometry*. Singapore : World Scientific, 2005. – ISBN 981–238–357–3

[CS16] CRAMPIN, Mike ; SAUNDERS, David J.: *Cartan Geometries and their Symmetries*. Atlantis Press, 2016. – ISBN 978–94–6239–192–5

[Fec11] FECKO, Mariàn: *Differential geometry and Lie groups for physicists*. Cambridge University Press, 2011. – ISBN 9780521187961, 9780521845076, 9780511242960

[Gar11] GARLING, D. J. H.: *Clifford Algebras: An Introduction*. Cambridge : Cambridge University Press, 2011. – ISBN 978–1–107–096638–7

[GM92] Gotay, Mark J. ; Marsden, Jerrold E.: Stress-energy-momentum tensors and the Belinfante-Rosenfeld formula. Version: 1992. http://dx.doi.org/10.1090/conm/132/1188448. In: *Mathematical aspects of classical field theory (Seattle, WA, 1991)* Bd. 132. Amer. Math. Soc., Providence, RI, 1992. – DOI 10.1090/conm/132/1188448, S. 367–392

[GMS09] Giachetta, Giovanni ; Mangiarotti, Luigi ; Sardanashvily, Gennadi: *Advanced classical field theory*. Singapore : World Scientific, 2009. – ISBN 978–981–283–895–7

[GQ20a] Gallier, Jean ; Quaintance, Jocelyn: *Differential Geometry and Lie Groups - A Computational Perspective*. Cham : Springer, 2020. http://dx.doi.org/10.1007/978-3-030-46040-2. http://dx.doi.org/10.1007/978-3-030-46040-2. – ISBN 978–3–030–46039–6

[GQ20b] Gallier, Jean ; Quaintance, Jocelyn: *Differential Geometry and Lie Groups - A Second Course*. Cham : Springer, 2020. http://dx.doi.org/10.1007/978-3-030-46047-1. http://dx.doi.org/10.1007/978-3-030-46047-1. – ISBN 978–3–030–46046–4

[Har90] Harvey, F. R.: *Spinors and Calibrations*. San Diego, London : Academic Press, 1990. – ISBN 0–12–329650–1

[HO01] Hehl, Friedrich W. ; Obukhov, Yuri: *Foundations of Classical Electrodynamics*. Boston : Birkhäuser, 2001. – ISBN 0817642226, 9780817642228

[Huy04] Huybrechts, Daniel: *Complex Geometry*. Berlin, Heidelberg : Springer, 2004. – ISBN 3–540–21290–6

[Jan21] Jancevicz, Bernard: *Directed Quantities in Electrodynamics*. Cham : Birkhäuser, 2021. http://dx.doi.org/10.1007/978-3-030-90471-5. http://dx.doi.org/10.1007/978-3-030-90471-5. – ISBN 978–3–030–90470–8

[KN63] Kobayashi, Shoshichi ; Nomizu, Katsumi: *Foundations of Differential Geometry*. New York : Interscience Publishers, 1963

[Kod86] Kodaira, Kunihiko: *Complex Manifolds and Deformation of Complex Structures*. New York : Springer, 1986. – ISBN 0–387–96188–7

[Kod07] Kodaira, Kunihiko: *Complex Analysis*. Cambridge : Cambridge University Press, 2007. – ISBN 987–052–180–937–5

[Kru15] Krupka, Demeter: *Introduction to Global Variational Geometry*. Atlantis Press, 2015. – ISBN 978–94–6239–072–0

[KS11] Kosmann-Schwarzbach, Yvette: *The Noether Theorems*. New York : Springer, 2011. – ISBN 978–0–387–87867–6

[KSM93] Kolář, Ivan ; Slovák, Jan ; Michor, Peter W.: *Natural Operations in Differential Geometry*. Berlin, Heidelberg : Springer, 1993. – ISBN 3–540–56235–4, 0–387–56235–4

[Lan85] Lang, Serge: *Differential Manifolds*. New York : Springer, 1985. http://dx.doi.org/10.1007/978-1-4684-0265-0. http://dx.doi.org/10.1007/978-1-4684-0265-0. – ISBN 978–0–387–96113–2, 978–1–4684–0265–0

[Lee13] Lee, John M.: *Introduction to Smooth Manifolds*. New York : Springer, 2013. – ISBN 978–1–4419–9981–8

[Lee18] Lee, John M.: *Introduction to Riemannian Manifolds*. New York : Springer, 2018. – ISBN 978–3–319–91754–2

