Differentialgeometrie I und II

Skript zur Vorlesung[1] von Prof T. Rivière

Wintersemester 2004/05 und Sommersemester 2005

[1] Dieses Skript wurde von Thiemo Kessel verfaßt und ist eine erste Fassung vom Dezember 2005. e-mail: *tkessel@math.ethz.ch*

Contents

1 Differential Calculus in Normed Vector Spaces

1.1 Differentiability and Tangent Map

Definition 1.1. *Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be two normed vector spaces. Let $U \subset E$ be open and let f be a map from U into F. Then f is called* ***differentiable*** *at $a \in U$ if there exists a linear and continuous map $df_a \in \mathcal{L}(E, F)$ such that*

$$\lim_{h \longrightarrow 0} \big\| f(a+h) - f(a) - df_a \cdot h \big\|_F \|h\|_E^{-1} = 0\,.$$

The map df_a is then called the ***differential*** *or the* ***tangent map*** *of f at a.*

Remarks. 1. Compared to the usual definition in the finite dimensional normed vector space $\mathbb{R}^n$ we require the differential to be also continuous.

2. Note that the space of linear and continuous maps from E into F, denoted by $\mathcal{L}(E, F)$, is a normed vector space for the norm

$$|||g|||_{\mathcal{L}(E,F)} = \sup_{\|x\|_E \leq 1} \|g(x)\|_F\,,$$

with $x \in E$ and $g \in \mathcal{L}(E, F)$.

With the same notations as in the previous definition we have the following

Definition 1.2. *The map $f : U \longrightarrow F$ is said to be* ***continuously differentiable*** *on U if for every $a \in U$ the differential df_a exists and the map*

$$\begin{aligned} df : U &\longrightarrow (\mathcal{L}(E, F), |||\cdot|||)\,, \\ a &\longmapsto df_a \end{aligned}$$

is continuous. We write $C^1(U, F)$ for the vector space of continously differentiable maps.

Next, we consider the case of *finite* dimensional normed vector spaces.

Proposition 1.3. *Let $E = \mathbb{R}^n$ and $U \subset \mathbb{R}^n$ open. Then $f \in C^1(U, F)$ if and only if for every $a \in U$ and $i = 1, \ldots, n$, we have that*

$$\frac{\partial f}{\partial x_i}(a) := \lim_{h \longrightarrow 0} \frac{f(a + h\, e_i) - f(a)}{h}$$

exist, and the ***partial derivatives***

$$\frac{\partial f}{\partial x_i} : U \longrightarrow F\,,$$

are continuous.

When we also assume that $F = \mathbb{R}^m$ is finite dimensional, then $f = (f_1, \dots, f_m) \in C^1(U, \mathbb{R}^m)$ if and only if $f_j \in C^1(U, \mathbb{R})$ for every $j = 1, \dots, m$. Then we also have for $h \in \mathbb{R}^n$ that

$$df_a \cdot h = J_a f \cdot h\,, \tag{1.1}$$

where $J_a f$ denotes the **Jacobian matrix** of f at the point a given by the following $(m \times n)$-matrix:

$$Jf_a = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \dots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}.$$

Note that in the case $m = 1$ the differential of f is simply given by

$$df_a \cdot h = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) \cdot h_i\,, \tag{1.2}$$

where again $h = (h_1, \dots, h_n) \in \mathbb{R}^n$. – Now, we quote a useful result:

Proposition 1.4 (Chain Rule). *Let $f : U \subset E \longrightarrow F$, U open, and let $g : F \longrightarrow G$ with a third normed vector space G. Let $a \in U$ and assume that f and g are differentiable at a, respectively at $f(a)$ (or $f \in C^1(U, F)$ and $g \in C^1(F, G)$). Then the composition $g \circ f$ is differentiable at a (or $g \circ f \in (U, G)$). Moreover, we have the formula*

$$d(g \circ f)_a = dg_{f(a)} \cdot df_a\,. \tag{1.3}$$

If $g = f^{-1} : F \longrightarrow E = G$ in the previous proposition, we obtain immediately

Corollary 1.5. *Let $U \subset E$ and $V \subset F$ two open subsets and $f : U \longrightarrow V$. Assume that f is differentiable at $a \in U$, $f^{-1} : V \longrightarrow U$ exists and is differentiable at $f(a) \in V$. Then*

$$d(f^{-1})_{f(a)} = (df_a)^{-1}\,.$$

From Definition 1.2 we arrive at a next step.

Definition 1.6. *Let $f \in C^1(U,F)$. We say that f is* ***twice continuously differentiable****, $f \in C^2(U,F)$, if the map $df : U \longrightarrow (\mathcal{L}(E,F), |||\cdot|||)$ is continuously differentiable. Similarly, we define $C^p(U,F)$,* ***p-times continuously differentiable*** *functions, and*

$$C^\infty(U,F) := \bigcap_{1\le p} C^p(U,F)\,.$$

Note that $d^2 f$ is symmetric, i.e., for $f \in C^2(U,F)$, we have that

$$d(df_a \cdot h)k = d(df_a \cdot k)h\,, \qquad \forall k,h \in E\,.$$

Example 1.7. The cut-off functions on $\mathbb{R}^n$ are an important example for C^∞-maps (see []).

Now, we generalize a well-known result for functions on $\mathbb{R}$.

Lemma 1.8 (Control Growth Lemma). *Let $f : E \longrightarrow F$ be differentiable in $U \subset E$ convex and open. Moreover, we assume that*

$$\sup_{c\in U} |||df_c||| = M < +\infty\,.$$

Then for every a, $b \in U$ the following inequality holds:

$$\|f(b) - f(a)\|_F \le M\,\|b-a\|_E\,. \tag{1.4}$$

Proof. Let a, $b \in U$ and $x_t := t\,b + (1-t)a$, for $t \in [0,1]$. By the convexity of $U \subset E$, it follows that $x_t \in U$. We then define

$$\begin{aligned} \varphi(t) : [0,1] &\longrightarrow F\,,\\ t &\longmapsto \varphi(t) := f(x_t) = f\big(t\,b + (1-t)a\big)\,. \end{aligned}$$

This map is differentiable for every $t \in [0,1]$ with $\varphi'(t) = df_{x_t} \cdot (b-a)$.

Next, let $\varepsilon > 0$ and for $0 \le s < 1$, we define

$$A_\varepsilon := \Big\{ s \in \mathbb{R} \;:\; \|\varphi(t) - \varphi(0)\|_F \le t(M\,\|b-a\|_E + \varepsilon) \;\; \forall t \in [0,s) \Big\}\,.$$

We want to show that $A_\varepsilon = [0,1]$. – It is clear from the definition that A_ε is a closed interval containing $s = 0$. Concerning the upper bound of A_ε, denoted by u, we assume for a contradiction that $0 < u < 1$. Since $\varphi(t)$ is differentiable at u, we deduce from Definition 1.1 that

$$\big\|\varphi\big(u + (t-u)\big) - \varphi(u) - \varphi'(u)(t-u)\big\|_F \le \varepsilon\,(t-u)\,,$$

where $u < t \le 1$ and close enough to u. The last inequality then gives

$$\begin{aligned} \big\|\varphi\big(u + (t-u)\big) - \varphi(u)\big\|_F &\le \left(\sup_{u\in(0,1)} |||\varphi'(u)||| + \varepsilon\right)(t-u)\\ &\le \big(M\|b-a\|_E + \varepsilon\big)(t-u)\,. \end{aligned}$$

On the other hand, from the fact that A_ε is a closed interval and that u is by assumption the upper bound, we have

$$\|\varphi(u) - \varphi(0)\|_F \le u(M\,\|b-a\|_E + \varepsilon)\,.$$

Thus putting the results together

$$\begin{aligned}\|\varphi(t) - \varphi(0)\|_F &\le \|\varphi(t) - \varphi(u)\|_F + \|\varphi(u) - \varphi(0)\|_F \\ &\le \big(M\|b-a\|_E + \varepsilon\big)(t-u) + u(M\,\|b-a\|_E + \varepsilon) \\ &\le t(M\,\|b-a\|_E + \varepsilon)\,,\end{aligned}$$

we obtain that u is *not* the upper bound of A_ε, since $t > u$ is also contained in A_ε. In other words, we get the openness in $[0,1]$ of the set A_ε.

Hence, the above shows that $A_\varepsilon = [0,1]$. Moreover, we have

$$\|f(b) - f(a)\|_F = \|\varphi(1) - \varphi(0)\|_F \le M\,\|b-a\|_E + \varepsilon$$

for all $\varepsilon > 0$, showing (1.4). □

Under the stronger assumption that $f \in C^1(U,F)$, the control growth lemma can be established in a straightforward manner. – Using the definition of φ and since $f \in C^1(U,F)$, we can write

$$\|f(b) - f(a)\|_F = \|\varphi(1) - \varphi(0)\|_F = \left\| \int_0^1 \varphi'(t)\,dt \right\|\,.$$

A short calculation, using Proposition 1.4, gives

$$\|\varphi'(t)\|_F = \big\|d\big(f(x_t)\big)\big\|_F = \big\|d\big(f(t\,b + (1-t)a)\big)\big\|_F = \|df_{x_t}(b-a)\|_F\,,$$

and the fact that $x_t \in U$ implies by assumption

$$\|df_{x_t}(b-a)\|_F \le M\,\|b-a\|_E\,.$$

Hence, we have directly established (1.4).

Definition 1.9. *Let $(E, \|\cdot\|_E)$ be a normed vector space. Then E is a* ***Banach space*** *whenever E is complete for $\|\cdot\|_E$, i.e., every Cauchy sequence in E converges to a limit for $\|\cdot\|_E$. An* ***isomorphism*** *of Banach spaces is a linear, continuous and bijective map.*

Using the Control Growth Lemma 1.8, we prove the following important theorem.

Theorem 1.10 (Local Inversion Theorem). *Let $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ be two Banach spaces, $U \subset E$ open, and let $f \in C^k(U,F)$ with $k \ge 1$. Assume that there exists $a \in U$ such that df_a is an isomorphism between E and F. Then there exists an open neighborhood V of a and an open neighborhood W of $f(a)$ such that*

(i) the map $f\big|_V \colon V \longrightarrow W$ *is invertible;*
(ii) for the inverse, we have $f\big|_V^{-1} \in C^k(W, E)$.

Remark. Since the differential df_a is by assumption invertible the open mapping theorem implies directly that $(df_a)^{-1}$ is continuous.

Proof. It suffices to proof the result for $E = F$, $a = 0$, $f(0) = 0$ and $df_0 = id_E$. Indeed, we can renormalize f by (note that $(df_a)^{-1}$ exists)

$$f_R(x) := (df_a)^{-1} \cdot \big(f(a + x) - f(a)\big)\,,$$

verifying the conditions above, and it is easily shown that if the local inversion theorem is proved for f_R then it is also proved for f. So we will prove the theorem for the renormalized map, also denoted by f.

First, note that since by assumption $f \in C^1(U, F)$, there exists $r > 0$ such that for all $\tilde{x} \in B_r(0)$, we have (see Definition 1.2)

$$\|df_{\tilde{x}} - df_0\| = |||df_{\tilde{x}} - id_E||| \leq \frac{1}{2}\,. \tag{1.5}$$

We show that f is locally invertible at $a = 0$, i.e., for all $y \in B_{r/2}(0)$ (recall that $f(0) = 0 \in E$), we want to find a unique $x \in B_r(0)$ such that $f(x) = y$. For this purpose, we define

$$h_y(x) := -f(x) + x + y$$

and show that h_y is a *contraction* from $B_r(0)$ into $B_r(0)$, i.e., there exists a constant $\alpha < 1$ such that for all x and x' in $B_r(0)$ the following inequality holds:

$$\|h_y(x) - h_y(x')\| \leq \alpha\,\|x - x'\|\,.$$

By assumption h_y is a C^k-map and for all $\tilde{x} \in B_r(0)$, we have

$$d(h_y)_{\tilde{x}} = -df_{\tilde{x}} + id_E\,.$$

Using (1.5), it follows

$$\sup_{\tilde{x} \in B_r(0)} |||d(h_y)_{\tilde{x}}||| \leq \frac{1}{2}\,.$$

Applying the Control Growth Lemma 1.8, this implies (note that $B_r(0)$ is convex)

$$\|h_y(x) - h_y(x')\| \leq \frac{1}{2}\,\|x - x'\|\,, \tag{1.6}$$

showing that the map h_y is a contraction from $B_r(0)$ into $B_r(0)$. From the Banach fixed-point theorem, we deduce the existence of a unique $x \in B_r(0)$ such that $h_y(x) = x$. In other words, we proved that for all $y \in B_{r/2}(0)$ there exists a unique $x \in B_r(0)$ such that $f(x) = y$. Moreover, we define a new map $f^{-1}(y) := x$.

In the next step, we show that the map f^{-1} is continuous in $B_{r/2}(0)$ using (1.6). In fact, for all y, $\tilde{y} \in B_{r/2}(0)$, we have

$$\begin{aligned}\|f^{-1}(y) - f^{-1}(\tilde{y})\| &= \|x - \tilde{x}\| \\ &\leq \|x - f(x) - \tilde{x} + f(\tilde{x})\| + \|f(x) - f(\tilde{x})\| \\ &\leq \frac{1}{2}\|x - \tilde{x}\| + \|y - \tilde{y}\| \\ &= \frac{1}{2}\|f^{-1}(y) - f^{-1}(\tilde{y})\| + \|y - \tilde{y}\| \,.\end{aligned}$$

This implies

$$\|f^{-1}(y) - f^{-1}(\tilde{y})\| \leq 2\|y - \tilde{y}\| \,.$$

Hence the map f^{-1} is Lipschitz in $B_{r/2}(0)$ and in particular continuous.

Considering only the case $k = 1$, it remains to show that f^{-1} is a C^1-map in $B_{r/2}(0)$. For this purpose, we first have to show that for arbitrary $y \in B_{r/2}(0)$ the differential $d(f^{-1})_y$ exists. As expected, we will actually see that $d(f^{-1})_y$ equals $(df_x)^{-1}$, for $x = f^{-1}(y)$, whose existence follows from (1.5) and the so-called Neumann-serie, and its continuity from the open mapping theorem. Then we are done, since Exercise 1.16 implies the continuity of the map $x \longmapsto \mathrm{Inv} \circ df_x = (df_x)^{-1}$ as composition of continuous maps (recall that $f \in C^1(U, F)$ by assumption and see also Definition 1.2).

Denoting small perturbations of $f^{-1}(y) = x$ by $f^{-1}(y + w) = x + v$, we see that

$$\begin{aligned}&f^{-1}(y + w) - f^{-1}(y) - (df_x)^{-1} \cdot w \\ &= (x + v) - x - (df_x)^{-1} \cdot w \\ &= (df_x)^{-1} \cdot (df_x \cdot v - w) \\ &= -(df_x)^{-1} \cdot \big[f(x + v) - f(x) - df_x \cdot v\big] \,. \qquad (1.7)\end{aligned}$$

Since f is a C^1-map by assumption, the expression in brackets is of order $o(\|v\|)$. Moreover, the map $(df_x)^{-1}$ is uniformly bounded on $B_r(0)$. Indeed, we can write

$$\begin{aligned}|||(df_x)^{-1}||| &= |||(df_x)^{-1} - id + id||| \\ &\leq |||(df_x)^{-1} - id||| + |||id||| \\ &\leq |||(df_x)^{-1}(id - df_x)||| + |||id||| \,.\end{aligned}$$

Using a result of functional analysis and again Equation (1.5), we obtain

$$\begin{aligned}|||(df_x)^{-1}||| &\leq |||(df_x)^{-1}|||\ |||(id - df_x)||| + |||id||| \\ &\leq |||(df_x)^{-1}|||\ \frac{1}{2} + |||id||| \,.\end{aligned}$$

Hence,

$$|||(df_x)^{-1}||| \leq 2|||id||| < +\infty$$

and the right-hand side of (1.7) is thus of order $o(\|v\|)$.

We have also that

$$\|v\| = \|x + v - x\| = \|f^{-1}(y + w) - f^{-1}(y)\| \leq 2\|w\| ,$$

implying that $o(\|v\|) = o(\|w\|)$. Recalling (1.7), we arrive at

$$\|f^{-1}(y + w) - f^{-1}(y) - (df_x)^{-1} \cdot w\| = o(\|w\|) .$$

By Definition 1.1 this implies that f^{-1} is differentiable at y, for all $y \in B_{r/2}(0)$, and

$$d(f^{-1})_y = \left(df_{x=f^{-1}(y)}\right)^{-1} .$$

□

Application

We go further with a list of important definitions:

Definition 1.11. *Let E, F be two Banach spaces, $U \subset E$ open and let $f : U \longrightarrow F$ a C^k-map with $k \geq 1$.*

*(i) Assume there exists $V =\subset F$ open such that $f^{-1} : V \longrightarrow U$ exists and $f^{-1} \in C^k(V, U)$. Then f is called a C^k-**diffeomorphism** between U and V. Note that in the case of $k = 0$ the map f is called a **homeomorphism**.*

*(ii) Assume that for $a \in U$ the differential df_a is invertible. Then f is called a **local** C^k-**diffeomorphism** about a.*

*(iii) Assume that for $a \in U$ the differential df_a is injective. Then f is called a **local** C^k-**immersion** about a.*

*(iv) Assume that for $a \in U$ the differential df_a is surjective. Then f is called a **local** C^k-**submersion** about a.*

*(v) Assume now that $E = \mathbb{R}^n$ and $F = \mathbb{R}^m$. Let $y \in \mathbb{R}^m$ and assume that for every $a \in f^{-1}(y)$ the differential df_a has maximal rank, i.e., for $n \geq m$ rank $df_a = m$ and for $n \leq m$ rank $df_a = n$. Then $y \in \mathbb{R}^m$ is called a **regular point** for f.*

Example 1.12 (Polar coordinates). Consider the following map:

$$\begin{aligned} f : \mathbb{R}_+ \setminus \{0\} \times \mathbb{R} &\longrightarrow \mathbb{R}^2 \setminus \{0, 0\} , \\ (\rho, \theta) &\longmapsto f(\rho, \theta) := (\rho \cos\theta, \rho \sin\theta) . \end{aligned}$$

It is a *local* diffeomorphism, since the differential

$$df_{(\rho,\theta)} = \begin{pmatrix} \cos\theta & -\rho \sin\theta \\ \sin\theta & \rho \cos\theta \end{pmatrix}$$

with $\det(df_{(\rho,\theta)}) = \rho > 0$, is invertible. On the other hand, $f(\rho, \theta) = f(\rho, \theta + 2\pi)$ implies that the map f is not injective; therefore not a *global* diffeomorphism.

Without proof we state the following

Theorem 1.13. *Two open sets $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ are locally diffeomorphic only if $m = n$. The same holds for homeomorphic.*

The next theorem due to Cauchy (1839) is an important consequence of the Local Inversion Theorem 1.10.

Theorem 1.14 (Implicit Function Theorem). *Let E, F, G be three Banach spaces and $\varphi \in C^k(U \times V, G)$, where $U \subset E$ open and $V \subset F$ open. Assume that there exists $(a,b) \in U \times V$ such that the differential $(d\varphi^a)_b$ of $\varphi^a(y) := \varphi(a,y)$ at b is an isomorphism between F and G. Similarly, we define $\varphi^b(x) := \varphi(x,b)$. Then there exist $U' \subset U$ open with $a \in U'$, $W \subset G$ open with $\varphi(a,b) \in W$ and a unique map $\psi \in C^k(U' \times W, V)$ such that*

$$z = \varphi\big(x, \psi(x,z)\big) . \tag{1.8}$$

for all $(x,z) \in U' \times W$. Moreover, the following holds for the differential of ψ:

$$\begin{aligned} d\psi_{(x,z)} \cdot h &= (d\varphi^x)^{-1}_{\psi(x,z)} \cdot \big((d\varphi^{\psi(x,z)})_x \cdot h\big) \\ d\psi_{(x,z)} \cdot l &= -(d\varphi^x)^{-1}_{\psi(x,z)} \cdot l\,. \end{aligned} \tag{1.9}$$

where $(h,l) \in U' \times W$.

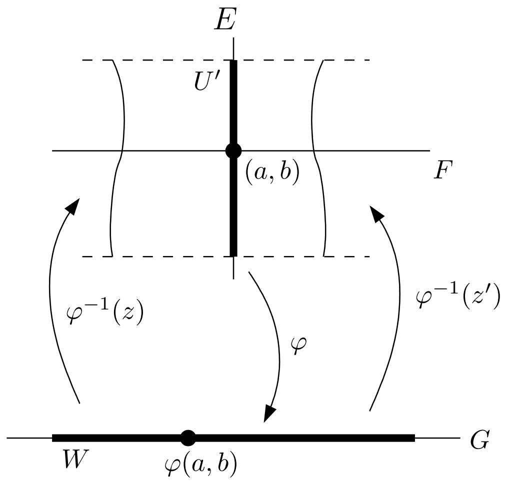

Fig. 1.1. For all $z \in W$ the map $\varphi^{-1}(z) = \big(x, \psi(x,z)\big)$ is a C^k-graph over U'.

Proof. For all $(x, y) \in U \times V$, we define the C^k-map

$$\begin{aligned} f : U \times V &\longrightarrow U \times G\,, \\ (x, y) &\longmapsto f(x, y) := \big(x, \varphi(x, y)\big)\,, \end{aligned}$$

and show that for $(a, b) \in U \times V$ the differential $df_{(a,b)}$ is an isomorphism. – For arbitrary $(h, k) \in U \times V$, we have[1]

$$df_{(a,b)} \cdot (h, k) = \left(h, \frac{\partial \varphi}{\partial x}(a, b) \cdot h + \frac{\partial \varphi}{\partial y}(a, b) \cdot k\right)\,,$$

where we write $(d\varphi^a)_b = \frac{\partial \varphi}{\partial y}(a, b)$ and $(d\varphi^b)_a = \frac{\partial \varphi}{\partial x}(a, b)$. In the following, we write the last equation as

$$df_{(a,b)} \cdot (h, k) = (h, \partial_x\varphi(a, b) \cdot h + \partial_y\varphi(a, b) \cdot k)\,.$$

Now, we assume that $df_{(a,b)} \cdot (h, k) = (0, 0)$. This clearly implies that $h = 0$ and $\partial_y\varphi \cdot k = 0$. Since by assumption $\partial_y\varphi$ is invertible, we get that k also vanishes. Hence $df_{(a,b)}$ is injective. In order to show that $df_{(a,b)}$ is surjective, we want to find for all $(v, w) \in U \times G$, pairs $(h, k) \in U \times V$ such that $df_{(a,b)} \cdot (h, k) = (v, w)$. It is easy to see that this implies that $h = v$ and $w = \partial_x\varphi \cdot h + \partial_y\varphi \cdot k$. By assumption the last equation has a solution for k, namely $k = (\partial_y\varphi)^{-1}(w - \partial_x\varphi \cdot h)$. Hence $df_{(a,b)}$ is also surjective. In summary, the differential $df_{(a,b)}$ is an isomorphism.

Next, we can apply the Local Inversion Theorem 1.10 showing that there exist $U' \subset U$ open with $a \in U'$ and $W \subset G$ open with $\varphi(a, b) \in W$ such that the map f has locally a C^k-inverse:

$$\begin{aligned} f^{-1} : U' \times W &\longrightarrow U \times V\,, \\ (x, z) &\longmapsto f^{-1}(x, z) =: \big(x, \psi(x, z)\big)\,. \end{aligned}$$

Thus $\psi : U' \times W \longrightarrow V$ is a C^k-map, and moreover

$$(x, z) = f\big(f^{-1}(x, z)\big) = f\big(x, \psi(x, z)\big) = \Big(x, \varphi\big(x, \psi(x, z)\big)\Big)\,,$$

showing (1.8).

With the notations introduced in the proof, we deduce by taking the differential of (1.8) that

$$id_z = \partial_y\varphi\big(x, \psi(x, z)\big) \cdot \partial_z\psi(x, z)\,,$$

and

$$0 = \partial_x\varphi\big(x, \psi(x, z)\big) + \partial_y\varphi\big(x, \psi(x, z)\big) \cdot \partial_x\psi(x, z)\,,$$

which is just (1.9). □

[1] We calculate exactly like in the case of Euclidean space (see (1.1)).

Application

In the local inversion theorem, the linear tangent map was assumed to be an isomorphism in one point. The next theorem treats the case of a tangent map being only surjective.

Theorem 1.15 (Submersion Theorem). *Let E, F Banach spaces, $U \subset E$ open, and let $\varphi \in C^k(U,F)$ with $k \geq 1$. Assume that there exist $a \in U$ and a closed subvector space E_1 of E such that $d\varphi_a$ is an isomorphism between E_1 and F. Moreover, assume also that $E = E_1 \oplus ker\, d\varphi_a$. Then there exist $U' \subset U$ open containing a, $W \subset F$ open containing $\varphi(a)$ and $\tilde{U} \subset ker\, d\varphi_a$ open containing 0 such that the map*

$$\begin{aligned} g : U' &\longrightarrow W \times ker\, d\varphi_a\,, \\ x &\longmapsto \big(\varphi(x), \pi(x-a)\big)\,, \end{aligned}$$

is a C^k-diffeomorphism from U' onto $g(U')$, where $\pi := E_1 \oplus ker\, d\varphi_a \longrightarrow ker\, d\varphi_a$ denotes the projection.

Remark. From Definition 1.11, we see that φ is a local C^k-submersion about a. Moreover, note that in the *finite* dimensional case (or in the case of an Hilbert space), it is not anymore necessary to assume that $E = E_1 \oplus \ker d\varphi_a$, since the subvector space E_1 always exists as orthogonal complement of $\ker d\varphi_a$ (with respect to an arbitrary scalar product).

Proof. We want to apply the Local Inversion Theorem 1.10. – For this purpose, we introduce the auxiliary map

$$\begin{aligned} g : U &\longrightarrow F \times \ker d\varphi_a\,, \\ x &\longmapsto g(x) := \big(\varphi(x), \pi(x-a)\big)\,. \end{aligned}$$

The differential of g at $a \in U$ is then given by

$$dg_a \cdot h = d\varphi_a \cdot h + \pi(h) \in F \times \ker d\varphi_a\,, \qquad h \in E\,,$$

which is an isomorphism by assumption. Hence, we can apply the local inversion theorem to get the existence of $U' \subset U$ open containing a, $W \subset F$ open containing $\varphi(a)$ and $\tilde{U} \subset \ker d\varphi_a$ open containing 0 such that g is a C^k-diffeomorphism between U' and $W \times \tilde{U}$. □

Remark. Although the three important theorems in this section are formulated in the most general case, they will often only be used in finite dimensional Euclidean space $\mathbb{R}^n$ in the next sections. Moreover, we emphasize that the three theorems are obtained from assumptions on the differential in only *one* point.

*   *   *

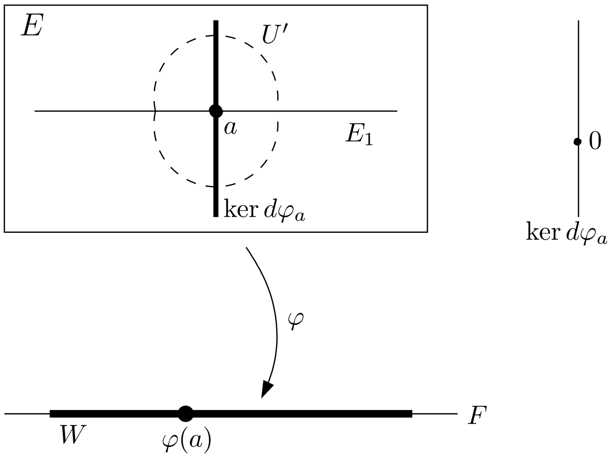

Fig. 1.2. Submersion theorem.

Exercises.

Exercise 1.16. Let $\mathrm{Aut}(E)$ denote the set of *automorphisms* of a Banach space E, i.e., the set of linear, continuous and bijective maps from E into E (see Definition 1.9). Show that the map

$$\begin{aligned} \mathrm{Inv} : (\mathcal{L}(E,E), |||\cdot|||) &\longrightarrow (\mathcal{L}(E,E), |||\cdot|||)\,, \\ f &\longmapsto \mathrm{Inv}(f) := f^{-1}\,, \end{aligned}$$

is continuous for $f \in \mathrm{Aut}(E)$.

1.2 The Flow of a Vector Field on $\mathbb{R}^n$

Definition 1.17. *Let U be an open subset of $\mathbb{R}^n$. Then the map $X : U \longrightarrow \mathbb{R}^n$ is called a* ***vector field*** *on U. If, in addition, the map X is C^k, then X is called a C^k-**vector field**.*

Definition 1.18. *Let X be a C^1-vector field on $U \subset \mathbb{R}^n$ open. A C^1-curve $\gamma : I \subset \mathbb{R} \longrightarrow U$, solving for all $t \in I$ the equation*

$$\dot{\gamma}(t) = X\big(\gamma(t)\big)\,,$$

is called an ***integral curve*** *of the vector field X.*

Definition 1.19. *Let X be a C^1-vector field on $U \subset \mathbb{R}^n$ open. A map*

$$\Gamma : I \times U \longrightarrow \mathbb{R}^n\,,$$

solving for all $(t,x) \in I \times U$ the equations

$$\frac{\partial \Gamma}{\partial t}(t,x) = X\big(\Gamma(t,x)\big)\,, \qquad \Gamma(0,x) = x\,,$$

is called the ***(local) flow*** *of* X*.*

Note that by $\frac{\partial}{\partial t}$ we mean usual partial derivative with respect to t. – The last definition can be reformulated in the following way: For all $x \in U$ fixed, the map $t \in I \longmapsto \Gamma(t,x)$ is an integral curve of X with initial condition x. – Using the Implicit Function Theorem 1.14, we now want to prove an important local existence result.

Theorem 1.20 (Local Existence of a Flow). *Let X be a C^k-vector field on $U \subset \mathbb{R}^n$ and let $a \in U$. Then there exist $U' \subset U$ open with $a \in U'$ and a C^k-map $\Gamma : [-T,T] \times U' \longrightarrow \mathbb{R}^n$ for some $T > 0$ such that for every $(t,x) \in [-T,T] \times U'$, we have*

$$\frac{\partial \Gamma}{\partial t}(t,x) = X\big(\Gamma(t,x)\big)\,, \qquad \Gamma(0,x) = x\,. \tag{1.10}$$

Integrating (1.10) gives

$$\Gamma(t,x) = x + \int_0^t X\big(\Gamma(s,x)\big)ds\,, \tag{1.11}$$

where x is a point in the neighborhood of a and t a time in the neighborhood of 0. We wish to apply the Implicit Function Theorem 1.14, and define therefore the map

$$\tilde{\varphi}\big(t,x,\Gamma(t,x)\big) = -\Gamma(t,x) + x + \int_0^t X\big(\Gamma(s,x)\big)ds\,.$$

Then (1.11) translates to

$$\tilde{\varphi}\big(t,x,\Gamma(t,x)\big) = 0\,. \tag{1.12}$$

Since Γ depends also on (x,t), it is not possible to apply directly the implicit function theorem. However, this idea will lead to the correct proof. But, in addition, we have to introduce a parameter $u \in [0,1]$.

Proof. First we introduce the space $F = C^0([0,1],\mathbb{R}^n)$ together with the L^∞-norm, i.e., $\|\alpha\|_\infty = \sup_{u\in[0,1]} |\alpha(u)|$ for $\alpha \in F$. It is not difficult to see that $(F, \|\cdot\|_\infty)$ is a Banach space[2]. Moreover, we define the map $\varphi : (-1,1) \times U \times F \longrightarrow F$ by

[2] Indeed, let $(\alpha_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in F, meaning that for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $\|\alpha_n - \alpha_m\|_\infty \le \varepsilon$, $\forall n,m \ge N$. This implies that $(\alpha_n(u))_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$ for all $u \in [0,1]$, since

$$|\alpha_n(u) - \alpha_m(u)| \le \|\alpha_n - \alpha_m\|_\infty \le \varepsilon\,, \qquad \forall n,m \ge N\,. \tag{1.13}$$

$$\varphi(t,x,\alpha)(u) := -\alpha(u) + t\int_0^u X\big(\alpha(\sigma)+x\big)\,d\sigma\,, \tag{1.14}$$

where $\sigma, u \in [0,1]$. Note that $(-1,1)\times U$ is an open set of $\mathbb{R}^{n+1}$ which is Banach, and therefore $(-1,1)\times U\times F$ is also an open subset of a Banach space since F was shown to be a Banach space.

Let $a \in U$ and, using the same notations as in the previous section, we compute $\partial_\alpha\varphi_{(0,a,0)}$. (As an exercise, one shows that $\varphi \in C^k\big((-1,1)\times U\times F, F\big)$ and hence the differential exists.) For $h \in F$, we have by definition $\varphi(0,a,h)(u) = -h(u)+0 = -h(u)$ and hence $\partial_\alpha\varphi_{(0,a,0)} = -id_F$, showing that it is an isomorphism.

We can then apply the Implicit Function Theorem 1.14 to the map φ at the point $(0,a,0)\in(-1,1)\times U\times F$ to get the existence of $(-T,T)\times U' \subset (-1,1)\times U$ and $W\subset F$ open containing $\varphi(0,a,0)$; moreover, the existence of a C^k-map

$$\psi : (-T,T)\times U'\times W \longrightarrow F\,,$$

such that for all $(t,x)\in(-T,T)\times U'$ and for all $z\in W$ the following equation holds:

$$\varphi\big(t,x,\psi(t,x,z)\big) = z\,.$$

Taking $z=0$, we thus get

$$\varphi\big(t,x,\psi(t,x,0)\big) = 0\,, \tag{1.15}$$

equivalently, using (1.14),

$$-\psi(t,x,0)(u) + t\int_0^u X\big(\psi(t,x,0)(\sigma)+x\big)d\sigma = 0\,,$$

for every $u\in[0,1]$.

In a next step, we want to establish a relation between ψ and Γ such that (1.15) implies that (1.12) holds. We claim that

$$\Gamma(t,x) = \psi\big(t,x,0\big)(1) + x \tag{1.16}$$

is the desired relation.

By the completeness of $\mathbb{R}$, we deduce that the Cauchy sequence $(\alpha_n(u))_{n\in\mathbb{N}}$ converges to a limit, which we denote by $\alpha(u)$. It remains to show the continuity of the resulting function $\alpha(u)$. For this purpose, we take the limit $m\longrightarrow\infty$ in (1.13) to obtain

$$|\alpha_n(u)-\alpha(u)|\le\varepsilon\,,\qquad \forall n\ge N\,,$$

where N depends only on $\varepsilon>0$ and *not* on $u\in[0,1]$. Hence, it follows

$$\lim_{n\longrightarrow\infty}\|\alpha_n-\alpha\|_\infty = 0\,.$$

Since α is the limit of a uniformly convergent sequence $(\alpha_n)_{n\in\mathbb{N}}$, it is indeed continuous and therefore F is Banach.

We first introduce the map

$$\begin{aligned} \tau_\lambda : F &\longrightarrow F\,, \\ \alpha(u) &\longmapsto \tau_\lambda \alpha(u) := \alpha(\lambda u)\,. \end{aligned}$$

For $\lambda \in [0,1]$ this map is well-defined. We calculate, for $\lambda\sigma = s$,

$$\begin{aligned} \varphi\big(\lambda t, x, \tau_\lambda \alpha\big)(u) &= -\alpha(\lambda u) + \lambda t \int_0^u X\big(\alpha(\lambda\sigma) + x\big) d\sigma \\ &= -\alpha(\lambda u) + t \int_0^{\lambda u} X\big(\alpha(s) + x\big) ds \\ &= \varphi\big(t, x, \alpha\big)(\lambda u) = \tau_\lambda \varphi\big(t, x, \alpha\big)(u)\,. \end{aligned}$$

Applying this formula to the continuous map $\alpha = \psi(t,x,0)$ leads to

$$\varphi\big(\lambda t, x, \tau_\lambda \psi(t,x,0)\big) = \tau_\lambda \varphi\big(t, x, \psi(t,x,0)\big) = \tau_\lambda 0 = 0\,, \tag{1.17}$$

where we used (1.15). Using again (1.15), the last equation can also be written as

$$\varphi\big(\lambda t, x, \psi(\lambda t, x, 0)\big) = 0\,. \tag{1.18}$$

Putting the results (1.17) and (1.18) together and because of the local uniqueness of the implicit function theorem, we arrive at the formula

$$\tau_\lambda \psi(t,x,0) = \psi(\lambda t, x, 0)\,. \tag{1.19}$$

Now, we are ready to show that (1.16) is the correct relation between Γ and ψ. Equation (1.19) implies

$$\begin{aligned} 0 &= \tau_\lambda \varphi\big(t, x, \psi(t,x,0)\big)(u) \\ &= -\psi(t,x,0)(\lambda u) + t \int_0^{\lambda u} X\big(\psi(t,x,0)(s) + x\big) ds \\ &= -\psi(\lambda t,x,0)(u) + t \int_0^{\lambda u} X\big(\psi(t,x,0)(s) + x\big) ds\,. \end{aligned}$$

For $u = 1$ and $ts = \sigma$, we finally obtain

$$\begin{aligned} 0 \quad &= \quad -\psi(\lambda t, x, 0)(1) + t \int_0^{\lambda} X\big(\psi(t,x,0)(s) + x\big) ds \\ &\overset{(1.19)}{=} \quad -\Gamma(t\lambda, x) + x + t \int_0^{\lambda} X\big(\psi(ts,x,0)(1) + x\big) ds \\ &= \quad -\Gamma(t\lambda, x) + x + t \int_0^{\lambda} X\big(\Gamma(ts,x) ds \\ &= \quad -\Gamma(t\lambda, x) + x + \int_0^{t\lambda} X\big(\Gamma(\sigma,x) d\sigma\,, \end{aligned}$$

which is precisely (1.12). This proves the existence of a local flow. $\square$

Theorem 1.21 (Uniqueness of the (Local) Flow). *Let X be a C^1-vector field on $U \subset \mathbb{R}^n$ open and let γ_1 and γ_2 denote two integral curves of X on $[0, T_1]$, respectively $[0, T_2]$, with $0 < T_1, T_2 \in \mathbb{R}$. If the integral curves have the same initial conditions, i.e., $\gamma_1(0) = \gamma_2(0)$, then they agree:*

$$\gamma_1(t) = \gamma_2(t)\,, \qquad \forall t \in [0, T_1] \cap [0, T_2]\,.$$

Proof. We take $T > 0$ such that $[0, T_1] \cap [0, T_2] = [0, T]$ and define

$$C := \big\{ t \in [0, T] \,:\, \gamma_1(t) = \gamma_2(t) \big\}\,.$$

By assumption $0 \in C$ and hence $C \subset [0, T]$ is non empty. Using a standard topological argument[3], it suffices to prove that C is both open and closed in $[0, T]$ to establish the uniqueness of the flow, i.e., $C = [0, T]$.

The set C is closed: Since the integral curves γ_1 and γ_2 are continuous, it is clear that C is closed.

The set C is open: Let $t_0 \in C$ with $x_0 = \gamma_1(t_0) = \gamma_2(t_0) \subset U$ and let $r > 0$ such that $\overline{B_r(x_0)} \subset U$. (The new initial time $t_0 \in C$ can be obtained by translation of $[0, T]$.) Since by assumption X is a C^1-vector field, all the partial derivatives of X are bounded on the compact set $\overline{B_r(x_0)}$. Thus, by the Control Growth Lemma 1.8, the vector field X is also Lipschitz on this compact set, i.e., there exists $k > 0$ such that

$$\|X(x) - X(\tilde{x})\| \le k\, \|x - \tilde{x}\|\,,$$

for all $x, \tilde{x} \in \overline{B_r(x_0)}$. On the other hand, since both integral curves are continuous, there exists $\delta > 0$ such that

$$\gamma_i\big((t_0 - \delta, t_0 + \delta)\big) \subset B_r(x_0)\,, \qquad i = 1, 2\,.$$

Next, we want to show that $(t_0 - \delta, t_0 + \delta) \subset C$. – For $t \in (t_0 - \delta, t_0 + \delta)$ the following estimate holds

$$\begin{aligned} \frac{d}{dt}\big\|\gamma_1(t) - \gamma_2(t)\big\|^2 &\le 2\, \|\gamma_1(t) - \gamma_2(t)\|\; \|\dot{\gamma}_1(t) - \dot{\gamma}_2(t)\| \\ &\le 2\, \|X\big(\gamma_1(t)\big) - X\big(\gamma_2(t)\big)\|\; \|\gamma_1(t) - \gamma_2(t)\| \\ &\le 2k\, \|\gamma_1(t) - \gamma_2(t)\|^2\,, \end{aligned} \tag{1.20}$$

where we also used the Definition 1.18 of an integral curve.

Now, we introduce the real valued function $f(t) = \|\gamma_1(t) - \gamma_2(t)\|^2 \ge 0$; then (1.20) translates to

$$\frac{d}{dt} f(t) \le 2k\, f(t)\,, \qquad f(t_0) = 0\,,$$

[3] If B is connected and $A \subset B$ open and closed, then A is empty or $A = B$. The same argument was already used in the proof of Control Growth Lemma 1.8.

implying that

$$\frac{d}{dt}\big(e^{-2kt}f(t)\big) \leq 0\,, \qquad \forall t \in [t_0 + \delta)\,.$$

Thus $f \equiv 0$ and $\gamma_1(t) = \gamma_2(t)$ on $[t_0 + \delta)$. Reversing time, we arrive at $(t_0 - \delta, t_0 + \delta) \subset C$. □

In words, this theorem shows that certainly no bifurcation of the integral curves will appear. – The next proposition indicates that passing along an integral to a time t and then to a time $T - t$ gives exactly the same than passing directly to a time T.

Proposition 1.22. *Let $\Gamma : [0,T] \times U \longrightarrow \mathbb{R}^n$ be a flow of a C^1-vector field X. Then for all $t \in (0,T)$ the following holds:*

$$\Gamma(T,x) = \Gamma\big(T-t, \Gamma(t,x)\big)\,, \qquad \forall x \in U\,. \tag{1.21}$$

Note that, writing $\Gamma_t(\cdot)$ for the map $\Gamma(t,\cdot) : U \longrightarrow \mathbb{R}^n$, (1.21) can be written as

$$\Gamma_T = \Gamma_{(T-t)+t} = \Gamma_{(T-t)} \circ \Gamma_t\,.$$

Especially, we have that $id_U = \Gamma_{-t} \circ \Gamma_t$ and hence we get $(\Gamma_t)^{-1} = \Gamma_{-t}$ for the inverse flow. Combining this observation with the fact the Γ_t is C^1, we get the following

Corollary 1.23. *Let $\Gamma : [0,T] \times U \longrightarrow \mathbb{R}^n$ be a flow of a C^1-vector field X. Then for all $t \in (0,T)$ the map*

$$\begin{aligned} \Gamma_t : U &\longrightarrow \mathbb{R}^n\,, \\ x &\longmapsto \Gamma_t(x) = \Gamma(t,x) \end{aligned}$$

is a (local) C^1-diffeomorphism onto $\Gamma_t(U)$.

We now introduce a very important concept in differential geometry, which will be used in the following at various places in a more general context.

Definition 1.24. *Let $U \subset \mathbb{R}^n$ open and $\varphi : U \longrightarrow V \subset \mathbb{R}^n$ a C^1-diffeomorphism. If the diffeomorphism φ is written as $\varphi(x) = \big(y_1(x), \ldots, y_n(x)\big)$ for $x \in U$, then the coordinates of $\big(y_1(x), \ldots, y_n(x)\big)$ with respect to the canonical basis $\{e_i\}_{1 \leq i \leq n}$ of $\mathbb{R}^n$ are called the* ***coordinates*** *on U for x. Moreover, the functions $(y_1, \ldots, y_n)$ are called the* ***coordinate functions*** *of φ for U and such a diffeomorphism φ is called a* ***chart*** *or* ***coordinate system*** *on U. And we define*

$$\frac{\partial}{\partial y_i}(x) := d(\varphi^{-1})_{(\varphi(x))} \cdot e_i\,, \qquad \forall i = 1, \ldots, n\,. \tag{1.22}$$

Remark. In the following, the same notation is often used for coordinate functions and coordinates.

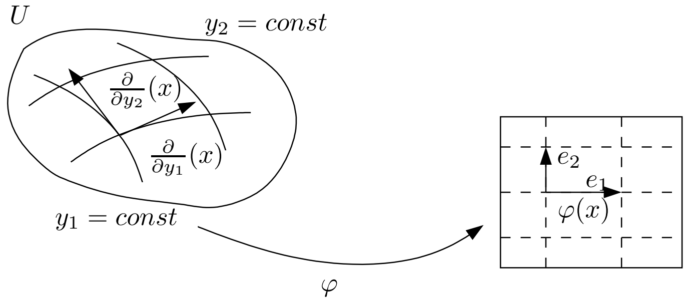

Fig. 1.3. Chart on Euclidean space.

Theorem 1.25 (Straightening Theorem). *Let X be a C^1-vector field on $a \in U \subset \mathbb{R}^n$ open such that $X(a) \neq 0$. Then there exist $W \subset U$ open containing a and a chart $\varphi = (y_1, y_2, \ldots, y_n)$ on W such that*

$$X(x) = \frac{\partial}{\partial y_1}(x)\,, \qquad \forall x \in W\,. \tag{1.23}$$

In other words, there exists a C^1-diffeomorphism $\varphi = (y_1, y_2, \ldots, y_n)$: $W \longrightarrow V \subset \mathbb{R}^n$ such that

$$d\varphi_x \cdot X(x) = e_1\,, \qquad \forall x \in W\,. \tag{1.24}$$

Proof. Let $\{e_i\}_{1\leq i\leq n}$ be the canonical basis of $\mathbb{R}^n$ with coordinate functions $(x_1, \ldots, x_n)$. We may assume that $a = 0$. After a linear transformation if necessary, we may also assume that $X(0) = e_1$.

By Theorem 1.20, there exists a neighborhood $\tilde{U} \subset U$ of 0 and $T > 0$ such that the (local) flow Γ_t of the vector field X exists in $(-T, T) \times \tilde{U}$. Consider now the map

$$\phi(x_1, x') = \Gamma_{x_1}\big((0, x')\big)\,, \tag{1.25}$$

where we write $x = (x_1, x') = (x_1, x_2, \ldots, x_n) \in \tilde{U}$ such that $|x_1| \leq T$.

Using Corollary 1.23 and again Theorem 1.20, we deduce that all partial derivatives of ϕ exist and are continuous. Hence, we conclude by Proposition 1.3 that $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^n)$.

In a next step, we compute the differential of ϕ at $(0,0)$. By Definition 1.19 of the flow, it follows

$$d\phi_{(0,0)} \cdot e_1 = \frac{\partial \phi}{\partial x_1}(0,0) = \frac{d}{dh}\bigg|_{h=0} \Gamma_h\big((0,0)\big) = X\big((0,0)\big) = e_1\,,$$

and, for $i = 2, \ldots, n$,

$$d\phi_{(0,0)} \cdot e_i = \frac{d}{dh}\bigg|_{h=0} \Gamma_0\big((0, h\, e_i)\big) = \frac{d}{dh}\bigg|_{h=0} h\, e_i = e_i\,.$$

Hence, the differential $d\phi_{(0,0)}$ is the identity and we can apply the Local Inversion Theorem 1.10, implying the existence of

$$V := \big\{(x_1,x') \in \mathbb{R}^n \ : \ (x_1,x') \in U' \subset \tilde{U} \ \text{ and } \ |x_1| \leq \tilde{T} \leq T\big\}\,,$$

and $W \subset \mathbb{R}^n$ neighborhood of $\phi\big((0,0)\big)$ such that ϕ is a local C^1-diffeomorphism from V into W.