[LM89] LAWSON, H. B. ; MICHELSOHN, Marie-Louise: *Spin Geometry.* Princeton : Princeton University Press, 1989. – ISBN 0–691–08542–0

[Lou01] LOUNESTO, Pertti: *Clifford Algebras and Spinors.* Cambridge : Cambridge University Press, 2001. – ISBN 0–521–00551–5

[MA94] MIRON, Radu ; ANASTASIEI, Mihai: *The Geometry of Lagrange Spaces: Theory and Applications.* Dordrecht : Springer, 1994. – ISBN 978–94–010–4338–0

[McC07] MCCABE, Gordon: *The Structure and Interpretation of the Standard Model.* Amsterdam : Elsevier, 2007. – ISBN 978–0–444–53112–4

[Mic08] MICHOR, Peter W.: *Topics in Differential Geometry.* Providence : American Mathematical Society, 2008. – ISBN 978–0–8218–2003–2

[Noe18] NOETHER, Emmy: Invariant Variation Problems. In: *Gott. Nachr.* 1918 (1918), S. 235–257. http://dx.doi.org/10.1080/00411457108231446. – DOI 10.1080/00411457108231446. – [Transp. Theory Statist. Phys.1,186(1971)]

[Olv86] OLVER, Peter J.: *Applications of Lie Groups to Differential Equations.* New York : Springer, 1986. – ISBN 0–387–94007–3

[O'N83] O'NEILL, Barrett: *Semi-Riemannian geometry.* New York, London : Academic Press, 1983. – ISBN 0–12–5267740–1

[Por95] PORTEOUS, Ian R.: *Clifford Algebras and the Classical Groups.* Cambridge : Cambridge University Press, 1995. – ISBN 0–521–55177–3

[Run59] RUND, Hanno: *The Differential Geometry of Finsler Spaces.* Berlin, Göttingen, Heidelberg : Springer, 1959. – ISBN 978–3–642–51612–2

[Sar13] SARDANASHVILY, Gennadi: *Advanced Differential Geometry for Theoreticians. Fiber bundles, jet manifolds and Lagrangian theory.* Saarbrücken : Lambert Academic Publishing, 2013. – ISBN 978–3–659–37815–7

[Sar16] SARDANASHVILY, Gennadi: *Noether's Theorems.* Atlantis Press, 2016. – ISBN 978–94–6239–170–3

[Sau89] SAUNDERS, David J.: *The geometry of jet bundles.* Cambridge : Cambridge University Press, 1989. – ISBN 0–521–36948–7

[Sch91] SCHWARTZ, Laurent: *Analyse.* Paris : Hermann, 1991. – ISBN 2–7056–6161–8, 2–7056–6162–5, 2–7056–6163–2, 2–7056–6186–7

[Sha97] SHARPE, R. W.: *Differential Geometry.* New York : Springer, 1997. – ISBN 0–387–94732–9

[She01a] SHEN, Zhongmin: *Differential Geometry of Spray and Finsler spaces.* Boston : Kluwer Academic Publishers, 2001. – ISBN 9780792368687

[She01b] SHEN, Zhongmin: *Lectures on Finsler Geometry.* Singapore : World Scientific, 2001. – ISBN 981–02–4530–0

[SLK14] SZILASI, József ; LOVAS, Rezső L. ; KERTÉSZ, Dávid C.: *Connections, Sprays and Finsler Structures.* Singapore : World Scientific, 2014. – ISBN 987–981–4440–09–7

[Spi99] SPIVAK, Michael: *A Comprehensive Introduction to Differential Geometry.* Houston : Publish or Perish, 1999. – ISBN 0914098705 9780914098706 0914098713 9780914098713 0914098721 9780914098720 091409873X 9780914098737 0914098748 9780914098744

[SS16] SHEN, Yi-Bing ; SHEN, Zhongmin: *Introduction to Modern Finsler Geometry.* Singapore : World Scientific, 2016. – ISBN 987–981–4713–16–0

[VR16] Vaz, Jayme ; Rocha, Roldão da: *An Introduction to Clifford Algebras and Spinors.* Oxford : Oxford University Press, 2016. – ISBN 978–0–19–878292–6

[Yan57] Yano, Kentaro: *The Theory of Lie Derivatives and its Applications.* Amsterdam : North-Holland, 1957. – ISBN 978–0–7204–2104–0

Index