Then, we define $\varphi = \phi^{-1}$. It remains to show that, for all $y \in W$ (this notation for the points in W is motivated by the fact that $\varphi = (y_1,\ldots,y_n)$ denote the coordinate functions on W),

$$d\varphi_y \cdot X(y) = e_1\,,$$

being equivalent to (see Corollary 1.5)

$$d\phi_{\varphi(y)} \cdot e_1 = X\big(\phi(x)\big)\,.$$

Using Proposition 1.22 and again the definition of the flow, we compute, for $x = (x_1,x') \in V$,

$$\begin{aligned} d\phi_{(x_1,x')} \cdot e_1 &= \left.\frac{d}{dh}\right|_{h=0} \Gamma_{x_1+h}\big((0,x')\big) = \left.\frac{d}{dh}\right|_{h=0} \Gamma_h\left(\Gamma_{x_1}\big((0,x')\big)\right) \\ &= \left.\frac{d}{dh}\right|_{h=0} \Gamma_h\big(\phi(x_1,x')\big) = X\big(\phi(x)\big)\,. \end{aligned}$$

This completes the proof of the straightening theorem. □

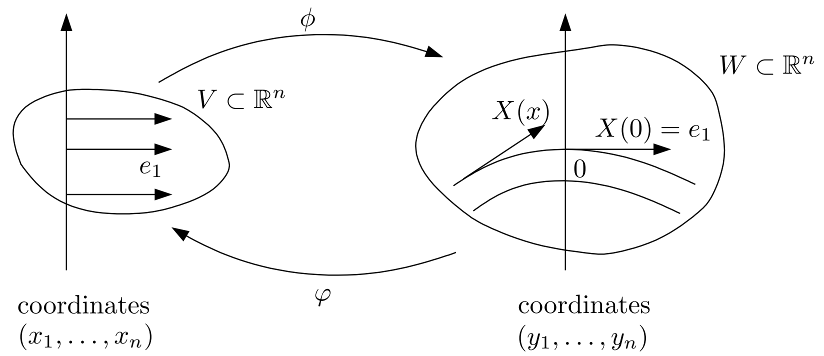

Fig. 1.4. Setting for the proof of the straightening theorem.

2 Differentiable Manifolds

2.1 Submanifolds of $\mathbb{R}^p$

Motivation

Definition 2.1. *A n-dimensional C^k-**submanifold** of $\mathbb{R}^p$, $n < p$, is a subset N^n of $\mathbb{R}^p$ such that for every $x \in N^n$ there exists $U \subset \mathbb{R}^p$ open containing x and a map $f : U \longrightarrow \mathbb{R}^p$ with the following properties:*

(i) $f \in C^k(U, \mathbb{R}^p)$,
(ii) the map f is a C^k-diffeomorphism from U into $f(U)$;
(iii) and $f(N^n \cap U) = \mathbb{R}^n \cap f(U)$.

*The map f is called a **straightening map** for N^n about x.*

Remark. We understand $\mathbb{R}^n$ as subset of $\mathbb{R}^p$ via the canonical inclusion

$$\begin{aligned} \iota : \mathbb{R}^n &\hookrightarrow \mathbb{R}^p\,, \\ (x_1, \ldots, x_n) &\longmapsto (x_1, \ldots, x_n, 0, \ldots, 0)\,. \end{aligned} \tag{2.1}$$

Thus, we can also write $\mathbb{R}^n \times \mathbb{R}^{p-n} = \mathbb{R}^p$. – Note also that, in the following, the upper indices for (sub)manifolds, subsets, etc. will always refer to their dimensions.

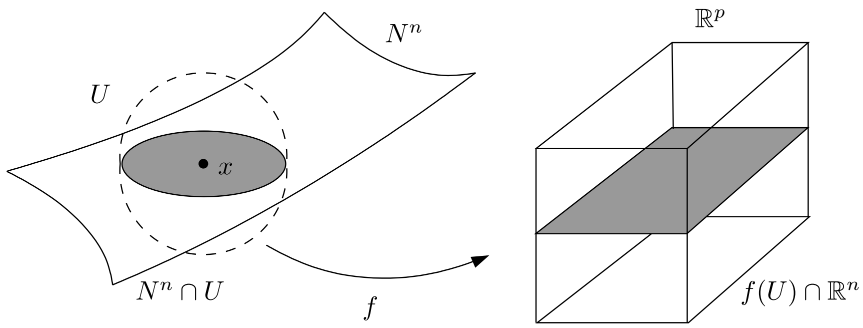

Fig. 2.1. Submanifold.

Proposition 2.2. *Let N^n be a subset of $\mathbb{R}^p$. Then the following assertions are equivalent:*

(i) The subset N^n is a n-dimensional C^k-submanifold of $\mathbb{R}^p$.
(ii) For all $x \in N^n$, there exists $U \subset \mathbb{R}^p$ open containing x and C^k-functions $f^1, \dots, f^{p-n} : U \longrightarrow \mathbb{R}$ such that $\{df_x^i\}_{i=1,\dots,(p-n)}$ is a free (linearly independent) family and

$$N^n \cap U = \bigcap_{i=1}^{p-n} \left(f^i\right)^{-1}(0)\,. \tag{2.2}$$

The functions f^i, $i = 1, \dots, p-n$, are called a ***family of constraints*** *to N^n about x.*
(iii) For all $x \in N^n$, there exists $U \subset \mathbb{R}^p$ open containing x and a C^k-submersion $g : U \longrightarrow \mathbb{R}^{p-n}$ such that $N^n \cap U = g^{-1}(0)$.

Remark. Let $p = 3$ and $n = 1$. Using (ii) we recover the well-known fact that *locally* a curve can be described as intersection of two surfaces in $\mathbb{R}^3$ given by f^1 and f^2. This is the local version of the characterization of a one-dimensional vector subspace in $\mathbb{R}^3$.

Proof. The equivalence of (ii) and (iii) is straightforward. – Let

$$(f^1, \dots, f^{p-n}) = g : U \longrightarrow \mathbb{R}^{p-n}\,.$$

By assumption, $\{df_x^i\}_{i=1,\dots,(p-n)}$ is a free family meaning that for every $x \in U$ the differential dg_x is surjective. Hence by Definition 1.11 the map g is a local submersion. Furthermore, we have

$$N^n \cap U = \bigcap_{i=1}^{p-n} \left(f^i\right)^{-1}(0) = g^{-1}(0)\,.$$

Now, we show that (iii) implies (i). – Let $x_0 \in N^n$ with $U \subset \mathbb{R}^n$ containing x_0 and $g : U \longrightarrow \mathbb{R}^{p-n}$ the C^k-submersion. Moreover, let $\pi : \mathbb{R}^p \longrightarrow \mathbb{R}^{p-n}$ denote the canonical projection. Then by the Submersion Theorem 1.15 there exist $U' \subset U$ and a local C^k-diffeomorphism $f : U' \longrightarrow \mathbb{R}^p$ such that

$$g(x) = \big(\pi \circ f\big)(x)\,, \qquad x \in U'\,.$$

The last equation gives, using (iii),

$$N^n \cap U' = g^{-1}(0) = \big(f^{-1} \circ \pi^{-1}\big)(0) = f^{-1}(\mathbb{R}^n) \cap U'\,,$$

showing that f is the straightening map for the submanifold N^n of $\mathbb{R}^p$.

It remains to show that (i) implies (iii). – Let $x_0 \in N^n$ and $f : U \longrightarrow \mathbb{R}^p$ the straightening map for N^n about $x_0 \in U$. By Definition 2.1, we have that

$N^n \cap U = f^{-1}(\mathbb{R}^n) \cap U$. Writing $f(x) = \big(f^1(x), \ldots, f^n(x), f^{n+1}(x), \ldots, f^p(x)\big)$, we define

$$\begin{aligned} g : U &\longrightarrow \mathbb{R}^{n-p}\,, \\ x &\longmapsto \big(f^{n+1}(x), \ldots, f^p(x)\big)\,. \end{aligned}$$

As a direct consequence, we note that $N^n \cap U = g^{-1}(0)$. Since f is a C^k-diffeomorphism, the matrix formed by the differentials $(df^i)_{x_0}$ is of maximal rank $p-n$ for all $i = n+1, \ldots, p$. Thus the map g gives the submersion of (iii). □

Example 2.3 (The Sphere S^{p-1}). We consider the sphere

$$S^{p-1} = \big\{x = (x_1, \ldots, x_p)\ :\ x_1^2 + \ldots + x_p^2 = 1\big\} \subset \mathbb{R}^p\,,$$

and the constraint

$$\begin{aligned} f^1 : \mathbb{R}^p &\longrightarrow \mathbb{R}\,, \\ x &\longmapsto x_1^2 + \ldots + x_p^2 - 1\,. \end{aligned} \tag{2.3}$$

Clearly $(f^1)^{-1}(0) = S^{p-1}$ and $f^1 \in C^\infty(\mathbb{R}^p, \mathbb{R})$. Moreover, for $x \in S^{p-1}$ the expression $df^1_x \cdot X = 2\sum_{i=1}^p x_i\, X_i$ with $X = (X_1, \ldots, X_p) \in \mathbb{R}^p \setminus \{0\}$ does not vanish and the differential of the constraint f^1 is therefore free. By Proposition 2.2 the sphere S^{p-1} is thus a $(p-1)$-dimensional C^∞-submanifold of $\mathbb{R}^p$.

Example 2.4 (The Orthogonal Group $O(k)$). Let $GL_k(\mathbb{R})$ denote the space of invertible $k \times k$ matrices identified with $\mathbb{R}^{k^2}$. The **orthogonal group** is defined by

$$O(k) = \{A \in GL_k(\mathbb{R})\ :\ A^T A = A\, A^T = \mathbb{1}\}\,,$$

where A^T denotes the transpose of the matrix A. We show that $O(k)$ is a C^∞-submanifold of $\mathbb{R}^{k^2}$ of dimension $\frac{k(k-1)}{2}$. – First, we define the C^∞-map

$$\begin{aligned} g : \mathbb{R}^{k^2} &\longrightarrow \mathbb{R}^{\frac{k(k+1)}{2}}\,, \\ A &\longmapsto A^T A - \mathbb{1}\,, \end{aligned}$$

where the space of symmetric matrices $S_k(\mathbb{R})$ is identified with $\mathbb{R}^{\frac{k(k+1)}{2}}$. For $A \in O(k)$ and $B \in GL_k(\mathbb{R})$ we have

$$dg_A \cdot B = B^T A + A^T B\,.$$

For $S \in S_k(\mathbb{R})$, i.e., $S^T = S$, we want to find $B \in GL_k(\mathbb{R})$ such that $dg_A \cdot B = S$. A short calculation gives $B = A\,S/2$. This shows that g is a submersion. Since

$$k^2 - \frac{k(k-1)}{2} = \frac{k(k+1)}{2}\,,$$

and also $O(k) = g^{-1}(0)$, Proposition 2.2 gives the result.

Example 2.5 (The Special Linear Group $SL(k,\mathbb{R})$). The **special linear group** is defined by

$$SL(k,\mathbb{R}) = \{A \in GL_k(\mathbb{R}) \ : \ \det A = 1\}\,.$$

It follows directly from Exercise 2.7 that the determinant map is a submersion. Moreover, we have that $SL(k,\mathbb{R}) = g^{-1}(0)$, where g is the map given by $A \longmapsto \det A - 1$. From Proposition 2.2 we then conclude that $SL(k,\mathbb{R})$ is a C^∞-submanifold of $\mathbb{R}^{k^2}$ of dimension k^2-1. – Note that the **special orthogonal group**

$$SO(k) = \{A \in O(k) \ : \ \det A = 1\} = O(k) \cap SL(k,\mathbb{R})$$

is also a C^∞-submanifold of $\mathbb{R}^{k^2}$ with $\dim SO(k) = \dim O(k)$.

Proposition 2.6. *Let N^n be a n-dimensional C^k-submanifold of $\mathbb{R}^p$.*

(i) If $\psi : \mathbb{R}^p \longrightarrow \mathbb{R}^p$ is a C^k-diffeomorphism, then $\psi(N^n)$ is again a n-dimensional C^k-submanifold of $\mathbb{R}^p$.
(ii) If for $q > p$ the canonical inclusion is denoted by

$$\begin{aligned} \iota : \mathbb{R}^p &\hookrightarrow \mathbb{R}^q\,, \\ (x_1,\dots,x_p) &\longmapsto (x_1,\dots,x_p,0,\dots,0)\,, \end{aligned}$$

then $\iota(N^n)$ is a C^k-submanifold of $\mathbb{R}^q$ of dimension n.
(iii) If M^m is a m-dimensional C^k-submanifold of $\mathbb{R}^q$ then $N^n \times M^m$ is a C^k-submanifold of $\mathbb{R}^p \times \mathbb{R}^q$ of dimension $n+m$.

Proof. The proof of (i) is straightforward. – Let f be the straightening map for N^n about x. Since ψ is by assumption a C^k-diffeomorphism it is clear that $f \circ \psi^{-1}$ gives a C^k-straightening map for $\psi(N^n)$ about $\psi(x)$.

Since N^n is by assumption a submanifold of $\mathbb{R}^p$, there exists by Proposition 2.2 a family of constraints $f^1,\dots,f^{p-n}$ about $\tilde{x} \in \mathbb{R}^p$ verifying (2.2). For $x = (x_1,\dots,x_q) \in U \times \mathbb{R}^{q-p}$ with $U \subset \mathbb{R}^p$ containing $\tilde{x}$, we then define the C^k-maps

$$f^i(x_1,\dots,x_p,x_{p+1},\dots,x_q) := f^i(x_1,\dots,x_p)\,, \qquad 1 \le i \le p-n\,,$$

and moreover

$$f^i(x_1,\dots,x_p,x_{p+1},\dots,x_q) := x_{i+n}\,, \qquad p-n+1 \le i \le q-n\,.$$

These two definitions obviously show that

$$\iota(N^n) \cap \big(U \times \mathbb{R}^{q-p}\big) = \bigcap_{i=1}^{q-n} \big(f^i\big)^{-1}(0)\,.$$

At the point $x_0 = \iota(\tilde{x}) = (\tilde{x}_1,\dots,\tilde{x}_p,0,\dots,0)$ the differential of the function

$$F := (f^1, \dots, f^p, f^{p+1}, \dots, f^{q-n}) : \mathbb{R}^q \longrightarrow \mathbb{R}^{q-p}$$

reads as

$$dF_{x_0} = \begin{pmatrix} A & 0 \\ 0 & \mathbb{1} \end{pmatrix} .$$

Since $\{df^i_{x_0}\}_{i=1,\dots,(p-n)}$ is free by assumption the matrix A has maximal rank $p-n$. We deduce that dF_{x_0} has also maximal rank. In summary, the family $\{df^i_{x_0}\}_{i=1,\dots,(q-n)}$ is a free family of constraints for $\iota(N^n)$ about x_0 and (ii) follows from Proposition 2.2.

For (iii), we note that the constraints for N^n and M^m can be combined to $q+p-(n+m)$ independent constraints for $N^n \times M^m$. Then again Proposition 2.2 gives the result. $\square$

From the last proposition we get many submanifolds from the one we already now. – For example, we showed that the "circle" S^1 is a submanifold of $\mathbb{R}^2$ and as an application of the second part of the proposition it is also a submanifold of $\mathbb{R}^3$. Moreover, the torus $S^1 \times S^1 \subset \mathbb{R}^2 \times \mathbb{R}^2$ is a two dimensional submanifold by the third part of the proposition.

* * *

Exercises.

Exercise 2.7. Let $GL_k(\mathbb{R})$ denote the space of invertible $k \times k$ matrices and consider the map

$$\begin{aligned} \det : GL_k(\mathbb{R}) &\longrightarrow \mathbb{R} \setminus \{0\} , \\ A &\longmapsto \det A . \end{aligned}$$

Show that its differential, for $A, B \in GL_k(\mathbb{R})$, is given by

$$d(\det)_A \cdot B = (\det A)\ \mathrm{Tr}(A^{-1}B) .$$

2.2 Differentiability on Submanifolds of $\mathbb{R}^p$

We want to give a meaning to a C^k-map, $k \geq 1$, from a C^k-submanifold N^n of $\mathbb{R}^p$ into $\mathbb{R}$ or into another C^k-submanifold M^m. This will be essentially done by the notion of a chart for a submanifold. – The following definition can be seen as generalization of Definition 1.24.

Definition 2.8. *Let N be a n-dimensional C^k-submanifold of $\mathbb{R}^p$ with C^k-straightening map f and $U \subset \mathbb{R}^p$ open containing $x \in N^n$. Moreover, let*

$$\begin{aligned} \pi : \mathbb{R}^p &\longrightarrow \mathbb{R}^n \\ (x_1, \dots, x_n, x_{n+1}, \dots, x_p) &\longmapsto (x_1, \dots, x_n) \end{aligned}$$

denote the canonical projection and $V \subset \mathbb{R}^n$ open. A ***chart*** *or* ***coordinate system*** *for N^n about x is a homeomorphism*

$$\varphi : N^n \cap U \longrightarrow V$$

such that the composition $\pi \circ f \circ \varphi^{-1}$ is a C^k- diffeomorphism from $\varphi(N^n \cap U) \subset \mathbb{R}^n$ into an open subset of $\mathbb{R}^n$. – If the chart φ is written as $\varphi(x) = \big(y_1(x), \ldots, y_n(x)\big)$ for $x \in N^n$, then the canonical coordinates of $\big(y_1(x), \ldots, y_n(x)\big)$ in $\mathbb{R}^n$ are called the ***coordinates*** *on the submanifold N^n for x; moreover $(y_1, \ldots, y_n)$ are called the* ***coordinate functions*** *of the chart φ for U.*

Remark. In the following, there will often be no difference in the notations for coordinates and coordinate functions.

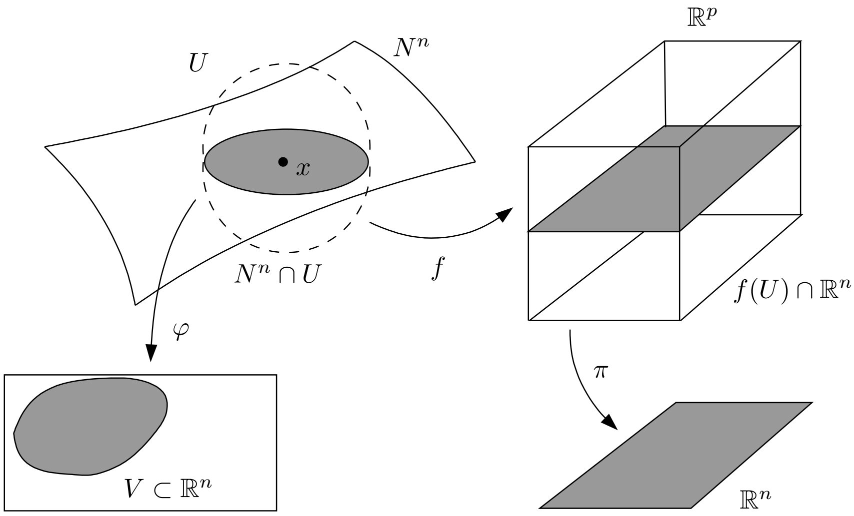

Fig. 2.2. A chart for a submanifold.

In order to show that this definition makes sense, we have to prove that it is independent of the straightening map f. – Let f and $\tilde{f}$ be two straightening maps for N^n about x. Let φ be a chart for N^n about x and assume that $\pi \circ f \circ \varphi^{-1}$ is a C^k-diffeomorphism. We write $\pi \circ \tilde{f} \circ \varphi^{-1} = \pi \circ \tilde{f} \circ f^{-1} \circ f \circ \varphi^{-1}$ and note that by Definition 2.1 the map $\tilde{f} \circ f^{-1} : \mathbb{R}^p \longrightarrow \mathbb{R}^p$ is a local C^k-diffeomorphism sending $\mathbb{R}^n \times \{0\}$ into itself. Hence, $\pi \circ \tilde{f} \circ \varphi^{-1}$ is also a C^k-diffeomorphism showing that the definition of a chart for a submanifold is independent of the straightening map.

Proposition 2.9. *Let N^n be a C^k-submanifold of $\mathbb{R}^p$ with two different charts φ and $\tilde{\varphi}$ about $x \in N^n$. Then for $U \subset \mathbb{R}^p$ containing x the* ***transition function***

$$\tilde{\varphi} \circ \varphi^{-1} : \varphi(N^n \cap U) \longrightarrow \tilde{\varphi}(N^n \cap U)$$

is a C^k-diffeomorphism.

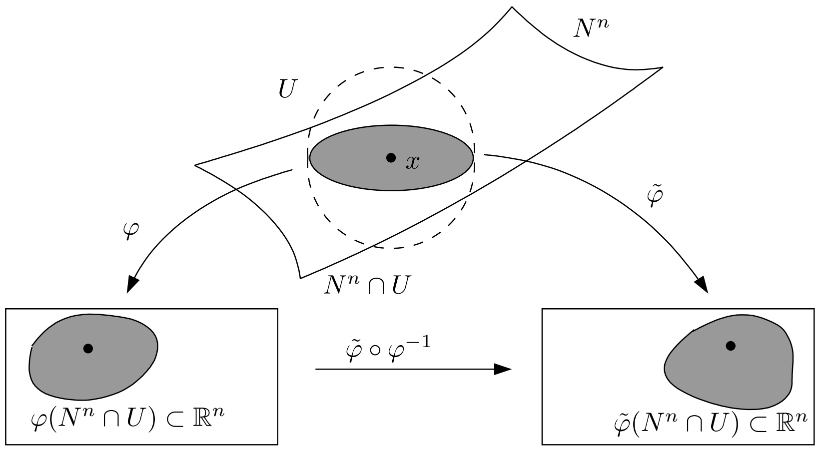

Fig. 2.3. Transition functions for submanifolds.

Proof. Let f be a straightening map for N^n about x. Let $\iota : \mathbb{R}^n \hookrightarrow \mathbb{R}^p$ and $\pi : \mathbb{R}^p \longrightarrow \mathbb{R}^n$ denote the canonical inclusion, respectively, projection. On $\varphi(N^n \cap U) \subset \mathbb{R}^n$, we can then write

$$\tilde{\varphi} \circ \varphi^{-1} = \tilde{\varphi} \circ f^{-1} \circ f \circ \varphi^{-1} = \tilde{\varphi} \circ f^{-1} \circ \iota \circ \pi \circ f \circ \varphi^{-1} , \qquad (2.4)$$

since $\iota \circ \pi\big|_{\mathbb{R}^n} = id\big|_{\mathbb{R}^n}$. The last three maps in (2.4) compose to a C^k-diffeomorphism by Definition 2.8 and the same argument holds for the other three maps writing $\tilde{\varphi} \circ f^{-1} \circ \iota$ as $\left(\pi \circ f \circ \tilde{\varphi}^{-1}\right)^{-1}$. Hence, the transition function $\tilde{\varphi} \circ \varphi^{-1}$ is indeed a C^k-diffeomorphism. □

Now, we are ready to define C^k-functions on submanifolds using charts.

Definition 2.10. *Let N^n be a C^l-submanifold of $\mathbb{R}^p$ and let $k \leq l$. A function $h : N^n \longrightarrow \mathbb{R}$ is called a* ***C^k-function*** *on N^n if for all $x \in N^n$ there exists a local chart φ about x such that*

$$h \circ \varphi^{-1} : \mathbb{R}^n \longrightarrow \mathbb{R}$$

is a C^k-function.

It is important to note that this definition is independent of the choice of the chart φ. Indeed, let φ and $\tilde{\varphi}$ be two charts about x. Assume that $h \circ \varphi^{-1}$ is a C^k-function. From Proposition 2.9 it follows directly that

$$h \circ \tilde{\varphi}^{-1} = h \circ \varphi^{-1} \circ \varphi \circ \tilde{\varphi}^{-1}$$

is also a C^k-function.

Example 2.11. Consider the sphere $S^n \subset \mathbb{R}^{n+1}$. We want to show that the canonical coordinate functions on $\mathbb{R}^{n+1}$ restricted to S^n are C^∞-functions on S^n. – The canonical coordinate functions[1], for $i = 1, \dots, n+1$, are given by

$$\begin{aligned} x_i : \mathbb{R}^{n+1} &\longrightarrow \mathbb{R}\,, \\ x &\longmapsto x_i\,. \end{aligned}$$

Note that in the notation there is no difference between coordinates and coordinate functions.

Let $x_0 = (0, \dots, 0, 1) \in \mathbb{R}^{n+1}$ denote the north pole of the sphere and consider the C^∞-map

$$\begin{aligned} f : U &\longrightarrow \mathbb{R}^{n+1}\,, \\ x &\longmapsto \left(x_1, \dots, x_n, \sum_{i=1}^{n+1} x_i^2 - 1 \right)\,, \end{aligned}$$

where $U \subset \mathbb{R}^{n+1}$ open containing x_0. It is clear that

$$f^{-1}\big((x_1, \dots, x_n, 0)\big) = S^n \cap U\,.$$

Moreover, a straightforward computation shows that $df_{x_0} = \operatorname{diag}(1, \dots, 1, 2)$; hence df_{x_0} is invertible and the Local Inversion Theorem 1.10 shows that f is a local C^∞-diffeomorphism in a neighborhood of x_0. This implies that f is a straightening map for S^n about x_0. – We already showed in Example 2.3 that the sphere is a C^∞-submanifold constructing a constraint. (Note that the straightening map is compatible with the constraint (2.3), in the sense that the last coordinate function of the straightening map equals the constraint.)

Consider now the C^∞-map

$$\begin{aligned} \varphi : S^n \cap U \subset \mathbb{R}^{n+1} &\longrightarrow V \subset \mathbb{R}^n\,, \\ (x_1, \dots, x_n, x_{n+1}) &\longmapsto (x_1, \dots, x_n)\,, \end{aligned}$$

where $V = \varphi(S^n \cap U)$ is an open neighborhood of $\varphi(x_0) = (0, \dots, 0) \in \mathbb{R}^n$. One easily sees that

[1] The corresponding chart is given by the identity map of $\mathbb{R}^{n+1}$, i.e., $\varphi(x) = (x_1, \dots, x_n, x_{n+1})$.

$$\varphi^{-1}(x_1,\dots,x_n) = \left(x_1,\dots,x_n,\sqrt{1-\sum_{i=1}^n x_i^2}\right) \in S^n \cap U\,.$$

Next, we compute, for $x \in V$,

$$\begin{aligned}\left(\pi\circ f\circ\varphi^{-1}\right)(x) &= \left(\pi\circ f\right)\left(x_1,\dots,x_n,\sqrt{1-\sum_{i=1}^n x_i^2}\right)\\ &= \pi(x_1,\dots,x_n,0)\\ &= (x_1,\dots,x_n)\,.\end{aligned}$$

Thus $\left(\pi\circ f\circ\varphi^{-1}\right)$ is the identity on $V\subset\mathbb{R}^n$ and therefore we have a chart φ for S^n about the north pole x_0.

In a next step, we note that the expressions for the canonical coordinate functions $(x_1,\dots,x_n,x_{n+1})$ in the chart φ read as

$$\begin{aligned}x_i\circ\varphi^{-1}(x_1,\dots,x_n) &= x_i\,, \qquad 1\le i\le n\,,\\ x_{n+1}\circ\varphi^{-1}(x_1,\dots,x_n) &= \sqrt{1-\sum_{i=1}^n x_i^2}\end{aligned}$$

showing by Definition 2.10 that all functions are in $C^\infty(U\cap S^n,\mathbb{R})$. This was the claim at the beginning of the example.

We extend the last definition for C^k-functions on submanifolds to maps between two submanifolds.

Definition 2.12. *Let N^n and M^m be two C^l-submanifolds of $\mathbb{R}^p$, respectively $\mathbb{R}^q$ and let $k\le l$. A map $h: N^n \longrightarrow M^m$ is called a **C^k-map** if h is continuous and if for all $x\in N^n$ there exist a chart φ for N^n about x and a chart $\tilde\varphi$ of M^m about $h(x)$ such that*

$$\tilde\varphi\circ h\circ\varphi^{-1}:\mathbb{R}^n\longrightarrow\mathbb{R}^m$$

is a C^k-map.

Remark. By a *continuous* map between submanifolds we mean the continuity of the map with respect to the topologies induced by the ambient Euclidean space on the submanifolds (restriction of open sets).

As a direct consequence of Proposition 2.9, we get that the definition is independent of the choice of the charts for N^n and M^m.

Definition 2.13. *Let N^n and M^m be two C^l-submanifolds of $\mathbb{R}^p$, respectively $\mathbb{R}^q$. A map $h: N^n\longrightarrow M^m$ is called a **C^k-diffeomorphism** if h is an homeomorphism, a C^k-map and if for all $x\in N^n$ and all charts φ about x, $\tilde\varphi$ about $h(x)$, we have that the linear map*

$$d\left(\tilde\varphi\circ h\circ\varphi^{-1}\right)_{\varphi(x)}$$

is invertible.

2.3 Abstract Manifolds

In this section, we introduce a large class of topological spaces, called manifolds, which are *locally* homeomorphic to open subsets of Euclidean spaces. However, these spaces don't require anymore the concept of "ambient space" which plays a crucial role in the definition of submanifolds.

Though they have no global vector space structure, there will be a notion of differentiability. In the next Section 2.4, we will then study in great detail the differentiability on manifolds. – Remember that the notion of differentiability was first defined on normed vector spaces in Section 1.1 and then extended to submanifolds of Euclidean spaces in the last section. – In the following, we give some motivations for the need of topological spaces generalizing submanifolds.

We introduce the set of straight lines in $\mathbb{R}^{n+1}$ passing through the origin and denote this set by $\mathbb{R}P^n$. In other words, we consider in $\mathbb{R}^{n+1} \setminus \{0\}$ the equivalence relation $x = (x_1, \ldots, x_{n+1}) \sim y = (y_1, \ldots, y_{n+1})$ if and only if there exists $\lambda \in \mathbb{R}$ such that $x = \lambda y$. Alternatively, the set $\mathbb{R}P^n$ can be considered as the quotient $S^n / \sim$, where the equivalence relation is defined by:

$$u \sim v \text{ iff } u = \pm v \qquad \text{for } u, v \in S^n \,.$$

Thus the equivalence classes $[u] = \{u, -u\} \in S^n / \sim$ define the set $\mathbb{R}P^n$.

Let $\pi : S^n \longrightarrow \mathbb{R}P^n$ be the canonical projection which gives $\mathbb{R}P^n$ the quotient topology, i.e., U open in $\mathbb{R}P^n$ if and only if $\pi^{-1}(U)$ is open in S^n. For this topology $\mathbb{R}P^n$ is called the **real projective space** of dimension n.

Note that *locally* the real projective space looks identical to S^n. More precisely, this identification can be made as long as the antipodal points are not considered simultaneously. Nevertheless, $\mathbb{R}P^n$ cannot be seen as a submanifold of $\mathbb{R}^{n+1}$.

We have seen in Example 2.3 that S^1 is a submanifold of $\mathbb{R}^2$. By Proposition 2.6 the torus $T^2 = S^1 \times S^1$ is thus a submanifold of $\mathbb{R}^4$. – Next, we consider the map

$$\begin{aligned} \Xi : \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \,, \\ (\theta, \psi) &\longmapsto \big(\cos\theta(2 + \cos\psi), \sin\theta, \sin\psi\big) \,. \end{aligned}$$

One can check that $N^2 := \Xi(\mathbb{R}^2)$ is a two-dimensional submanifold of $\mathbb{R}^3$. Moreover,

$$\begin{aligned} \tilde{\Xi} : S^1 \times S^1 &\longrightarrow \Xi(\mathbb{R}^2) \,, \\ (\mathrm{e}^{i\theta}, \mathrm{e}^{i\psi}) &\longmapsto \big(\cos\theta(2 + \cos\psi), \sin\theta, \sin\psi\big) \,, \end{aligned}$$

is a C^∞-diffeomorphism between the two submanifolds T^2 and N^2. – We want to construct a topological space without using the notion of "ambient space", whose embedding (realization) in $\mathbb{R}^4$ and $\mathbb{R}^3$ is given by T^2 and N^2, respectively.

Consider $\mathbb{R}^2/\sim$ with the equivalence relation:

$$(a,b) \sim (a',b') \text{ iff } (a',b') = (a,b) + (k,l)\,,$$

where $(k,l) \in \mathbb{Z}^2$. We denote this set by $\mathbb{R}^2/\mathbb{Z}^2$. The canonical projection $\pi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2/\mathbb{Z}^2$ defines the quotient topology, i.e., U open in $\mathbb{R}^2/\mathbb{Z}^2$ if and only if $\pi^{-1}(U)$ open in $\mathbb{R}^2$. Geometrically, this quotient space leads to a certain identification of the edges of a square with unit length as described in Figure 2.4. Hence, the quotient space $\mathbb{R}^2/\mathbb{Z}^2$ is homeomorphic to T^2 and N^2.

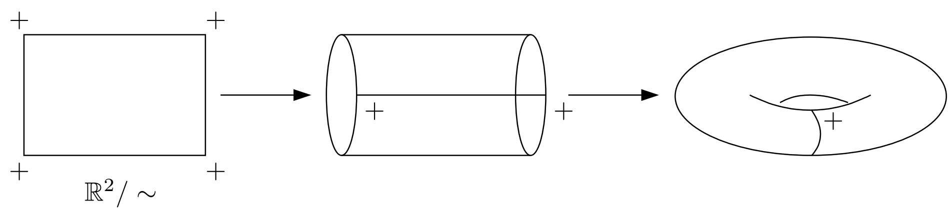

Fig. 2.4. Torus.

Next, we mention the example of the *Klein bottle* being not a submanifold. Starting from $\mathbb{R}^2/\mathbb{Z}^2$ with the quotient topology, we introduce an additional equivalence relation $\sim'$ on $[0,\frac{1}{2}] \times [0,1] \subset \mathbb{R}^2/\mathbb{Z}^2$. Namely, for $a,b,a',b' \in [0,\frac{1}{2}] \times [0,1]$, we have

$$(a,b) \sim' (a',b') \text{ iff } (a',b') = \left(a + \frac{1}{2}, -b\right)\,.$$

As described in Fig. 2.5, this equivalence relation leads again to identifications of the edges in the square $[0,\frac{1}{2}] \times [0,1] \subset \mathbb{R}^2/\mathbb{Z}^2$.

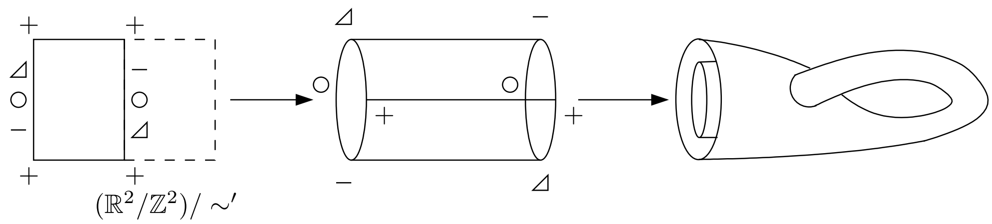

Fig. 2.5. Klein bottle.

Now, we give a precise definition of a manifold being an important mathematical object.

Definition 2.14. *Let X be a topological space. Then X is called* ***separated*** *or* ***Hausdorff*** *if for all x, $\tilde{x} \in X$ there exist open neighborhoods U_x and $U_{\tilde{x}}$ with $U_x \cap U_{\tilde{x}} = \emptyset$.*

Definition 2.15. *A n-dimensional* ***topological manifold*** *or C^0-**manifold** is a topological separated space X such that for all $x \in X$ there exist an open neighborhood U_x and an homeomorphism $\varphi : U_x \longrightarrow \varphi(U_x) \subset \mathbb{R}^n$.*

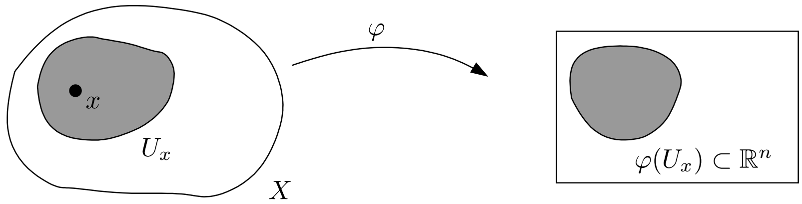

Fig. 2.6. Topological manifold.

Remark. A topological manifold is thus by definition *locally* homeomorphic to $\mathbb{R}^n$ which does not imply that the space is separated (see Exercise 2.31).

Next, we define C^1-manifolds. Naively, one would simply replace the defining homeomorphism by a diffeomorphism. But, we have to be more careful.

Definition 2.16. *Let X be a topological separated space and $k \geq 1$. A C^k-**atlas** or C^k-**system of charts** of dimension n is a family $(U_i, \varphi_i)_{i \in I}$, where the $(U_i)_{i \in I}$ are open subsets of X and the $(\varphi_i)_{i \in I}$ are homeomorphisms from U_i into $\mathbb{R}^n$, with the following properties:*

(i) The $(U_i)_{i \in I}$ cover the topological space X, i.e., $\bigcup_{i \in I} U_i = X$.
(ii) For every $i, j \in I$ such that $U_i \cap U_j \neq \emptyset$, the ***transition functions***

$$\varphi_j \circ \varphi_i^{-1} : \varphi_i(U_i \cap U_j) \longrightarrow \varphi_j(U_i \cap U_j)$$

are C^k-diffeomorphisms.

The previous definition shows that we can only understand the topological space X from a differential point of view via a system of charts. Moreover, if X admits a C^k-system of charts of dimension n then it cannot admit a system of charts of dimension p for $n \neq p$. Indeed, let $(U_i, \varphi_i)_{i \in I}$ and $(V_j, \psi_j)_{j \in J}$ denote the two systems of charts with different dimensions. For $i \in I$, $j \in J$ such that $U_i \cap V_j \neq \emptyset$ we have that $\varphi_i(U_i \cap V_j) \subset \mathbb{R}^n$ and $\psi_j(U_i \cap V_j) \subset \mathbb{R}^p$ are both open. By Definition 2.16 (ii) we obtain a homeomorphism between two open sets of Euclidean spaces with different dimensions. This contradicts Corollary 1.13.

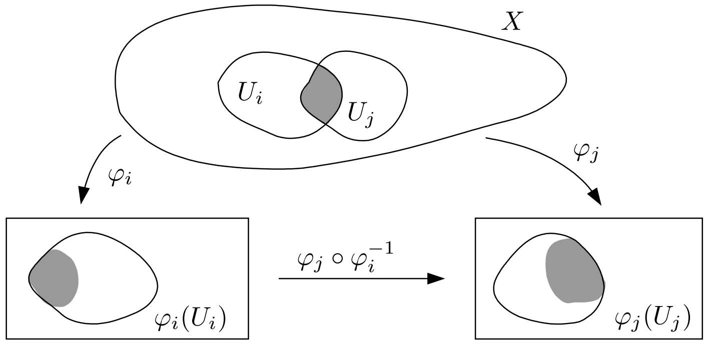

Fig. 2.7. System of charts and transition functions.

Definition 2.17. *Let X be a topological separated space admitting two C^k-systems of charts $(U_i, \varphi_i)_{i\in I}$ and $(V_j, \psi_j)_{j\in J}$ of same dimension. These two systems of charts are called* ***equivalent*** *if the union of them is still a C^k-system of charts, i.e., if for every $i \in I$, $j \in J$ such that $U_i \cap V_j \neq \emptyset$ the transition functions*

$$\psi_j \circ \varphi_i^{-1} : \varphi_i(U_i \cap V_j) \longrightarrow \psi_j(U_i \cap V_j)$$

are C^k-diffeomorphisms.

It is clear that the union of two charts still covers the topological space X as required in the Definition 2.16. Moreover, the relation “being equivalent” of the previous definition defines an equivalence relation on the systems of charts and we arrive at the important definition of n-dimensional C^k-manifolds:

Definition 2.18. *Let X be a topological separated space. An equivalence class of C^k-systems of charts is called a C^k-**differentiable structure** on X. Assume, in addition, that the systems of charts are n-dimensional. Then X together with a differentiable structure is called a n-dimensional C^k-**manifold**.*

Examples of Manifolds

Example 2.19 (The Euclidean Space). The separated topological space $\mathbb{R}^n$ becomes a manifold for the single chart (U, φ), where $U = \mathbb{R}^n$ and $\varphi = id_{\mathbb{R}^n}$.

Example 2.20 (A Submanifold as Manifold). Let N^n be a n-dimensional C^k-submanifold of $\mathbb{R}^p$. We consider the coordinate system $\varphi_x : N^n \cap U_x \longrightarrow \mathbb{R}^n$ of Definition 2.8 where $U_x \subset \mathbb{R}^p$ is an open neighborhood of x. By Definition 2.1 such coordinate systems exist for every $x \in N^n$. Indeed, we set $\varphi_x := \pi \circ f_x\big|_{N^n \cap U_x}$ with f_x the straightening map about x and $\pi : \mathbb{R}^p \longrightarrow \mathbb{R}^n$ the canonical projection. Using Proposition 2.9 the family $(U_x, \varphi_x)_{x\in N^n}$ defines a C^k-system of charts for the submanifold N^n. Thus N^n becomes a C^k-manifold of dimension n.

Example 2.21 (The Real Projective Space). The real projective space $\mathbb{R}P^n = \big(\mathbb{R}^{n+1} \setminus \{0\}\big)/\sim$ was already defined at the beginning of this section. In order to show that this space is a C^∞-manifold of dimension n, we have to find an atlas for it.

Let

$$\begin{aligned} \pi : \mathbb{R}^{n+1} \setminus \{0\} &\longrightarrow \mathbb{R}P^n\,, \\ x = (x_1, \ldots, x_{n+1}) &\longmapsto [x] = [x_1, \ldots, x_{n+1}] \end{aligned}$$

denote the canonical projection. For $i = 1, \ldots, n+1$, we then define $U_i := \{[x] \,:\, x_i \neq 0\}$ and

$$\varphi_i\big([x]\big) := \left(\frac{x_1}{x_i}, \ldots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \ldots, \frac{x_{n+1}}{x_i}\right) \subset \mathbb{R}^n\,.$$

It is easy to check that $\varphi_i : U_i \longrightarrow \mathbb{R}^n$ is an homeomorphism for all $i = 1, \ldots, n+1$ and that $\bigcup_{i=1}^{n+1} U_i = \mathbb{R}P^n$. By Definition 2.16, it remains to show that for all i, j with $U_i \cap U_j \neq \emptyset$ the transition functions

$$\varphi_j \circ \varphi_i^{-1} : \varphi_i(U_i \cap U_j) \longrightarrow \varphi_j(U_i \cap U_j)$$

are C^∞-diffeomorphisms.

We have, for $z = (z_1, \ldots, z_n) \in \varphi_i(U_i \cap U_j) \subset \mathbb{R}^n$ assuming that $i > j$,

$$\begin{aligned} \varphi_j \circ \varphi_i^{-1}(z_1, \ldots, z_n) &= \varphi_j\big([z_1, \ldots, z_{i-1}, 1, z_{i+1}, \ldots, z_n]\big) \\ &= \left(\frac{z_1}{z_j}, \ldots, \frac{z_{j-1}}{z_j}, \frac{z_{j+1}}{z_j}, \ldots, \frac{z_{i-1}}{z_j}, \frac{1}{z_j}, \frac{z_{i+1}}{z_j}, \ldots, \frac{z_n}{z_j}\right) \in \mathbb{R}^n\,. \end{aligned}$$

Since by assumption $z_j \neq 0$, we deduce that $\varphi_j \circ \varphi_i^{-1} \in C^\infty(\varphi_i(U_i \cap U_j), \varphi_j(U_i \cap U_j))$. One can check that this implies that $(U_i, \varphi_i)_{i=1,\ldots,n+1}$ is an atlas for the real projective space $\mathbb{R}P^n$.

Example 2.22 (The Torus). We consider again the example of the torus $T^2 = \mathbb{R}^2/\mathbb{Z}^2$. Let $x \in \mathbb{R}^2$ and $\alpha \in \mathbb{R}$ with $0 < \alpha < 1/4$. We define $U_x := \pi(B_\alpha(x)) \subset \mathbb{R}^2/\mathbb{Z}^2$, where π is the canonical projection. Since

$$\pi^{-1}(U_x) = \bigcup_{(k,l)\in\mathbb{Z}^2} B_\alpha\big(x + (k,l)\big)\,,$$

we deduce that U_x is open and $(U_x)_{x\in\mathbb{R}^2}$ is thus an open covering of T^2. Let $[y] \in U_x$ and let y denote the unique representant in $B_\alpha(x)$ of the equivalence class $[y]$. Then, we define

$$\begin{aligned} \varphi_x : U_x &\longrightarrow \mathbb{R}^2\,, \\ [y] &\longmapsto y \in B_\alpha(x)\,. \end{aligned}$$

It is easy to show that $\varphi_x : U_x \longrightarrow B_\alpha(x)$ is an homeomorphism.

In a next step, we show that $(U_x, \varphi_x)_{x\in\mathbb{R}^2}$ is an atlas for T^2. – Let $x \neq \tilde{x} \in \mathbb{R}^2$ such that $U_x \cap U_{\tilde{x}} \neq \emptyset$. By definition of U_x this means that there exists $(k,l) \in \mathbb{Z}^2$ such that $B_\alpha(x) \cap B_\alpha(\tilde{x}+(k,l)) \neq$ and because $\alpha \in \mathbb{R}$ small enough the pair $(k,l) \in \mathbb{Z}^2$ is unique. Take $z \in B_\alpha(x) \cap B_\alpha(\tilde{x}+(k,l))$, then an easy computation shows that (note that $(k,l) \in \mathbb{R}^2$ is independent of $z \in \mathbb{R}^2$)

$$\varphi_{\tilde{x}} \circ \varphi_x^{-1}(z) = z - (k,l)\,,$$

hence a C^∞-diffeomorphism and T^2 is a C^∞-manifold of dimension 2 for the atlas $(U_x, \varphi_x)_{x\in\mathbb{R}^2}$.

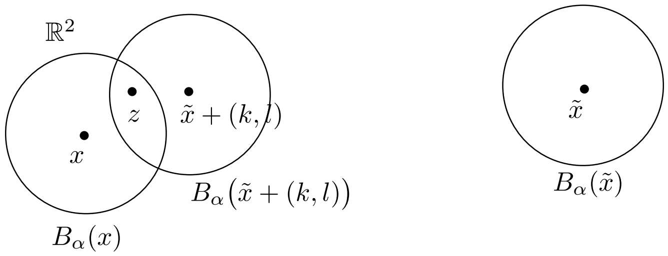

Fig. 2.8. Atlas for the torus.

The following proposition allows us to construct further examples of manifolds.

Proposition 2.23. *Let M^m and N^n be two C^k-manifolds for the atlas $(U_i, \varphi_i)_{i\in I}$ and $(V_j, \psi_j)_{j\in J}$, respectively. Then the product $M^m \times N^n$ is a C^k-manifold of dimension $m+n$ for the atlas*

$$(U_i \times V_j, \varphi_i \times \psi_j)_{(i,j)\in I\times J}\,,$$

where we define $\varphi_i \times \psi_j(x,y) := \big(\varphi_i(x), \psi_j(y)\big)$ for all $(x,y) \in M^m \times N^n$.

Proof. The proof is a straightforward application of the definitions. □

2.3.1 Some Topological Properties of Differentiable Manifolds

Proposition 2.24. *A topological manifold M^m is a locally compact topological space.*

Proof. We have to show that for all $x \in M^m$ there exist a compact set K_x containing x and a open neighborhood $\tilde{U}_x$ such that $\tilde{U}_x \subset K_x$. – Let $x \in M^n$ and (U_x, φ_x) a chart about x. By Definition 2.15, the map φ_x is an homeomorphism from U_x into $\varphi_x(U_x) \subset \mathbb{R}^m$. Hence $\varphi_x(U_x)$ is open and there exists $\rho > 0$ such that $B_\rho(\varphi_x(x)) \subset \varphi_x(U_x)$. The set

$$\varphi_x^{-1}\big(\bar{B}_\rho(\varphi(x))\big) := K_x$$

is then compact. Moreover, for $\varphi_x^{-1}\big(B_\rho(\varphi(x))\big) := \tilde{U}_x$ open, we have $\tilde{U}_x \subset K_x$. □

Proposition 2.25. *A topological manifold M^m is a locally connected space, i.e., every point has a connected open neighborhood.*

Proof. Again, we use a chart (U_x, φ_x) about $x \in M^m$. Clearly, $\varphi_x(U_x)$ is an open neighborhood of $\varphi_x(x) \in \mathbb{R}^m$ and it contains a connected neighborhood $C \subset \varphi_x(U_x) \subset \mathbb{R}^m$ of $\varphi_x(x)$. Since φ_x is an homeomorphism into $\varphi_x(U_x)$, the set $\varphi_x^{-1}(C) \subset M^m$ is also connected. □

Proposition 2.26. *A topological manifold M^m is connected if and only if it is path connected.*

Proof. It is a well-known fact that a *path* connected topological space is also connected. – Conversely, let $x \in M^m$ and set

$$C_x = \Big\{ y \in M^m \;:\; \exists \gamma \in C^0\big([0,1], M^m\big) \text{ with } \gamma(0) = x\,, \gamma(1) = y \Big\}\,.$$

Clearly, $C_x \neq \emptyset$ since $x \in C_x$. We will show that C_x is both open and closed in M^m. Then, since by assumption M^m is connected, we deduce that $C_x = M^m$ indicating that the topological manifold M^m is path connected.

The set C_x is open: Let $y \in C_x$ and (U, φ) a chart about y. Since $\varphi(U)$ is open there exists $\rho > 0$ such that $B_\rho(\varphi(y)) \subset \varphi(U)$. Consider for $z \in B_\rho(\varphi(y))$ the map

$$\tilde{\gamma}_z(t) := \varphi^{-1}\big(t\, z + (1-t)\varphi(y)\big) \in C^0([0,1], U)\,.$$

Clearly, we have $\tilde{\gamma}_z(0) = y$ and $\tilde{\gamma}_z(1) = \varphi^{-1}(z)$, thus a path joining y and $\varphi^{-1}(z)$. Since by assumption there is a path between x and y, we can also join x and $\varphi^{-1}(z)$ by a path for arbitrary $z \in B_\rho(\varphi(y))$. Thus $y \in C_x$ implies that $\varphi^{-1}\big(B_\rho(\varphi(y))\big) \subset C_x$, which is also open, since φ a homeomorphism. Hence, C_x is open.

The set C_x is closed: Let (U, y) be a chart about y and let $y \in \bar{C}_x$. We show that $y \in C_x$. – Since $\varphi(y) \in \varphi(U)$ which is open, we deduce the existence of $\rho > 0$ such that $B_\rho(\varphi(y)) \subset \varphi(U)$. Thus $\varphi^{-1}\big(B_\rho(\varphi(y))\big) \subset U$; moreover, it is open and contains y. Therefore, there exists $z \in C_x \cap \varphi^{-1}\big(B_\rho(\varphi(y))\big)$. Since $\varphi(z) \in B_\rho(\varphi(y))$ there is a path in $\mathbb{R}^m$ joining $\varphi(y)$ and $\varphi(z)$. By φ^{-1} also a path joining y and z. From the fact that $z \in C_x$, we obtain that y and x are also path connected showing that $y \in C_x$. □

As an application of the previous proposition, we consider the topological space

$$E = \left\{ \left(x, \sin\frac{1}{x} \right) \;:\; x \in \mathbb{R}^+ \setminus \{0\} \right\} \cup \Big\{ (0, y) \;:\; y \in [-1, 1] \Big\} \subset \mathbb{R}^2\,,$$

which is connected but *not* path connected. Hence, by Proposition 2.26 the topological space E is *not* a topological manifold.

2.4 Differentiability on a Manifolds

As announced before, we now want to give a meaning to a C^k-map between two manifolds. As in the case of submanifolds, this can be done using charts.

Definition 2.27. *Let M^m and N^n be two C^k-manifolds with $k \geq 1$. A map $f : M^m \longrightarrow N^n$ is a C^k-**map** if f is continuous and if for all $(i,j) \in I \times J$ the map*

$$\psi_j \circ f \circ \varphi_i^{-1} : \varphi_i\big(U_i \cap f^{-1}(V_j)\big) \longrightarrow \psi_j(V_j)$$

*is a C^k-map, where $(U_i, \varphi_i)_{i \in I}$ and $(V_j, \psi_j)_{j \in J}$ are systems of charts for M^m and N^n, respectively. If f is a C^k-map, we write $f \in C^k(M^m, N^n)$. The map $\psi_j \circ f \circ \varphi_i^{-1}$ is often called the **coordinate expression** for f.*

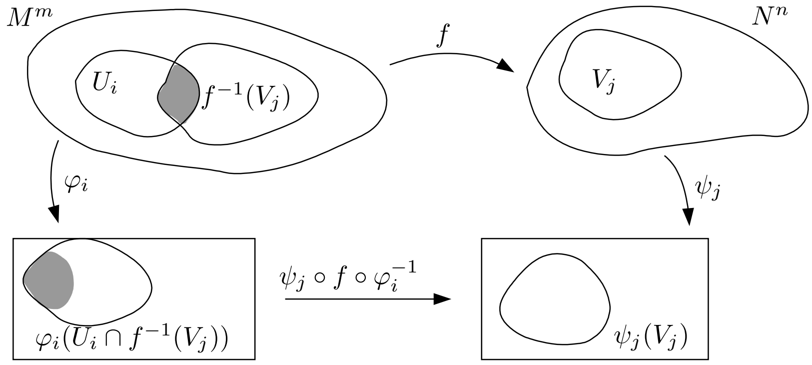

Fig. 2.9. A C^k-map between manifolds.

This definition is independent of the choice of $(U_i, \varphi_i)_{i \in I}$ and $(V_j, \psi_j)_{j \in J}$ representing the fixed differentiable structure of the manifold. To see this, let $(\tilde{U}_i, \tilde{\varphi}_i)_{i \in \tilde{I}}$ and $(\tilde{V}_j, \tilde{\psi}_j)_{j \in \tilde{J}}$ be two other charts of the fixed differentiable structures on M^m and N^n, respectively. (In other words, we choose two charts equivalent to the initial ones.) Assume that the map $f : M^m \longrightarrow N^n$ is a C^k-map for the charts $(U_i, \varphi_i)_{i \in I}$ and $(V_j, \psi_j)_{j \in J}$. Writing, for all $\tilde{i} \in \tilde{I}$ and $\tilde{j} \in \tilde{J}$,

$$\tilde{\psi}_{\tilde{j}} \circ f \circ \tilde{\varphi}_{\tilde{i}}^{-1} = \tilde{\psi}_{\tilde{j}} \circ \psi_j^{-1} \circ \psi_j \circ f \circ \varphi_i^{-1} \circ \varphi_i \circ \tilde{\varphi}_{\tilde{i}}^{-1} \,,$$

we deduce, using Definition 2.17, that the map f is again a C^k-map.

Definition 2.28. *A map $f : M^m \longrightarrow N^n$ is called a C^k-**diffeomorphism** if f is an homeomorphism and if f and f^{-1} are both C^k-maps. If such a map exists then M^m and N^n are called **diffeomorphic**.*

Example 2.29. Consider the topological space $\mathbb{R}$ with the single chart (U, φ), where $U = \mathbb{R}$ and $\varphi = id_{\mathbb{R}}$. This gives the canonical differentiable structure on $\mathbb{R}$ and denote the resulting manifold by M. On the other hand, consider again $\mathbb{R}$ with the single chart (V, ψ) where $V = \mathbb{R}$ and $\psi(x) = x^{1/3}$ and denote the resulting manifold by N. It is clear that the two charts are not equivalent. Indeed, the map

$$\begin{aligned} \psi \circ \varphi^{-1} : \mathbb{R} &\longrightarrow \mathbb{R}\,, \\ x &\longmapsto x^{1/3} \end{aligned}$$

is not continuously differentiable.

Nevertheless, the manifolds M and N are diffeomorphic. In fact, the map

$$\begin{aligned} f : M &\longrightarrow N\,, \\ x &\longmapsto x^3 \end{aligned}$$

is a C^∞-diffeomorphism, since $\psi \circ f \circ \varphi^{-1} = id_{\mathbb{R}}$. Thus, the two charts define two different differentiable structures but the corresponding two different manifolds are diffeomorphic to each other.

We end up with the interesting question if there exists on $\mathbb{R}^n$, considered as topological space, a differentiable structure which is *not* diffeomorphic to the canonical one? (Note that "diffeomorphic" refers to the corresponding manifold.) We mention that only for $n = 4$ such a differentiable structure exists (see []).

Definition 2.30. *Let M^m and N^n be two C^k-manifolds with $k \geq 1$. Moreover, let $f \in C^k(M^m, N^n)$. Then f is called **immersion**, respectively, **submersion** if for all $(i, j) \in I \times J$ the differential*

$$d(\psi_j \circ f \circ \varphi_i^{-1}) : \mathbb{R}^m \longrightarrow \mathbb{R}^n$$

is injective, respectively, surjective, where the usual notation is used for the systems of charts for M^m and N^n.

This definition is again independent of the choice of $(U_i, \varphi_i)_{i \in I}$ and $(V_j, \psi_j)_{j \in J}$ representing the fixed differentiable structure of the manifold.

* * *

Exercises.

Exercise 2.31.

2.5 The Tangent Manifold to a Differentiable Manifold

2.5.1 The Tangent Space to a Submanifold of $\mathbb{R}^p$

Let N^n be a n-dimensional C^1-submanifold of $\mathbb{R}^p$. We first consider the case $n = 1$.

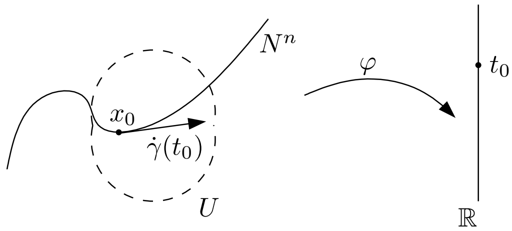

Fig. 2.10. Tangent space to a 1-dimensional submanifold.

Let $x_0 \in N^1$ and $U \subset \mathbb{R}^p$ open neighborhood of x_0. Consider a chart $\varphi : N^1 \cap U \longrightarrow V$ about x_0, where $V \subset \mathbb{R}$ open (see Definition 2.8), and define

$$\gamma(t) := \varphi^{-1}(t)\,, \qquad t \in \mathbb{R}\,. \tag{2.5}$$

For $\gamma(t) \in U$ we clearly have that $|\dot\gamma(t)| \neq 0$. Hence, the map γ is a regular parameterization of a curve lying in the 1-dimensional C^1-submanifold $N^1 \cap U \subset \mathbb{R}^p$. The *tangent space* at $\gamma(t_0) = x_0$ is the space generated by the velocity vectors of the curve γ at t_0:

$$\dot\gamma(t_0) = \left.\frac{d}{dt}\right|_{t=t_0} \gamma(t)\,. \tag{2.6}$$

Let $\sigma : \mathbb{R} \longrightarrow N^1 \cap U$ be another C^1-curve such that

$$\sigma(s) = \gamma \circ \psi(s)\,,$$

where $\psi \in C^1(\mathbb{R}, \mathbb{R})$ verifying $\psi(s_0) = t_0$; hence $\sigma(s_0) = x_0$. The velocity vector of the curve σ at s_0 is given by

$$\dot\sigma(s_0) = \left.\frac{d}{ds}\right|_{s=s_0} \sigma(s) = \dot\gamma\big(\psi(s_0)\big)\,\left.\frac{d}{ds}\right|_{s=s_0} \psi(s)\,.$$

Showing that $\dot\gamma(t_0)$ is only multiplied by a real number. – Therefore, we can interpret the tangent space to a 1-dimensional submanifold N^1 at a point as the velocity vectors of *arbitrary* curves in N^1 passing through this point.

This motivates the following definition for arbitrary dimensions.

Definition 2.32. *Let N^n be a n-dimensional C^1-submanifold of $\mathbb{R}^p$ and $x_0 \in N^n$. The **tangent space** to N^n at x_0 is defined as the set of velocity vectors of arbitrary curves $C^1([0,1], N^n)$ passing through x_0. The space is denoted by $T_{x_0}N^n$.*

Proposition 2.33. *Let N^n be a n-dimensional C^1-submanifold of $\mathbb{R}^p$ and $x_0 \in N^n$. Then the tangent space $T_{x_0}N^n$ is a n-dimensional vector subspace of $\mathbb{R}^p$. Moreover, if U denote an open neighborhood of x_0 and $f^1, \ldots, f^{p-n}$:*

$U \longrightarrow \mathbb{R}$ *denote a free family of constraints for* N^n *about* x_0 *(see Proposition 2.2), then*

$$T_{x_0}N^n = \bigcap_{i=1}^{p-n} ker\, df^i(x_0)\,. \tag{2.7}$$

Proof. Since the differential is a linear map, the expression $\ker df^i(x_0)$, $i = 1, \ldots, p-n$, characterizes $(p-n)$ hyperplanes in $\mathbb{R}^p$. By assumption, they are all free, hence

$$\dim \bigcap_{i=1}^{p-n} \ker df^i(x_0) = p - (p-n) = n\,.$$

Therefore, if we prove the second assertion the first follows.

Let $\gamma \in C^1([0,1], N^n)$ with $\gamma(t_0) = x_0$ and let $U \subset \mathbb{R}^p$ be an open neighborhood of $x_0 \in N^n$. Fix $\alpha > 0$ such that $\gamma(t) \in N^n \cap U$ for $t \in [t_0 - \alpha, t_0 + \alpha]$. Using (2.2), we see that $f^i\big(\gamma(t)\big) = 0$, for $i = 1, \ldots, p-n$ and for all $t \in [t_0 - \alpha, t_0 + \alpha]$. Differentiating this identity at t_0 gives $(df^i)_{\gamma(t)} \cdot \dot{\gamma}(t_0) = 0$, hence

$$\dot{\gamma}(t_0) \in \bigcap_{i=1}^{p-n} \ker df^i(x_0)\,,$$

and so

$$T_{x_0}N^n \subset \bigcap_{i=1}^{p-n} \ker df^i(x_0)\,.$$

Conversely, let f be a straightening map for N^n about x_0 compatible with the constraints, i.e., if we write $f = (\tilde{f}^1, \ldots, \tilde{f}^p)$, then $\tilde{f}^j = f^{j-n}$ for $j = n+1, \ldots, p$. For $X \in \bigcap_{i=1}^{p-n} \ker df^i(x_0) \subset \mathbb{R}^p$, we thus have that $Y := df_{x_0} \cdot X$ lies in $\mathbb{R}^n$. Note that, using Definition 2.1 (iii), we have that $f(x_0) + t\,Y \in \mathbb{R}^n \cap f(U)$ for t small enough, say $|t| < \alpha$, $\alpha > 0$. Consider now the curve

$$\gamma(t) := f^{-1}\big(f(x_0) + t\,Y\big)\,.$$

We see that $\gamma \in C^1\big((-\alpha, \alpha), N^n \cap U\big)$. Also, we have

$$\dot{\gamma}(0) = df^{-1}_{f(x_0)} \cdot Y = X\,,$$

showing that

$$\bigcap_{i=1}^{p-n} \ker df^i(x_0) \subset T_{x_0}N^n\,.$$

□

This proposition shows that once we have the constraints the equation for the tangent plane is easy to find.

Example 2.34. Consider again the sphere $S^n \in \mathbb{R}^{n+1}$ with the constraint (see (2.3))

$$f^1(x) = x_1^2 + \ldots + x_{n+1}^2 - 1\,.$$

Let $x_0 = (x_1^0, \ldots, x_{n+1}^0) \in S^n$; then $(df^1)_{x_0} \cdot X = 2\sum_{i=1}^{n+1} x_i^0\, X_i$ for the vector $X = (X_1, \ldots, X_{n+1}) \in \mathbb{R}^{n+1}$. Thus for the tangent space of the sphere, we get

$$T_{x_0}S^n = \ker (df^1)_{x_0} = \{X \in \mathbb{R}^{n+1} \ : \ \langle X, x_0\rangle = 0\}\,,$$

where $\langle \cdot, \cdot\rangle$ means the usual scalar product in $\mathbb{R}^{n+1}$.

2.5.2 The Tangent Space to a Manifold

Let M^m be a m-dimensional C^1-manifold with a system of charts $(U_i, \varphi_i)_{i\in I}$ and $x_0 \in M^m$. Still in this more general case there exists the notion of a smooth path in M^m passing through x_0. Indeed, we have that $\gamma \in C^1([0,1], M^m)$ passing through x_0 if, for every $i \in I$,

$$\varphi_i \circ \gamma : \gamma^{-1}(U_i) \subset [0,1] \longrightarrow \mathbb{R}^n \in C^1([0,1], \mathbb{R}^n)\,,$$

and there exists $t_0 \in [0,1]$ such that $\gamma(t_0) = x_0$ (see Definition 2.27). The velocity vector $\dot{\gamma}$ of the smooth curve, however, has a priori no meaning anymore, since there is not an "ambient Euclidean space" as in the case of submanifolds. Using the concept of charts, we can only look at the expressions

$$\left.\frac{d}{dt}\right|_{t=t_0} \big(\varphi_i \circ \gamma\big)(t)\,. \tag{2.8}$$

Unfortunately, these velocity vectors depend on the system of charts $(U_i, \varphi_i)_{i\in I}$ and *not* only on the differentiable structure. More precisely, the velocities of two different charts are related as (we need that $x_0 \in U_i \cap U_j$)

$$\begin{aligned}\left.\frac{d}{dt}\right|_{t=t_0} \big(\varphi_j \circ \gamma\big)(t) &= \left.\frac{d}{dt}\right|_{t=t_0} \big(\varphi_j \circ \varphi_i^{-1} \circ \varphi_i \circ \gamma\big)(t)\\ &= d(\varphi_j \circ \varphi_i^{-1})_{\varphi_i(x_0)}\left.\frac{d}{dt}\right|_{t=t_0} \big(\varphi_i \circ \gamma\big)(t)\,.\end{aligned}$$

Obviously, there is no reason for the differential of the transition function to be $id_{\mathbb{R}^m}$. Therefore, we cannot define the tangent space $T_{x_0}M^m$ of a manifold as being the velocities of curves passing through x_0 as in the case of submanifolds.

We now try to define a tangent space as a vector space which is independent of the choice of the representant of the differentiable structure. – Let γ_1, $\gamma_2 \in C^1([0,1], M^m)$ be two paths with $\gamma_1(t_0) = \gamma_2(t_0) = x_0$. We observe that if there exists a chart (U_i, φ_i) such that

$$\frac{d}{dt}\bigg|_{t=t_0} \big(\varphi_i \circ \gamma_1\big)(t) = \frac{d}{dt}\bigg|_{t=t_0} \big(\varphi_i \circ \gamma_2\big)(t)\,, \tag{2.9}$$

then for another chart (U_j, φ_j), we have (note that $x_0 \in U_i \cap U_j$)

$$\begin{aligned}
\frac{d}{dt}\bigg|_{t=t_0} \big(\varphi_j \circ \gamma_1\big)(t) &= \frac{d}{dt}\bigg|_{t=t_0} \big(\varphi_j \circ \varphi_i^{-1} \circ \varphi_i \circ \gamma_1\big)(t) \\
&= d(\varphi_j \circ \varphi_i^{-1})_{\varphi_i(x_0)}\, \frac{d}{dt}\bigg|_{t=t_0} \big(\varphi_i \circ \gamma_1\big)(t) \\
&= d(\varphi_j \circ \varphi_i^{-1})_{\varphi_i(x_0)}\, \frac{d}{dt}\bigg|_{t=t_0} \big(\varphi_i \circ \gamma_2\big)(t) \\
&= \frac{d}{dt}\bigg|_{t=t_0} \big(\varphi_j \circ \gamma_2\big)(t)\,.
\end{aligned}$$

Thus (2.9) holds also for the second chart.

We denote the set of smooth paths passing through $x_0 \in M^m$ by

$$\mathcal{P}_{x_0}(M^m) = \Big\{\gamma \in C^1([-\delta, \delta], M^m) \;:\; \delta > 0 \text{ and } \gamma(0) = x_0\Big\}\,. \tag{2.10}$$

On $\mathcal{P}_{x_0}(M^m)$ we introduce the following equivalence relation: $\gamma_1 \sim \gamma_2$ if and only if there exists $i \in I$ such that

$$\frac{d}{dt}\bigg|_{t=0} \big(\varphi_i \circ \gamma_1\big)(t) = \frac{d}{dt}\bigg|_{t=0} \big(\varphi_i \circ \gamma_2\big)(t)\,. \tag{2.11}$$

The calculation before shows that if (2.11) holds for one chart (U_i, φ_i) about x_0 it holds for every chart about x_0.

Next, we define on the set $\mathcal{P}_{x_0}(M^m)/\sim$ a scalar multiplication by

$$\alpha[\gamma] := [\gamma_\alpha]\,, \tag{2.12}$$

where $\alpha \in \mathbb{R}$ and $\gamma_\alpha(t) := \gamma(\alpha\, t) \in \mathcal{P}_{x_0}(M^m)$. If $\gamma \sim \tilde{\gamma}$, we obtain that

$$\begin{aligned}
\frac{d}{dt}\bigg|_{t=t_0} \big(\varphi_i \circ \gamma_\alpha\big)(t) &= \frac{d}{dt}\bigg|_{t=t_0} \big(\varphi_i \circ \gamma\big)(\alpha\, t) \\
&\overset{(2.11)}{=} \frac{d}{dt}\bigg|_{t=t_0} \big(\varphi_i \circ \tilde{\gamma}\big)(\alpha\, t) = \frac{d}{dt}\bigg|_{t=t_0} \big(\varphi_i \circ \tilde{\gamma}_\alpha\big)(t)\,,
\end{aligned}$$

showing that the definition (2.12) is independent of the representant of the equivalence class $[\gamma]$.

Moreover, on $\mathcal{P}_{x_0}(M^m)/\sim$ we define an addition by

$$[\gamma_1] + [\gamma_2] := [\gamma_{1+2}]\,, \tag{2.13}$$

where

$$\gamma_{1+2}(t) := \varphi_i^{-1}\left(\varphi_i(x_0) + t\, \frac{d}{dt}\Big|_{t=0} \big(\varphi_i \circ \gamma_1\big)(t) + t\, \frac{d}{dt}\Big|_{t=0} \big(\varphi_i \circ \gamma_2\big)(t)\right) .$$

Note that $\gamma_{1+2} \in \mathcal{P}_{x_0}(M^m)$. We claim that this operation is moreover well-defined, i.e.,

a) independent of the chart φ_i;
b) independent of the representants of the classes $[\gamma_1]$ and $[\gamma_2]$.

The second claim follows from an analogous computation as for the scalar multiplication. For the first claim, we have to show that the path γ_{1+2} expressed in two different charts lies in the same equivalence class, i.e.,

$$\frac{d}{dt}\Big|_{t=0} \big(\varphi_i \circ \gamma_{1+2}\big)(t) = \frac{d}{dt}\Big|_{t=0} \big(\varphi_i \circ \tilde{\gamma}_{1+2}\big)(t) ,$$

where $\tilde{\gamma}_{1+2}$ means the path expressed in another chart φ_j about x_0. Hence, we have to compare the following two expressions

$$\begin{aligned}
&\frac{d}{dt}\Big|_{t=0} \big(\varphi_i \circ \gamma_{1+2}\big)(t) \\
&= \frac{d}{dt}\Big|_{t=0} \left\{\varphi_i\left(\varphi_i^{-1}\left(\varphi_i(x_0) + t\, \frac{d}{dt}\Big|_{t=0} \big(\varphi_i \circ \gamma_1\big)(t) + t\, \frac{d}{dt}\Big|_{t=0} \big(\varphi_i \circ \gamma_2\big)(t)\right)\right)\right\} , \\
&= \frac{d}{dt}\Big|_{t=0} \big(\varphi_i \circ \gamma_1\big)(t) + \frac{d}{dt}\Big|_{t=0} \big(\varphi_i \circ \gamma_2\big)(t) ,
\end{aligned}$$

and

$$\begin{aligned}
&\frac{d}{dt}\Big|_{t=0} \big(\varphi_i \circ \tilde{\gamma}_{1+2}\big)(t) \\
&= \frac{d}{dt}\Big|_{t=0} \left\{\varphi_i\left(\varphi_j^{-1}\left(\varphi_j(x_0) + t\, \frac{d}{dt}\Big|_{t=0} \big(\varphi_j \circ \gamma_1\big)(t) + t\, \frac{d}{dt}\Big|_{t=0} \big(\varphi_j \circ \gamma_2\big)(t)\right)\right)\right\} \\
&= d(\varphi_i \circ \varphi_j^{-1})_{\varphi_j(x_0)}\left(\frac{d}{dt}\Big|_{t=0} \big(\varphi_j \circ \gamma_1\big)(t) + \frac{d}{dt}\Big|_{t=0} \big(\varphi_j \circ \gamma_2\big)(t)\right) \\
&= \frac{d}{dt}\Big|_{t=0} \big(\varphi_i \circ \gamma_1\big)(t) + \frac{d}{dt}\Big|_{t=0} \big(\varphi_i \circ \gamma_2\big)(t) .
\end{aligned}$$

Thus, the two expressions are equal. – Note that the constant path $\gamma \equiv x_0$ gives the neutral element for the addition defined in (2.13).

So far we have shown that $\mathcal{P}_{x_0}(M^m)/\sim$ with the scalar multiplication (2.12) and the addition (2.13) is a vector space. Now, we will see that its dimension is m. – Let (U_i, φ_i) be a chart about x_0 and define the map[2]

$$\begin{aligned}
d\varphi_i : \mathcal{P}_{x_0}(M^m)/\sim\ &\longrightarrow \mathbb{R}^m , \\
[\gamma] &\longmapsto d\varphi_i \cdot [\gamma] := \frac{d}{dt}\Big|_{t=0} \big(\varphi_i \circ \gamma\big)(t) .
\end{aligned} \tag{2.14}$$

[2] The notation used for this map will become clear in Section 2.6.

In order to show that $d\varphi_i$ is an isomorphism, we first observe that if $d\varphi_i \cdot [\gamma] = 0$, then $\varphi_i \circ \gamma$ has vanishing velocity at $t = 0$, which is also the velocity of the constant path at $t = 0$. Therefore, $[\gamma] = 0$ and $d\varphi_i$ is injective. It is also surjective, since for every $X \in \mathbb{R}^n$, we have

$$d\varphi_i\big[\varphi^{-1}(\varphi(x_0) + t\,X)\big] = X\,.$$

This implies that $\mathcal{P}_{x_0}(M^m)/\sim$ has dimension m.

Putting all these results together, we end up with the definition for the tangent space to a manifold M^m.

Definition 2.35. *The **tangent space** $T_{x_0}M^m$ to a C^1-manifold M^m at the point x_0 is defined by*

$$T_{x_0}M^m := \mathcal{P}_{x_0}(M^m)/\sim\,,$$

*which is an m-dimensional vector space. And the **tangent vectors** are equivalence classes of paths $[\gamma] \in \mathcal{P}_{x_0}(M^m)/\sim$.*

Concerning the basis of the tangent space, we introduce the following useful notations[3]: Consider a chart (U, x) with $x : U \longrightarrow \mathbb{R}^m$. Then, for all $p \in M^m$, we denote by

$$\left(\frac{\partial}{\partial x_1}(p), \ldots, \frac{\partial}{\partial x_m}(p)\right) \tag{2.15}$$

the basis of T_pM^m, given by

$$\frac{\partial}{\partial x_i}(p) := \big[\gamma_i\big]\,, \qquad i = 1, \ldots, m\,, \tag{2.16}$$

where $\{e_i\}_{i=1,\ldots,m}$ is the canonical basis of $\mathbb{R}^m$ and

$$\gamma_i := x^{-1}\big(x(p) + t\,e_i\big)\,.$$

The following proposition can be interpreted as motivation for the previous notations.

Proposition 2.36. *Let M^m be a C^k-manifold of dimension m and $U \subset M^m$ open. Moreover, let (U, x) and (U, y) be two charts about $p \in U$. Then the following formula holds for all $p \in U$:*

$$\frac{\partial}{\partial x_k}(p) = \sum_{i=1}^{m} \frac{\partial\big(y_i \circ x^{-1}\big)}{\partial x_k}\big(x(p)\big)\,\frac{\partial}{\partial y_i}(p)\,, \qquad k = 1, \ldots, m\,, \tag{2.17}$$

where the notation $y = (y_1, \ldots, y_m) : U \longrightarrow \mathbb{R}^m$ for the coordinate functions is used. The formula can also be written as

$$\frac{\partial}{\partial x_k} = \frac{\partial y_i}{\partial x_k}\frac{\partial}{\partial y_i}$$

[3] At this stage, we change slightly the notations: The charts are now often denoted by (U, x), instead of (U, φ), and points of a manifold by p, instead of x.

Proof. We can write

$$\begin{aligned}\frac{\partial}{\partial x_k}(p) &= \big[x^{-1}\big(x(p)+t\,e_k\big)\big] \\ &= \Big[y^{-1}\Big(y\circ x^{-1}\big(x(p)+t\,e_k\big)\Big)\Big]\end{aligned}$$

We calculate for the path $y\circ x^{-1}\big(x(p)+t\,e_k\big)$ in $\mathbb{R}^m$, using (1.1),

$$\begin{aligned}\frac{d}{dt}\Big|_{t=0} y\circ x^{-1}\big(x(p)+t\,e_k\big) &= d(y\circ x^{-1})_{x(p)}\cdot e_k \\ &= J_{x(p)}\big(y\circ x^{-1}\big)\cdot e_k = \sum_{i=1}^{m}\frac{\partial\big(y_i\circ x^{-1}\big)}{\partial x_k}\,(x(p))\,e_i\,,\end{aligned}$$

where $J_{x(p)}\big(y\circ x^{-1}\big)$ denotes the Jacobian of the map $y\circ x^{-1}:\mathbb{R}^m\longrightarrow\mathbb{R}^m$ at the point $x(p)$. From this, we deduce that

$$\frac{\partial}{\partial x_k}(p) = \Big[y^{-1}\Big(y(p)+t\sum_{i=1}^{m}\frac{\partial\big(y_i\circ x^{-1}\big)}{\partial x_k}\,(x(p))\,e_i\Big)\Big]\,.$$

By definition, this implies (2.17). □

2.5.3 The Tangent Bundle to a Manifold

Let M^m be a m-dimensional C^k-manifold with $k\geq 2$. From a "set" point of view the **tangent bundle** to M^m is defined by

$$TM^m = \bigcup_{p\in M^m} T_pM^m\,,$$

i.e., by the collection of all tangent vectors at all points of M^m. Moreover, we introduce the following projection map

$$\begin{aligned}\pi: TM^m &\longrightarrow M\,,\\ [\gamma] &\longmapsto \pi([\gamma]) := p\,,\ \text{ if } [\gamma]\in T_pM\,.\end{aligned}\tag{2.18}$$

The point $p\in M^m$ is called the **base point** for $[\gamma]$.

In what follows, we prove that the tangent bundle can be made to a $2m$-dimensional C^{k-1}-manifold. – Let $(U_i,\varphi_i)_{i\in I}$ be an atlas for M^m, and define

$$TU_i := \bigcup_{p\in U_i} T_pM^m = \pi^{-1}(U_i)\,.$$

Obviously, we have that $\bigcup_{i\in I} TU_i = TM^m$. On $TU_i\subset TM^m$, we consider the map

$$\begin{aligned}\Phi_i : TU_i &\longrightarrow \mathbb{R}^m \times \mathbb{R}^m = \mathbb{R}^{2m}\,, \\ [\gamma] &\longmapsto \big(\varphi_i \circ \pi([\gamma]), d\varphi_i \cdot [\gamma]\big)\,. \end{aligned} \tag{2.19}$$

Recall that (see (2.14))

$$d\varphi_i \cdot [\gamma] = \frac{d}{dt}\bigg|_{t=0} \big(\varphi_i \circ \gamma\big)(t)\,.$$

It is clear that the map Φ_i is a bijection from TU_i into $\varphi_i(U_i) \times \mathbb{R}^m$. We then say that $\Omega \subset TM^m$ is *open* in TM^m if and only if for all $i \in I$ the set $\Phi_i(\Omega \cap TU_i)$ is open in $\mathbb{R}^{2m}$.

In order to establish that this defines a topology on the tangent bundle TM^m, we take, in a first step, $(\Omega_k)_{k\in K}$ all open and show that $\bigcup_{k\in K} \Omega_k$ is also open and, in a second step, $\Omega_1, \ldots, \Omega_N$ open and show that $\bigcap_{l=1}^N \Omega_l$ is open. – For all $i \in I$, we have

$$\Phi_i\left(\bigcup_{k\in K} \Omega_k \cap TU_i\right) = \bigcup_{k\in K} \Phi_i(\Omega_k \cap TU_i)\,.$$

Since the right-hand side is open in $\mathbb{R}^{2m}$, it follows by definition of openness for TM^m that $\bigcup_{k\in K} \Omega_k$ is indeed open. Consider now $\Omega_1, \ldots, \Omega_N$ open and since the Φ_i are bijections, we can write, for all $i \in I$,

$$\Phi_i\left(\bigcap_{l=1}^N \Omega_l \cap TU_i\right) = \bigcap_{l=1}^N \Phi_i(\Omega_l \cap TU_i)\,.$$

Therefore, we have a topology on TM^m. In addition, this topology is separated (the proof is straightforward).

Proposition 2.37. *This topology on the tangent bundle TM^m is independent of the atlas $(U_i, \varphi_i)_{i\in I}$ representing the fixed differentiable structure on M^m. Moreover, the maps Φ_i, $i \in I$, defined in (2.19) are homeomorphisms for this topology and the projection π is continuous.*

Proof. Let $(V_j, \psi_j)_{j\in J}$ be another equivalent atlas for M^m and $(TV_j, \Psi_j)_{j\in J}$ as defined before (see (2.19)). Assume that $\Omega \subset TM^m$ is open for this atlas, i.e., $\Psi_j(\Omega \cap TV_j)$ is open in $\mathbb{R}^{2m}$ for all $j \in J$. We then have to prove that $\Phi_i(\Omega \cap TU_i)$ is open in $\mathbb{R}^{2m}$ for all $i \in I$.

We first write

$$\Phi_i(\Omega \cap TU_i) = \bigcup_{j\in J} \Phi_i\Big(\Psi_j^{-1}\big(\Psi_j(\Omega \cap TU_i \cap TV_j)\big)\Big)\,.$$

Take $j \in J$ such that $TU_i \cap TV_j \neq \emptyset$. Since Ψ_j is a bijection, we get

$$\Psi_j(\Omega \cap TU_i \cap TV_j) = \Psi_j(\Omega \cap TV_j) \cap \Psi_j(TU_i \cap TV_j)\,.$$

The first term on the right-hand side is open by assumption. The second equals $\psi_j(U_i \cap V_j) \times \mathbb{R}^m$ being thus also open in $\mathbb{R}^{2m}$. Using the fact that the charts $(U_i, \varphi_i)_{i \in I}$ and $(V_j, \psi_j)_{j \in J}$ are equivalent, we will show that the transition functions

$$\Phi_i \circ \Psi_j^{-1} : \Psi_j(TU_i \cap TV_j) \longrightarrow \Phi_i(TU_i \cap TV_j)$$

are C^{k-1}-maps with $k \geq 2$ (this will be done in the next Proposition 2.38). We then deduce that

$$\Phi_i\Big(\Psi_j^{-1}\big(\Psi_j(\Omega \cap TU_i \cap TV_j)\big)\Big)\,,$$

and hence

$$\bigcup_{j \in J} \Phi_i\Big(\Psi_j^{-1}\big(\Psi_j(\Omega \cap TU_i \cap TV_j)\big)\Big)$$

are open in $\mathbb{R}^{2m}$. This result holds for all $i \in I$, and the topology on TM^m thus only depends on the differentiable structure on M^m. – The remaining assertions are left as an exercise. □

We end up with the following proposition summarizing the results for the tangent bundle.

Proposition 2.38. *Let M^m be a m-dimensional C^k-manifold with $k \geq 2$, and let $(U_i, \varphi_i)_{i \in I}$ be an atlas for M^m. Then $(TU_i, \Phi_i)_{i \in I}$ defines a C^{k-1}-differentiable structure for the tangent bundle TM^m depending only on the differentiable structure on M^m. And TM^m is a C^{k-1}-manifold of dimension $2m$ for this differentiable structure. Moreover, the projection π is a C^{k-1}-submersion for this differentiable structure.*

Proof. Let $(U_i, \varphi_i)_{i \in I}$ be an atlas for M^m and $(TU_i, \Phi_i)_{i \in I}$ defined as before. For $i,\, j \in I$ such that $TU_i \cap TU_j \neq \emptyset$ (note that this also implies that $U_i \cap U_j \neq \emptyset$), we consider the transition functions

$$\begin{aligned} \Phi_j \circ \Phi_i^{-1} : \varphi_i(U_i \cap U_j) \times \mathbb{R}^m \subset \mathbb{R}^{2m} &\longrightarrow \varphi_j(U_i \cap U_j) \times \mathbb{R}^m \subset \mathbb{R}^{2m}\,, \\ (x, \xi) &\longmapsto \Phi_j \circ \Phi_i^{-1}(x, \xi)\,. \end{aligned} \tag{2.20}$$

Using (2.19), we get the following explicit expression for the transition functions (see Fig. 2.11):

$$\begin{aligned} \Phi_j \circ \Phi_i^{-1}(x, \xi) &= \Phi_j\Big(\big[\varphi_i^{-1}(x + t\,\xi)\big]\Big) \\ &= \left(\varphi_j \circ \varphi_i^{-1}(x),\, \frac{d}{dt}\bigg|_{t=0} \big(\varphi_j \circ \varphi_i^{-1}\big)(x + t\,\xi)\right) \\ &= \big(\varphi_j \circ \varphi_i^{-1}(x), d(\varphi_j \circ \varphi_i^{-1})_x \cdot \xi\big)\,. \end{aligned} \tag{2.21}$$

By assumption $(U_i, \varphi_i)_{i \in I}$ defines a C^k-differential structure on M^m, hence the transition functions $\varphi_j \circ \varphi_i^{-1} : \varphi_j(U_i \cap U_j) \longrightarrow \varphi_i(U_i \cap U_j)$ are

C^k-maps and moreover the differentials $d(\varphi_j \circ \varphi_i^{-1})$ are C^{k-1}-maps. Therefore using (2.21), we have that

$$\Phi_j \circ \Phi_i^{-1} \in C^{k-1}(\varphi_i(U_i \cap U_j) \times \mathbb{R}^m, \varphi_j(U_i \cap U_j) \times \mathbb{R}^m)\,.$$

Similarly, we show that $\Phi_i \circ \Phi_j^{-1} = (\Phi_j \circ \Phi_i^{-1})^{-1}$ is a C^k-map, and thus $\Phi_j \circ \Phi_i^{-1}$ is a C^{k-1}-diffeomorphism. Hence, by Definition 2.16 the family $(TU_i, \Phi_i)_{i\in I}$ defines a C^{k-1}-atlas on the tangent bundle. – Note that we loose one degree of regularity.

Now, let $(\psi_j, V_j)_{j\in J}$ be another atlas equivalent to $(U_i, \varphi_i)_{i\in I}$. From the explicit expression (2.21) for the transition functions $\Phi_i \circ \Psi_j^{-1}$, we get that they are C^{k-1}-diffeomorphisms. By Definition 2.17 the systems of charts $(TU_i, \Phi_i)_{i\in I}$ and $(TV_j, \Psi_j)_{j\in J}$ are equivalent, showing that the differentiable structure on the tangent bundle TM^n does *not* depend on the representant of the fixed differentiable structure on the manifold M^m.

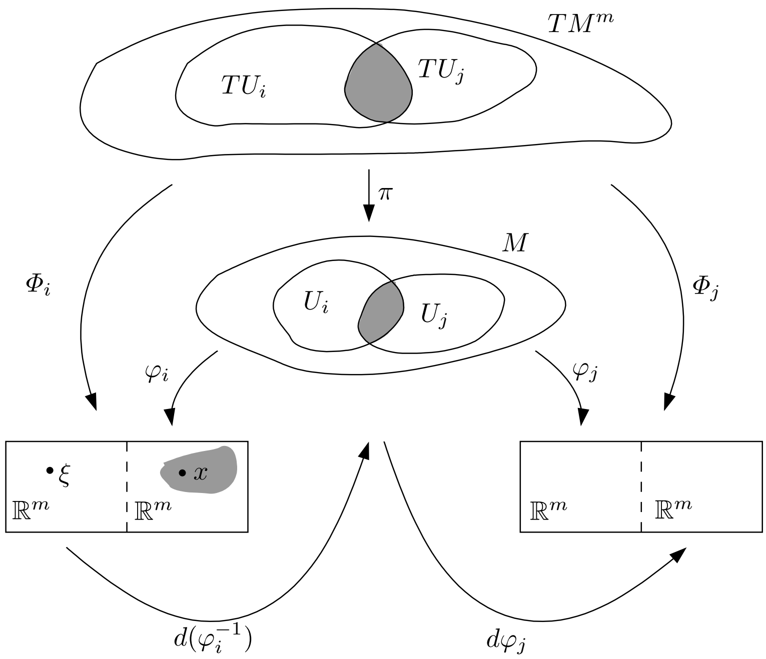

Fig. 2.11. Transition functions for the tangent bundle.

Next, we show that $\pi : TM^m \longrightarrow M$ defined in (2.18) is a C^{k-1}-submersion. – Clearly, the coordinate expression for π is given by

$$\varphi_j \circ \pi \circ \Phi_i^{-1}(x, \xi) = \varphi_j \circ \varphi_i^{-1}(x)\,,$$

where $(x, \xi) \in \varphi_i(U_i) \times \mathbb{R}^m$. Since π is already continuous by Proposition 2.37, we deduce from the last equation that the projection π is also a C^{k-1}-map (see Definition 2.27). Moreover, we have

$$\operatorname{rank} d(\varphi_j \circ \pi \circ \Phi_i^{-1}) = \operatorname{rank} d(\varphi_j \circ \varphi_i^{-1}) = m\,,$$

implying that the differential has maximal rank, since we deal with a m-dimensional target manifold. Hence, the projection is a C^{k-1}-submersion (see Definition 2.30). $\square$

We close this section with a definition which will be useful in the following.

Definition 2.39. *Let M^m be a C^k-manifold with $k \geq 2$ and $s : M^m \longrightarrow TM^m$ a C^l-map, $l \leq k-1$. Then the map s is called a C^l-**section** of the tangent bundle TM^m if it satisfies $\pi \circ s = id_{M^m}$, where π is the projection defined in (2.18).*

2.6 The Tangent Map Between two Manifolds

Now, we come to an important definition, which was already used in a special case in order to show that the tangent bundle TM^m is a differentiable manifold (see (2.14)).

Definition 2.40. *Let M^m and N^n be two C^k-manifolds and let $f : M^m \longrightarrow N^n$ be a C^k-map with $k \geq 2$. The **tangent map** to f is the C^{k-1}-map*

$$\begin{aligned} df : TM^m &\longrightarrow TN^n\,, \\ [\gamma] &\longmapsto df_p \cdot [\gamma] := [f \circ \gamma]\,, \end{aligned}$$

where $\gamma \in C^1([0,1], M^m)$ with $\gamma(0) = p$.

Remark. The tangent map therefore acts on equivalence classes of curves by transporting then with the map itself to the target tangent bundle. In addition, the map $df_p : T_pM^m \longrightarrow T_{f(p)}N^n$ is *linear*. – Note that if $N^n = \mathbb{R}^n$, then the tangent map reduces to

$$df_p \cdot [\gamma] = \frac{d}{dt}\bigg|_{t=0} \big(f \circ \gamma\big)(t)\,.$$

This is exactly what we already introduced in (2.14).

The definition clearly makes sense. – Indeed, if $(U_i, \varphi_i)_{i \in I}$ and $(V_j, \psi_j)_{j \in J}$ are two systems of charts for M^m and N^n, respectively, and if $\gamma \sim \tilde{\gamma}$, where $\tilde{\gamma} \in C^1([0,1], M^m)$ with $\tilde{\gamma}(0) = p$, we obtain (see (2.11))

$$\begin{aligned} \frac{d}{dt}\bigg|_{t=0} \big(\psi_i \circ f \circ \gamma\big)(t) &= \frac{d}{dt}\bigg|_{t=0} \big(\psi_i \circ f \circ \varphi_j^{-1} \circ \varphi_j \circ \gamma\big)(t) \\ &= d\big(\psi_i \circ f \circ \varphi_j^{-1}\big)_{\varphi_j(p)} \frac{d}{dt}\bigg|_{t=0} \big(\varphi_j \circ \gamma\big)(t) \\ &= d\big(\psi_i \circ f \circ \varphi_j^{-1}\big)_{\varphi_j(p)} \frac{d}{dt}\bigg|_{t=0} \big(\varphi_j \circ \tilde{\gamma}\big)(t) \end{aligned}$$

Hence, the desired result $f \circ \gamma \sim f \circ \tilde{\gamma}$.

In a next step, we prove that the tangent map $df : TM^m \longrightarrow TN^n$ is indeed a C^{k-1}-map. – Let $(U_i, \varphi_i)_{i \in I}$ and $(V_j, \psi_j)_{j \in J}$ be two systems of charts for M^m and N^n, respectively, and $(TU_i, \Phi_i)_{i \in I}$ and $(TV_j, \Psi_j)_{j \in J}$ the corresponding systems of charts for TM^m and TN^n, respectively. First, we show that df is continuous.

For $\Omega \subset TN^n$ open, we can write, since Φ_i, $i \in I$, is a bijection[4],

$$(df)^{-1}(\Omega) = \bigcup_{i \in I} \Phi_i^{-1} \Phi_i [(df)^{-1}(\Omega) \cap TU_i] \,, \tag{2.22}$$

and, since Ψ_j, $j \in J$, is also a bijection,

$$\begin{aligned} \Phi_i[(df)^{-1}(\Omega) \cap TU_i] &= \Phi_i \left[(df)^{-1} \left(\bigcup_{j \in J} \Psi_j^{-1} \big(\Psi_j(\Omega \cap TV_j)\big) \right) \cap TU_i \right] \\ &= \Phi_i \left[\bigcup_{j \in J} \Big(\big(\Psi_j \circ df\big)^{-1} \big(\Psi_j(\Omega \cap TV_j)\big) \Big) \cap TU_i \right] \\ &= \bigcup_{j \in J} \Phi_i \left[\big(\Psi_j \circ df\big)^{-1} \big(\Psi_j(\Omega \cap TV_j)\big) \cap TU_i \right] \,. \end{aligned} \tag{2.23}$$

This is equivalent to

$$\bigcup_{j \in J} \Big(\big(\Psi_j \circ df \circ \Phi_i^{-1}\big)^{-1} \big(\Psi_j(\Omega \cap TV_j)\big) \Big) \,.$$

By definition, the map $\Psi_j \circ df \circ \Phi_i^{-1}$ is defined on the set $\varphi_i\big(f^{-1}(V_j) \cap U_i\big) \times \mathbb{R}^m$ which is *open* in $\mathbb{R}^{2m}$. The coordinate expression for the tangent map (see (2.25) below) reads, for $(x, \xi) \in \varphi_i\big(f^{-1}(V_j) \cap U_i\big) \times \mathbb{R}^m \subset \mathbb{R}^{2m}$ open,

$$\big(\Psi_j \circ df \circ \Phi_i^{-1}\big)(x, \xi) = \big(\psi_j \circ f \circ \varphi_i^{-1}(x), d(\psi_j \circ f \circ \varphi_i^{-1})_x \cdot \xi\big) \in \mathbb{R}^{2n} \,, \tag{2.24}$$

The fact that f is by assumption a C^k-map translates to $\psi_j \circ f \circ \varphi_i^{-1} \in C^k(\mathbb{R}^m, \mathbb{R}^n)$. Thus, we deduce that the expression (2.24) is at least continuous since $k \geq 2$.

Moreover, by definition of the topology on TN^n, we deduce that by assumption $\Psi_j(\Omega \cap TV_j)$ is *open* in $\mathbb{R}^{2n}$. From this and using (2.23), we then get the openness of $\Phi_i\big[(df)^{-1}(\Omega) \cap TU_i\big]$. With (2.22), we deduce that $(df)^{-1}(\Omega)$ is open in TM^m and the continuity of the tangent map.

In order to show that df is a C^{k-1}-map, it suffices to prove that the coordinate expression of the tangent map df is a C^{k-1}-map (see Definition 2.27).

[4] Note that the brackets always indicate that we are dealing with elements of the tangent bundle. Moreover, by $(df)^{-1}(\Omega)$ we mean the inverse image from a set point of view, i.e., the set $\{[\gamma] \in TM^m \ : \ df \cdot ([\gamma]) \in \Omega \subset TN^n\}$

– Let $(x, \xi) \in \varphi_i\big(f^{-1}(V_j) \cap U_i\big) \times \mathbb{R}^m$. Then, we claim that the coordinate expression for df reads as

$$\big(\Psi_j \circ df \circ \Phi_i^{-1}\big)(x, \xi) = \big(\psi_j \circ f \circ \varphi_i^{-1}(x), d(\psi_j \circ f \circ \varphi_i^{-1})_x \cdot \xi\big) \in \mathbb{R}^{2n}\,, \quad (2.25)$$

The claim then immediately implies the result.

To show the coordinate expression (2.25) we compute, using the Definition 2.40 of the tangent map,

$$df \circ \Phi_i^{-1}(x, \xi) \overset{(2.19)}{=} df\big([\varphi_i^{-1}(x + t\,\xi)]\big) = \big[f \circ \varphi_i^{-1}(x + t\,\xi)\big]\,.$$

This gives, using again (2.19),

$$\begin{aligned}
\big(\Psi_j \circ df \circ \Phi_i^{-1}\big)(x, \xi) &= \Psi_j\big([f \circ \varphi_i^{-1}(x + t\,\xi)]\big) \\
&= \Big(\psi_j \circ f \circ \varphi_i^{-1}(x), d\psi_j\,\big[f \circ \varphi_i^{-1}(x + t\,\xi)\big]\Big) \\
&= \left(\psi_j \circ f \circ \varphi_i^{-1}(x), \left.\frac{d}{dt}\right|_{t=0} \big(\psi_j \circ f \circ \varphi_i^{-1}\big)(x + t\,\xi)\right) \\
&= \big(\psi_j \circ f \circ \varphi_i^{-1}(x), d(\psi_j \circ f \circ \varphi_i^{-1})_x \cdot \xi\big)\,.
\end{aligned}$$

In summary, the tangent map df is indeed a C^{k-1}-map.

We go a step further and introduce some useful notations for the tangent map acting on the basis (2.15) of the tangent space T_pM^m of M^m at the point p. Let (U, x) a local chart and $f : U \longrightarrow N^n$. Then, for all $p \in U$, the expression (see (2.16))

$$df_p \cdot \frac{\partial}{\partial x_i}(p) = \big[f\big(x^{-1}(x(p) + t\,e_i)\big)\big]\,, \qquad i = 1, \ldots, m\,,$$

is denoted by

$$\frac{\partial f}{\partial x_i}(p) \in T_{f(p)}N^n\,. \quad (2.26)$$

Let (U, y) be another chart, and we set $y = (y_1, \ldots, y_m)$ with $y_i : U \longrightarrow \mathbb{R}$, for $i = 1, \ldots, m$. Then we have, for $k = 1, \ldots, m$,

$$\begin{aligned}
\frac{\partial y_i}{\partial x_k}(p) &\overset{(2.26)}{=} (dy_i)_p \cdot \frac{\partial}{\partial x_k} \\
&= \big[y_i\big(x^{-1}(x(p) + t\,e_k)\big)\big] \\
&= \left.\frac{d}{dt}\right|_{t=0} \Big(y_i\big(x^{-1}(x(p) + t\,e_k)\big)\Big) \\
&= d(y_i \circ x^{-1})_{x(p)} \cdot e_k = \frac{\partial}{\partial x_k}\big(y_i \circ x^{-1}\big)(x(p))\,.
\end{aligned}$$

This shows that we recover the notation introduced in Proposition 2.36. – With these notations objects on manifolds look *locally* like in flat space (see also Exercise 2.42).

Proposition 2.41. *Let $f \in C^k(M^m, N^n)$ and $g \in C^k(N^n, L^l)$. Then $g \circ f \in C^k(M^m, L^l)$ and for the tangent map of the composition the following formula holds:*

$$d(g \circ f) = dg \cdot df\,.$$

Proof. For $[\gamma] \in TM^m$ we compute, using Definition 2.40,

$$\begin{aligned} d(g \circ f)[\gamma] &= [g \circ f(\gamma)] = [g(f \circ \gamma)] \\ &= dg[f \circ \gamma] = dg(df[\gamma])\,. \end{aligned}$$

The fact that $g \circ f \in C^k(M^m, L^l)$ is left as an exercise. □

* * *

Exercises.

Exercise 2.42. Let $f : M^m \longrightarrow \mathbb{R}^n$. Using charts and the lemma of Schwartz, show that for all $i, j = 1, \ldots, m$ we have

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)\,.$$

Exercise 2.43. Let (U, x) be a local chart of M^m. Show that for $p \in U$, we have

$$dx_p \cdot \frac{\partial}{\partial x_i}(p) = e_i\,,$$

where $\{e_i\}_{i=1,\ldots,m}$ denotes the canonical basis of $\mathbb{R}^m$.

2.7 Vector Fields on a Manifold

We already studied vector fields on Euclidean spaces in Section 1.2. Now, we want to study the more general case of vector fields on manifolds . Apart from some additional results, we will recover those of Euclidean spaces. We start with the following basic definition.

Definition 2.44. *Let M^m be a C^k-differentiable manifold of dimension m, $k \geq 2$. A C^l-* ***vector field*** *on M^m with $l \leq k-1$ is a C^l-map*

$$X : M^m \longrightarrow TM^m\,,$$

such that $\pi \circ X = id_{M^m}$, where $\pi : TM^m \longrightarrow M^m$ is the canonical projection of (2.18). Equivalently, a C^l-vector field can be defined as C^l-section of the tangent bundle (see Definition 2.39). Moreover, we denote C^∞-vector fields on M^m by $\mathcal{X}(M)$.

Remark. At each point of the manifold we thus assign a tangent vector in such a way that the dependence on the base point is C^l-smooth. The basis of the tangent spaces introduced in (2.15) gives the standard example for a C^l-vector field.

Let (U, x) be a chart about $p \in M^m$. Then a vector field can be written locally, i.e., in terms of the chart (U, x), as

$$X(p) = \sum_{i=1}^{m} X_i(p) \frac{\partial}{\partial x_i}(p)\,, \tag{2.27}$$

where the $X_i : U \longrightarrow \mathbb{R}$ are functions and $\{\frac{\partial}{\partial x_i}(p)\}_{i=1,\ldots,m}$ denotes the basis of T_pM^m. Note that the functions X_i are uniquely defined by the chart (U, x). The regularity of the functions X_i, $i = 1, \ldots, m$, can be deduced from the following

Proposition 2.45. *Let M^m be a C^k-manifold and let $X : M^m \longrightarrow TM^m$ such that $\pi \circ X = id_{M^m}$. Then the map X is a C^l-vector field with $l \leq k-1$ if and only if, for all charts (U, x), we have that $X_i \in C^l(U, \mathbb{R})$, $i = 1, \ldots, m$, where the functions X_i are given by (2.27).*

Proof. Consider the map

$$dx \circ X \circ x^{-1} : x(U) \subset \mathbb{R}^m \longrightarrow \mathbb{R}^{2m}\,.$$

Hence, the map X is a C^l-vector field, $l \leq k-1$, if and only if, for all charts (U, x), we have that $dx \cdot X \in C^l(U, \mathbb{R}^{2m})$. From Exercise 2.43 and the local expression (2.27), we then deduce by linearity

$$dx_p \cdot X(p) = dx_p \left(\sum_{i=1}^{m} X_i(p) \frac{\partial}{\partial x_i}(p) \right) = \sum_{i=1}^{m} X_i(p)\, e_i\,.$$

Therefore, X is a C^l-vector field if and only if $X_i \in C^l(U, \mathbb{R})$, $i = 1, \ldots, m$, for all charts (U, x). $\square$

We come to two other important definitions.

Definition 2.46. *Let M^m be a C^k-manifold, $U \subset M^m$ open and let X be a C^l-vector field on M^m with $l \leq k-1$. A map*

$$\Gamma^X : I \times U \longrightarrow M^m\,, \qquad I \subset \mathbb{R}\,,$$

solving for all $(t, p) \in I \times U$ the equations

$$\frac{\partial \Gamma^X}{\partial t}(t, p) = X\big(\Gamma^X(t, p)\big)\,, \qquad \Gamma^X(0, p) = p\,, \tag{2.28}$$

*is called the **(local) flow** of the vector field X.*

Remark. Note that (2.28) can be written precisely as, using the notation introduced in (2.26),

$$\frac{\partial \Gamma^X}{\partial t}(t,p) = d\Gamma^X_{(t,p)} \cdot \frac{\partial}{\partial t}(t,p) = X\big(\Gamma^X(t,p)\big)\,.$$

Moreover, we will often write Γ^X_t for $\Gamma^X(t,\cdot)$ with $t \in I$.

Definition 2.47. *Let $f : M^m \longrightarrow N^n$ be a C^k-diffeomorphism and let X be a C^k-vector field on M^m. We define the* ***push-forward*** *of X by f to be the following C^{k-1}-vector field on N^n:*

$$f_*X(\tilde{p}) := df_{f^{-1}(\tilde{p})}X\big(f^{-1}(\tilde{p})\big)\,, \qquad \tilde{p} \in N^n\,. \tag{2.29}$$

Remark. Note that the definition requires the map $f : M^m \longrightarrow N^n$ to be a diffeomorphism, since otherwise the right-hand side of (2.29) is *not* necessarily a vector field. Again we loose one degree of regularity because of the defining tangent map.

The next proposition gives an explicit expression for the flow of the push-forward in terms of the flow of the vector field itself.

Proposition 2.48. *Let $f : M^m \longrightarrow N^n$ be a diffeomorphism and X a vector field on M^m with corresponding flow Γ^X_t. Then, the map*

$$\Gamma^{f_*X}_t := f \circ \Gamma^X_t \circ f^{-1} : N^n \longrightarrow N^n \tag{2.30}$$

is a flow for the push-forward f_*X.

Proof. For $\tilde{p} \in N^n$ we compute, using Proposition 2.41 and the fact that Γ^X_t is a flow for X,

$$\begin{aligned}
\frac{\partial \Gamma^{f_*X}_t}{\partial t}(\tilde{p}) &= \frac{\partial}{\partial t}\big(f \circ \Gamma^X_t \circ f^{-1}\big)(\tilde{p}) \\
&= d\Big(f \circ \Gamma^X_t\big(f^{-1}(\tilde{p})\big)\Big) \cdot \frac{\partial}{\partial t}\big(f^{-1}(\tilde{p})\big) \\
&= df_{\Gamma^X_t(f^{-1}(\tilde{p}))}\Big(d(\Gamma^X_t)_{f^{-1}(\tilde{p})} \cdot \frac{\partial}{\partial t}\big(f^{-1}(\tilde{p})\big)\Big) \\
&= df_{\Gamma^X_t(f^{-1}(\tilde{p}))} \cdot X\Big(\Gamma^X_t\big(f^{-1}(\tilde{p})\big)\Big)\,.
\end{aligned}$$

The right-hand side can be written differently

$$df_{f^{-1}(f\circ\Gamma^X_t\circ f^{-1}(\tilde{p}))} \cdot X\Big(f^{-1}\big(f \circ \Gamma^X_t \circ f^{-1}(\tilde{p})\big)\Big)\,.$$

By Definition 2.47, we just obtain

$$\begin{aligned}
\frac{\partial \Gamma^{f_*X}_t}{\partial t}(\tilde{p}) &= df_{f^{-1}(f\circ\Gamma^X_t\circ f^{-1}(\tilde{p}))}X\Big(f^{-1}\big(f \circ \Gamma^X_t \circ f^{-1}(\tilde{p})\big)\Big) \\
&= f_*X\big(f \circ \Gamma^X_t \circ f^{-1}(\tilde{p})\big)\,,
\end{aligned}$$

showing that $\Gamma_t^{f_*X}$ is the flow for f_*X (since $\Gamma_0^{f_*X}(\tilde{p}) = \tilde{p}$ is obviously verified). □

This result allows us to generalize Theorem 1.20 for $\mathbb{R}^m$.

Theorem 2.49 (Local Existence of a Flow). *Let M^m be a C^{k+1}-manifold and let X be a C^k-vector field on M^m with $k \geq 1$. Then there exist an open neighborhood U of $p \in M^m$ and a C^k-map $\Gamma^X : (-T, T) \times U \longrightarrow M^m$ for some $0 < T \in \mathbb{R}$ such that, for every $(t, p) \in (-T, T) \times U$, we have*

$$\frac{\partial \Gamma^X}{\partial t}(t, p) = X\big(\Gamma^X(t, p)\big)\,, \qquad \Gamma^X(0, p) = p\,.$$

Proof. Let $p \in M^m$ and (U, x) a chart about p. We define $Y := x_*X$, which is a C^k-vector field on $x(U) \subset \mathbb{R}^m$. (Here, we assume that M^m is a C^{k+1}-manifold in order for the push-forward to be a C^k-vector field.) By Theorem 1.20 there exists a C^k-map $\Gamma_t^Y : U' \longrightarrow \mathbb{R}^m$, $U' \subset x(U)$, being the local flow for Y. Using the previous Proposition 2.48 and since $X = (x^{-1})_*Y$, we have that the C^k-map

$$\Gamma_t^X := x^{-1} \circ \Gamma_t^Y \circ x$$

is the local flow for X. □

Remark 2.50. Using Proposition 2.48 in the same way as in the proof of Theorem 2.49, we can extend the following results for $\mathbb{R}^m$ to manifolds:

a) the Uniqueness Theorem for local flows of C^1-vector fields (see Theorem 1.21);
b) the Proposition 1.22 concerning the composition of flows and
c) the Straightening Map Theorem 1.25.

For *compact* manifolds, we have also global existence.

Theorem 2.51 (Global Existence of a Flow). *Let M^m be a compact C^{k+1}-manifold with $k \geq 1$ and let X be a C^k-vector field. Then there exists a global flow for X, i.e., a C^k-map $\Gamma^X : \mathbb{R} \times M^m \longrightarrow M^m$ such that, for every $(t, p) \in \mathbb{R} \times M^m$, we have*

$$\frac{\partial \Gamma^X}{\partial t}(t, p) = X\big(\Gamma^X(t, p)\big)\,, \qquad \Gamma^X(0, p) = p\,.$$

Moreover, the map Γ_t^X is a C^k-diffeomorphism from M^m into M^m, for all $t \in \mathbb{R}$.
Proof. □

2.7.1 The Bracket of two Vector Fields on Manifolds

Definition 2.52. *Let X and Y be two C^k-vector fields on a C^{k+1}-manifold M^m with $k \geq 1$. We define the* **bracket** *of X and Y to be the following C^{k-1}-vector field:*

$$[X,Y](p) = \left.\frac{d}{dt}\right|_{t=0} \big((\Gamma_t^Y)_* X\big)(p)\,, \qquad p \in M^m\,. \tag{2.31}$$

Remark. Again we loose one degree of regularity (see Definition 2.47).

Recall that we already know that for small enough "time" t the flow Γ_t^Y locally exists and is unique. Moreover, observe that

$$\begin{aligned}\left.\frac{d}{dt}\right|_{t=0} \big((\Gamma_t^Y)_* X\big)(p) &= \lim_{t\longrightarrow 0} \frac{\big((\Gamma_t^Y)_* X\big)(p) - \big((\Gamma_0^Y)_* X\big)(p)}{t} \\ &= \lim_{t\longrightarrow 0} \frac{\big((\Gamma_t^Y)_* X\big)(p) - X(p)}{t}\,,\end{aligned}$$

and using the Definition 2.47 of the push-forward, we arrive at

$$[X,Y](p) = \lim_{t\longrightarrow 0} \frac{1}{t}\left(\big(d\Gamma_t^Y\big)_{\Gamma_{-t}^Y(p)} \cdot X\big(\Gamma_{-t}^Y(p)\big) - X(p)\right)\,. \tag{2.32}$$

With these equivalent expressions for the bracket, it is clear that its definition makes sense, since we compare vectors in the *same* tangent plane, i.e., with the *same* base point. When this is not the case anymore, we need the theory of connections which we will treat later. – Note that the bracket $[X,Y]$ is often also denoted by $L_Y X$ being the **Lie derivative** of X in the direction Y (see Section 9.92 for more details).

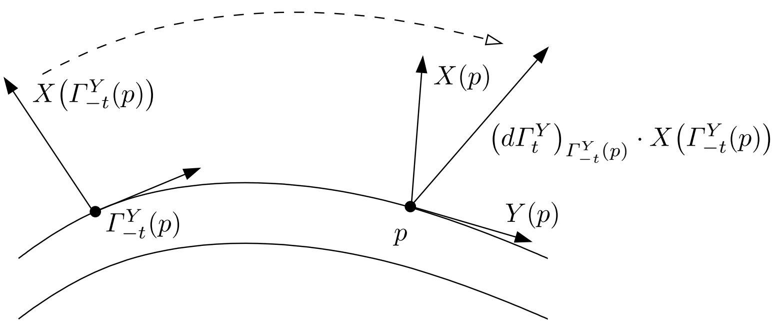

Fig. 2.12. The bracket of two vector fields.

The following proposition gives an explicit formula for the bracket in a local chart.

Proposition 2.53. *Let M^m be a C^{k+1}-manifold and (U,x) be a local chart. Moreover, let X and Y be two C^k-vector fields on U, written for $p \in U$ as*

$$X(p) = \sum_{i=1}^m X_i(p) \frac{\partial}{\partial x_i}(p)\,, \qquad Y(p) = \sum_{i=1}^m Y_i(p) \frac{\partial}{\partial x_i}(p)\,.$$

Then, we have for the bracket

$$[X,Y](p) = \sum_{i=1}^m \sum_{j=1}^m \left(X_j(p) \frac{\partial Y_i}{\partial x_j}(p) - Y_j(p) \frac{\partial X_i}{\partial x_j}(p) \right) \frac{\partial}{\partial x_i}(p)\,. \tag{2.33}$$

Proof. For all $f \in C^{k+1}(M^m, \mathbb{R})$, we have

$$\begin{aligned} df_p \cdot \big((\Gamma_t^Y)_* X\big)(p) &= df_p \cdot \Big(\big(d\Gamma_t^Y\big)_{\Gamma_{-t}^Y(p)} \cdot X\big(\Gamma_{-t}^Y(p)\big)\Big) \\ &= d\big(f \circ \Gamma_t^Y\big)_{\Gamma_{-t}^Y(p)} \cdot X\big(\Gamma_{-t}^Y(p)\big)\,. \end{aligned} \tag{2.34}$$

Now, we expand $f \circ \Gamma_t^Y(p) : \mathbb{R} \longrightarrow \mathbb{R}$ with respect to the "time" t:

$$\begin{aligned} f \circ \Gamma_t^Y(p) &= f \circ \Gamma_0^Y(p) + \int_0^1 \frac{d}{ds}\big(f \circ \Gamma_{ts}^Y\big)(p)\, ds \\ &= f(p) + t\, H(t,p)\,, \end{aligned}$$

where

$$H(t,p) := \int_0^1 df_{\Gamma_{ts}^Y(p)} \cdot \frac{\partial \Gamma_{ts}^Y}{\partial s}(p)\, ds\,,$$

and $H(0,p) = df_p \cdot Y(p)$. Inserting this expansion in (2.34), we get

$$df_p \cdot \big((\Gamma_t^Y)_* X\big)(p) = df_{\Gamma_{-t}^Y(p)} \cdot X\big(\Gamma_{-t}^Y(p)\big) + t\, d\big(H(t,\cdot)\big)_{\Gamma_{-t}^Y(p)} \cdot X\big(\Gamma_{-t}^Y(p)\big)\,.$$

From this we obtain, with the definition of the bracket,

$$\begin{aligned} df_p \cdot [X,Y](p) &= \left.\frac{d}{dt}\right|_{t=0} \Big(df_p \cdot \big((\Gamma_t^Y)_* X\big)(p)\Big) \\ &= \left.\frac{d}{dt}\right|_{t=0} \Big(df_{\Gamma_{-t}^Y(p)} \cdot X\big(\Gamma_{-t}^Y(p)\big)\Big) \\ &\quad + \left.\frac{d}{dt}\right|_{t=0} \Big(t\, d\big(H(t,\cdot)\big)_{\Gamma_{-t}^Y(p)} \cdot X\big(\Gamma_{-t}^Y(p)\big)\Big)\,. \end{aligned} \tag{2.35}$$

The first term equals (chain rule)

$$-d\big(df \cdot X\big)_p \cdot Y(p)\,, \tag{2.36}$$

where the definition of the flow is used. The second term becomes (product rule)

$$d\big(H(0,\cdot)\big)_{\Gamma_0^Y(p)} \cdot X\left(\Gamma_0^Y(p)\right) = d\big(df \cdot Y\big)_p \cdot X(p)\,. \tag{2.37}$$

Inserting (2.36) and (2.37) into (2.35), it follows that

$$df_p \cdot [X,Y](p) = -d\big(df \cdot X\big)_p \cdot Y(p) + d\big(df \cdot Y\big)_p \cdot X(p)\,. \tag{2.38}$$

Let (U, x) be a local chart about $p \in M^m$ and write $x = (x_1, \ldots, x_m) : U \longrightarrow \mathbb{R}^m$. We set $f = x_i \in C^{k+1}(U, \mathbb{R})$, and deduce from (2.38) that

$$(dx_i)_p \cdot [X,Y](p) = (dY_i)_p \cdot X(p) - (dX_i)_p \cdot Y(p)\,,$$

where we used that $(dx_i)_p \cdot X(p) = X_i(p)$ and $(dx_i)_p \cdot Y(p) = Y_i(p)$ since $(dx_i)_p \cdot \frac{\partial}{\partial x_j}(p) = \delta_{ij}$ (see Exercise 2.43). Using again the local representations for X, Y and the notation introduced in (2.26), we end up with (remember also the remark after Definition 2.40)

$$\begin{aligned}(dx_i)_p \cdot [X,Y](p) &= (dY_i)_p \cdot \left(\sum_{j=1}^m X_j(p)\frac{\partial}{\partial x_j}(p)\right) - (dX_i)_p \cdot \left(\sum_{j=1}^m Y_j(p)\frac{\partial}{\partial x_j}(p)\right)\\ &= \sum_{j=1}^m \left(X_j(p)\frac{\partial Y_i}{\partial x_j}(p) - Y_j(p)\frac{\partial X_i}{\partial x_j}(p)\right)\,. \end{aligned} \tag{2.39}$$

This implies the desired formula (2.33). □

As a direct consequence of the last proposition, we get the

Corollary 2.54. *With the usual notations the bracket of two vector fields is anti-symmetric, i.e.,*

$$[X,Y] = -[Y,X]\,.$$

Proposition 2.55. *Let $\varphi : M^m \longrightarrow N^n$ be a diffeomorphism and let X, Y be two vector fields on the manifold M^m. Then we have*

$$\varphi_*[X,Y] = [\varphi_* X, \varphi_* Y]\,. \tag{2.40}$$

Proof. For the proof it is useful to have the following composition in mind:

$$N^n \xrightarrow{\varphi^{-1}} M^m \xrightarrow{\Gamma_t^Y} M^m \xrightarrow{\varphi} N^n\,.$$

We then compute, for $p \in N^n$,

$$\begin{aligned}\varphi_*\big((\Gamma_t^Y)_* X\big)(p) &= \varphi_*\Big(\big(d\Gamma_t^Y\big)_{\Gamma_{-t}^Y(\cdot)} \cdot X\left(\Gamma_{-t}^Y(\cdot)\right)\Big)(p)\\ &= d\varphi_{\varphi^{-1}(p)} \circ \big(d\Gamma_t^Y\big)_{\Gamma_{-t}^Y(\varphi^{-1}(p))} \cdot X\left(\Gamma_{-t}^Y(\varphi^{-1}(p))\right)\\ &= d\big(\varphi \circ \Gamma_t^Y\big)_{\Gamma_{-t}^Y \circ \varphi^{-1}(p)} \cdot X\left(\Gamma_{-t}^Y \circ \varphi^{-1}(p)\right)\,.\end{aligned}$$

This right-hand side can also be written as

$$d\big(\varphi \circ \Gamma_t^Y \circ \varphi^{-1}\big)_{\varphi \circ \Gamma_{-t}^Y \circ \varphi^{-1}(p)} \circ d\varphi_{\Gamma_{-t}^Y \circ \varphi^{-1}(p)} \cdot X\left(\Gamma_{-t}^Y \circ \varphi^{-1}(p)\right). \tag{2.41}$$

We remember that $\Gamma_t^{\varphi_* Y} = \varphi \circ \Gamma_t^Y \circ \varphi^{-1}$ (see Proposition 2.48). Thus (2.41) reads as

$$\varphi_*\big((\Gamma_t^Y)_* X\big)(p) = \left(d\Gamma_t^{\varphi_* Y}\right)_{\Gamma_{-t}^{\varphi_* Y}(p)} \circ d\varphi_{\Gamma_{-t}^Y \circ \varphi^{-1}(p)} \cdot X\left(\Gamma_{-t}^Y \circ \varphi^{-1}(p)\right),$$

or equivalently, for the right-hand side,

$$\left(d\Gamma_t^{\varphi_* Y}\right)_{\Gamma_{-t}^{\varphi_* Y}(p)} \circ d\varphi_{\varphi^{-1}\left((\varphi \circ \Gamma_{-t}^Y \circ \varphi^{-1})(p)\right)} \cdot X\left(\varphi^{-1}\big(\varphi \circ \Gamma_{-t}^Y \circ \varphi^{-1}(p)\big)\right).$$

Using again Proposition 2.48, we arrive at

$$\begin{aligned}
\varphi_*\big((\Gamma_t^Y)_* X\big)(p) &= \left(d\Gamma_t^{\varphi_* Y}\right)_{\Gamma_{-t}^{\varphi_* Y}(p)} \circ d\varphi_{\varphi^{-1}\left((\Gamma_{-t}^{\varphi_* Y})(p)\right)} \cdot X\left(\varphi^{-1}\big(\Gamma_{-t}^{\varphi_* Y}(p)\big)\right) \\
&= \left(d\Gamma_t^{\varphi_* Y}\right)_{\Gamma_{-t}^{\varphi_* Y}(p)} \cdot \varphi_* X\left(\Gamma_{-t}^{\varphi_* Y}(p)\right).
\end{aligned} \tag{2.42}$$

For later use, we use again Definition (2.47) of the push-forward to obtain

$$\varphi_*\big((\Gamma_t^Y)_* X\big)(p) = (\Gamma_t^{\varphi_* Y})_*\big(\varphi_* X\big)(p). \tag{2.43}$$

Taking the derivative with respect to the "time" t in (2.42), leads to

$$\begin{aligned}
\left.\frac{d}{dt}\right|_{t=0} \varphi_*\big((\Gamma_t^Y)_* X\big)(p) &= \left.\frac{d}{dt}\right|_{t=0} \left(d\Gamma_t^{\varphi_* Y}\right)_{\Gamma_{-t}^{\varphi_* Y}(p)} \cdot \varphi_* X\left(\Gamma_{-t}^{\varphi_* Y}(p)\right) \\
&= [\varphi_* X, \varphi_* Y](p),
\end{aligned}$$

where the Definition 2.52 for the bracket is used. □

Remark. Concerning the regularity of the vector field in (2.40), the bracket and the push-forward both reduce the degree of regularity by one. And this obviously holds for both sides of the equality (2.40).

Proposition 2.56 (Jacobi Identity). *Let M^m be a C^{k+1}-manifold and let X, Y, Z be three C^k-vector fields on M^m with $k \geq 2$. Then we have*

$$\big[X,[Y,Z]\big] + \big[Z,[X,Y]\big] + \big[Y,[Z,X]\big] = 0. \tag{2.44}$$

Proof. First, we define the C^{k-1}-vector field W on M^m to be the bracket $[X,Y]$ of X and Y. By Definition 2.52, we then have

$$\begin{aligned}
[W,Z](p) &= \left.\frac{d}{dt}\right|_{t=0} \big((\Gamma_t^Z)_* W\big)(p) \\
&= \left.\frac{d}{dt}\right|_{t=0} \big((\Gamma_t^Z)_* [X,Y]\big)(p).
\end{aligned} \tag{2.45}$$

On the other hand, Proposition 2.55 implies that

$$(\Gamma_t^Z)_*[X,Y] = \big[(\Gamma_t^Z)_*X, (\Gamma_t^Z)_*Y\big](p)\,.$$

Inserting this into (2.45) leads to

$$\big[[X,Y],Z\big](p) = \frac{d}{dt}\bigg|_{t=0} \big[(\Gamma_t^Z)_*X, (\Gamma_t^Z)_*Y\big](p)\,. \tag{2.46}$$

Differentiating the right-hand side like a product and using again the definition of the bracket, we end up with

$$\big[[X,Y],Z\big](p) = \big[[X,Z],Y\big](p) + \big[X,[Y,Z]\big](p)\,.$$

Due to the anti-symmetry of the bracket, this is equivalent to (2.44). □

The bracket of two vector fields measures the extent to which their flows fail to commute.

Proposition 2.57. *Let M^m be a manifold and let X, Y be two vector fields on M^m with corresponding flows Γ_t^X and Γ_s^Y. Then we have that $[X,Y](p) = 0$, for all $p \in M^m$, if and only if the corresponding flows commute, i.e.,*

$$\Gamma_t^X \circ \Gamma_s^Y = \Gamma_s^Y \circ \Gamma_t^X\,, \tag{2.47}$$

for all times for which the flows exist.

Proof. From Theorem 2.49 the flows of the two vector fields X, Y exist and are (local) diffeomorphisms (see Corollary 1.23 which can be extended to manifolds). – First, we show that

$$(\Gamma_t^X)_*[X,Y] = \frac{d}{dt}(\Gamma_t^X)_*Y\,. \tag{2.48}$$

Indeed, since $\Gamma_{t+s}^X = \Gamma_t^X \circ \Gamma_s^X$ (see Proposition 1.22), we deduce from the chain rule that

$$(\Gamma_{t+s}^X)_*Y = (\Gamma_t^X)_*(\Gamma_s^X)_*Y\,.$$

Then, it follows

$$\frac{d}{dt}(\Gamma_t^X)_*Y = \frac{d}{ds}\bigg|_{s=0}(\Gamma_{t+s}^X)_*Y = (\Gamma_t^X)_*\,\frac{d}{ds}\bigg|_{s=0}(\Gamma_s^X)_*Y\,.$$

Using Definition 2.52 of the bracket, we obtain

$$\frac{d}{dt}(\Gamma_t^X)_*Y = (\Gamma_t^X)_*[X,Y]\,,$$

as claimed.

Now, we assume that $[X, Y] = 0$. Then from (2.48) and Proposition 2.48, we deduce that

$$\Gamma_s^Y = \Gamma_s^{(\Gamma_t^X)_* Y} = \Gamma_t^X \circ \Gamma_s^Y \circ \Gamma_{-t}^X ,$$

so that (2.47) follows.

Conversely, assume that the flows commute, i.e., $\Gamma_t^X \circ \Gamma_s^Y = \Gamma_s^Y \circ \Gamma_t^X$, or equivalently,

$$\Gamma_t^X \circ \Gamma_s^Y \circ \Gamma_{-t}^X = \Gamma_s^Y .$$

Applying again Proposition 2.48 to Γ_t^X, we deduce that $(\Gamma_t^X)_* Y = Y$. Hence, by (2.48) the bracket of X and Y vanishes. □

The next theorem gives a geometric interpretation of the bracket.

Theorem 2.58 ("Double" Straightening Theorem). *Let M^m be a manifold and let X, Y be two C^1-vector fields on M^m. Moreover, assume that there exists $\tilde{p} \in M^m$ such that $X(\tilde{p})$ and $Y(\tilde{p})$ are linearly independent. Then the following assertions are equivalent:*

(i) There exists a local chart (U, y) about $\tilde{p}$ such that, for all $p \in U$, we have

$$X(p) = \frac{\partial}{\partial y_1}(p) \quad \textit{and} \quad Y(p) = \frac{\partial}{\partial y_2}(p) .$$

(ii) For all $p \in U$ the bracket vanishes, i.e., $[X, Y](p) = 0$.

Proof. For (i) implies (ii): We apply the local coordinate expression (2.33) for the bracket to $X(p) = \frac{\partial}{\partial y_1}(p)$ and $Y(p) = \frac{\partial}{\partial y_2}(p)$, $p \in U$. A straightforward computation then shows that $[X, Y](p) = 0$, for all $p \in U$.

For (ii) implies (i): Let (U', x) be a local chart about $\tilde{p} \in M^m$ such that $\tilde{p} = 0$. (Note that by a slight abuse of notation we identify points $(x_1, \ldots, x_m) \in \mathbb{R}^m$ with points $p \in U'$ via the chart.) After applying a linear transformation, we may assume that

$$X(0) = \frac{\partial}{\partial x_1}(0) \quad \text{and} \quad Y(0) = \frac{\partial}{\partial x_2}(0) .$$

This is possible since $X(0)$ and $Y(0)$ are linearly independent by assumption.

By Theorem 2.49, there exist local flows Γ_t^X and Γ_s^Y for the C^1-vector fields X, respectively, Y. Consider now the map

$$\begin{aligned} \phi : \tilde{U} \subset \mathbb{R}^m &\longrightarrow M^m , \\ (x_1, x_2, x') &\longmapsto \phi(x_1, x_2, x') := \Gamma_{x_1}^X \Gamma_{x_2}^Y \big((0, 0, x')\big) , \end{aligned}$$

where we write $x = (x_1, x_2, x') = (x_1, x_2, x_3, \ldots, x_m) \in \tilde{U}$ and $\tilde{U}$ is prescribed by the local flows Γ_t^X and Γ_s^Y. From Theorem 2.49 and Corollary 1.23, which can be extended to manifolds, we deduce that $\phi \in C^1(\mathbb{R}^m, M)$.

In a next step, we calculate the differential of ϕ at $(0, 0, 0) \in \mathbb{R}^m$. Using Definition 2.46 of the local flow, it follows

$$d\phi_{(0,0,0)} \cdot e_1 = \frac{d}{dh}\bigg|_{h=0} \Gamma_h^X \Gamma_0^Y\big((0,0,0)\big) = X(0) = \frac{\partial}{\partial x_1}(0)\,.$$

It is important to note at this stage that since $[X, Y] = 0$, the corresponding flows Γ_t^X and Γ_s^Y commute by Proposition 2.57. Hence, we get

$$d\phi_{(0,0,0)} \cdot e_2 = \frac{d}{dh}\bigg|_{h=0} \Gamma_0^X \Gamma_h^Y\big((0,0,0)\big) = \frac{d}{dh}\bigg|_{h=0} \Gamma_h^Y \Gamma_0^X\big((0,0,0)\big) = Y(0) = \frac{\partial}{\partial x_2}(0)\,.$$

Moreover, for $i = 3, \ldots, m$, we have

$$d\phi_{(0,0,0)} \cdot e_i = \frac{d}{dh}\bigg|_{h=0} \Gamma_0^X \Gamma_0^Y\big((0,0,h\,e_i)\big) = \frac{d}{dh}\bigg|_{h=0} h\,e_i = \frac{\partial}{\partial x_i}(0)\,.$$

(Note again that we write $h\,e_i$ for the point in M^m corresponding to $h\,e_i \in \mathbb{R}^m$ via the chart.) In summary, we see that $d\phi_{(0,0,0)}$ is an isomorphism and thus, by the Local Inversion Theorem 1.10, the map ϕ is a local C^1-diffeomorphism, say, onto $U \subset M^m$ open.

Then, we define $\varphi = \phi^{-1}$. For $x = (x_1, x_2, x') \in \phi^{-1}(U)$, we compute, using Proposition 1.22 extended to manifolds,

$$\begin{aligned} d\phi_{(x_1,x_2,x')} \cdot e_1 &= \frac{d}{dh}\bigg|_{h=0} \Gamma_{x_1+h}^X \Gamma_{x_2}^Y\big((0,0,x')\big) = \frac{d}{dh}\bigg|_{h=0} \Gamma_h^X \Gamma_{x_1}^X \Gamma_{x_2}^Y\big((0,0,x')\big) \\ &= \frac{d}{dh}\bigg|_{h=0} \Gamma_h^X\big(\phi(x)\big) = X\big(\phi(x)\big)\,, \end{aligned}$$

and similarly, because the flows Γ_t^X and Γ_s^Y commute by assumption,

$$\begin{aligned} d\phi_{(x_1,x_2,x')} \cdot e_2 &= \frac{d}{dh}\bigg|_{h=0} \Gamma_{x_1}^X \Gamma_{x_2+h}^Y\big((0,0,x')\big) = \frac{d}{dh}\bigg|_{h=0} \Gamma_{x_1}^X \Gamma_h^Y \Gamma_{x_2}^Y\big((0,0,x')\big) \\ &= \frac{d}{dh}\bigg|_{h=0} \Gamma_h^Y \Gamma_{x_1}^X \Gamma_{x_2}^Y\big((0,0,x')\big) = Y\big(\phi(x)\big)\,. \end{aligned}$$

This completes the proof, since $\varphi = (y_1, \ldots, y_m) : U \subset M^m \longrightarrow \mathbb{R}^m$ gives rise to local coordinates on U satisfying, for all $p \in U$,

$$X(p) = \frac{\partial}{\partial y_1}(p) \quad \text{and} \quad Y(p) = \frac{\partial}{\partial y_2}(p)\,.$$

□

Remark. Note the similarity in the proofs of the two "Straightening Theorems" 1.25 and 2.58.

3 Differential Calculus on Manifolds

3.1 Differential Calculus on $\mathbb{R}^n$

3.1.1 Skew-symmetric Forms on $\mathbb{R}^n$

Definition 3.1. *A* ***skew-symmetric*** *or* ***alternating*** p***-form*** α *on* $\mathbb{R}^n$ *is a map*

$$\alpha : \underbrace{\mathbb{R}^n \times \ldots \times \mathbb{R}^n}_{p-\mathit{times}} \longrightarrow \mathbb{R}$$

such that

(i) it is linear in each factor, i.e., with λ, $\mu \in \mathbb{R}$ we have, for all $k = 1, \ldots, p$,

$$\begin{aligned} \alpha(v_1, \ldots, \lambda\, v_k + \mu\, w_k, \ldots, v_p) = \lambda\, \alpha(v_1, \ldots, v_k, \ldots, v_p) \\ +\mu\, \alpha(v_1, \ldots, w_k, \ldots, v_p)\,, \end{aligned} \tag{3.1}$$

where $v_1, \ldots, v_p \in \mathbb{R}^n$ and $w_k \in \mathbb{R}^n$;

(ii) it is skew-symmetric, i.e., for all permutations $\sigma \in \mathcal{S}_p$ of $\{1, \ldots, p\}$, we have

$$\alpha(v_{\sigma(1)}, \ldots, v_{\sigma(p)}) = (-1)^{|\sigma|}\, \alpha(v_1, \ldots, v_p)\,, \tag{3.2}$$

where $(-1)^{|\sigma|}$ denotes the signature of the permutation with $|\sigma| = p - k$, k being the number of orbits.

Moreover, we denote the set of skew-symmetric p-forms on $\mathbb{R}^n$ by $\bigwedge^p \mathbb{R}^n$.

We use the following conventions:

$$\begin{aligned} \bigwedge\nolimits^0 \mathbb{R}^n &= \mathbb{R}\,, \\ \bigwedge\nolimits^1 \mathbb{R}^n &= (\mathbb{R}^n)^*\,. \end{aligned}$$

Lemma 3.2. *The set $\bigwedge^p \mathbb{R}^n$ has the following properties:*

(i) If there exist $i \neq j$ such that $v_i = v_j \in \mathbb{R}^n$, then $\alpha(v_1, \ldots, v_p) = 0$.

(ii) For $p > n$, we have $\bigwedge^p \mathbb{R}^n = \{0\}$.

(iii) The set $\bigwedge^p \mathbb{R}^n$ *is an* $\mathbb{R}$*-vector space.*

Proof. For (i): Let $\sigma \in \mathcal{S}_p$ such that $\sigma(i) = j$, $\sigma(j) = i$ and $\sigma(k) = k$ for $k \neq i, j$. From (3.2) and the fact that $v_i = v_j$ by assumption, we deduce

$$\begin{aligned}\alpha(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_p) &= -\alpha(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_p) \\ &= -\alpha(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_p)\,.\end{aligned}$$

This clearly implies $\alpha(v_1, \ldots, v_p) = 0$.

For (ii): Let $v_1, \ldots, v_p \in \mathbb{R}^n$ with $p > n$. Then there exists v_i such that $v_i = \sum_{j \neq i} \lambda_j v_j$ for some $\lambda_j \in \mathbb{R}$. Assuming, for instance, $i = p$ we compute, using (i),

$$\begin{aligned}\alpha(v_1, \ldots, v_p) &= \alpha\left(v_1, \ldots, v_{p-1}, \sum_{j \neq i} \lambda_j v_j\right) \\ &= \sum_{j \neq i} \lambda_j\, \alpha(v_1, \ldots, v_{p-1}, v_j) = 0\,.\end{aligned}$$

For (iii): The vector space structure for $\bigwedge^p \mathbb{R}^n$ is straightforward. □

Example 3.3. The map $\alpha : \mathbb{R}^3 \longrightarrow \mathbb{R}$, $v = (v^1, v^2, v^3) \longmapsto v^1$ is a skew-symmetric 1-form on $\mathbb{R}^3$, and the map $\alpha : \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}$, $(v, w) \longmapsto v^2 w^3 - v^3 w^2$ is a skew-symmetric 2-form on $\mathbb{R}^3$. – Note that the determinant is the standard example for an alternating n-form on $\mathbb{R}^n$.

With the following operation we can pass from one "degree" of a skew-symmetric form to another.

Definition 3.4. *Let* $\alpha \in \bigwedge^p \mathbb{R}^n$ *and* $\beta \in \bigwedge^q \mathbb{R}^n$*. We define the* ***wedge*** *or* ***exterior product*** *of* α *and* β*, denoted by* $\alpha \wedge \beta \in \bigwedge^{p+q} \mathbb{R}^n$*, in the following way:*

$$\alpha \wedge \beta(v_1, \ldots, v_{p+q}) = \sum_{\sigma \in \mathcal{S}_{p+q}} \frac{(-1)^{|\sigma|}}{p!q!}\, \alpha(v_{\sigma(1)}, \ldots, v_{\sigma(p)})\, \beta(v_{\sigma(p+1)}, \ldots, v_{\sigma(p+q)})\,. \tag{3.3}$$

Remark. It is easy to check that $\alpha \wedge \beta$ is indeed a skew-symmetric $(p+q)$-form. Note also that for $\alpha \in \bigwedge^0 \mathbb{R}^n$ the exterior product reduces to usual scalar multiplication.

Example 3.5. Let $\alpha \in \bigwedge^1 \mathbb{R}^n$ and $\beta \in \bigwedge^1 \mathbb{R}^n$, then we have

$$\alpha \wedge \beta(v, w) = \alpha(v)\beta(w) - \alpha(w)\beta(v)\,. \tag{3.4}$$

In particular, let $\{e_i\}_{i=1,\ldots,n}$ denote the canonical basis of $\mathbb{R}^n$. Then we denote by $\{e_i^*\}_{i=1,\ldots,n}$ the dual basis of $\bigwedge^1 \mathbb{R}^n$ defined, for $i = 1, \ldots, n$, by

$$e_i^*(v) := v^i \,, \qquad v = \sum_{j=1}^{n} v^j \, e_j \,.$$

From this and (3.4), we see that

$$e_i^* \wedge e_j^*(v, w) = e_i^*(v) e_j^*(w) - e_i^*(w) e_j^*(v) = v^i \, w^j - w^i \, v^j \,. \tag{3.5}$$

Proposition 3.6. *Let $\alpha \in \bigwedge^p \mathbb{R}^n$ and $\beta \in \bigwedge^q \mathbb{R}^n$. Then the exterior product*

a) is "anti-commutative", i.e.,

$$\alpha \wedge \beta = (-1)^{pq} \beta \wedge \alpha \,;$$

b) is associative, i.e., if $\gamma \in \bigwedge^r \mathbb{R}^n$, then

$$(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma) \,;$$

c) is distributive, i.e., if $\gamma \in \bigwedge^q \mathbb{R}^n$, then

$$\alpha \wedge (\beta + \gamma) = \alpha \wedge \beta + \alpha \wedge \beta \,.$$

Proof. □

As an application of the previous proposition, (3.5) extends to

$$e_{i_1}^* \wedge \ldots \wedge e_{i_p}^*(v_1, \ldots, v_p) = \sum_{\sigma \in \mathcal{S}_p} (-1)^{|\sigma|} v_1^{i_{\sigma(1)}} \ldots v_p^{i_{\sigma(p)}} \,. \tag{3.6}$$

This result has an important consequence.

Proposition 3.7. *Let $p \leq n$ and $\{e_i\}_{i=1,\ldots,n}$ the canonical basis of $\mathbb{R}^n$. Then the set*

$$\{e_{i_1}^* \wedge \ldots \wedge e_{i_p}^*\}_{1 \leq i_1 < \ldots < i_p \leq n}$$

form a basis of $\bigwedge^p \mathbb{R}^n$. And $\bigwedge^p \mathbb{R}^n$ is thus of dimension $C_n^p = \frac{n!}{p!(n-p)!}$.

Proof. Let $e_{k_1}, \ldots, e_{k_p} \in \mathbb{R}^n$ be p basis vectors. Using (3.6), we then obtain

$$e_{i_1}^* \wedge \ldots \wedge e_{i_p}^*(e_{k_1}, \ldots, e_{k_p}) = \sum_{\sigma \in \mathcal{S}_p} (-1)^{|\sigma|} \delta_{k_1}^{i_{\sigma(1)}} \ldots \delta_{k_p}^{i_{\sigma(p)}} \,.$$

This expression is different from zero if and only if $\{i_1, \ldots, i_p\} = \{k_1, \ldots, k_p\}$, showing that $\{e_{i_1}^* \wedge \ldots \wedge e_{i_p}^*\}_{i_1 < \ldots < i_p}$ is a free (linearly independent) family.

Moreover, for $\alpha \in \bigwedge^p \mathbb{R}^n$, we compute

$$\begin{aligned}\alpha(v_1,\dots,v_p) &\overset{(3.1)}{=} \sum_{i_1,\dots,i_p} v_1^{i_1}\dots v_p^{i_p}\,\alpha(e_{i_1},\dots,e_{i_p})\\ &= \sum_{i_1<\dots<i_p}\sum_{\sigma\in\mathcal{S}_p} v_1^{i_{\sigma(1)}}\dots v_p^{i_{\sigma(p)}}\,\alpha(e_{i_{\sigma(1)}},\dots,e_{i_{\sigma(p)}})\,,\end{aligned}$$

where we used Lemma 3.2 (i). Using again (3.6), we obtain

$$\begin{aligned}\alpha(v_1,\dots,v_p) &\overset{(3.2)}{=} \sum_{i_1<\dots<i_p}\sum_{\sigma\in\mathcal{S}_p} v_1^{i_{\sigma(1)}}\dots v_p^{i_{\sigma(p)}}\,(-1)^{|\sigma|}\alpha(e_{i_1},\dots,e_{i_p})\\ &= \sum_{i_1<\dots<i_p}\alpha(e_{i_1},\dots,e_{i_p})\,e_{i_1}^*\wedge\dots\wedge e_{i_p}^*(v_1,\dots,v_p)\,.\end{aligned}$$

Thus $\{e_{i_1}^*\wedge\dots\wedge e_{i_p}^*\}_{i_1<\dots<i_p}$ is also a generating family and we have

$$\alpha = \sum_{i_1<\dots<i_p}\alpha(e_{i_1},\dots,e_{i_p})\,e_{i_1}^*\wedge\dots\wedge e_{i_p}^*\,. \tag{3.7}$$

□

Next, we define an important natural operation on alternating forms.

Definition 3.8. *Let $\varphi:\mathbb{R}^n\longrightarrow\mathbb{R}^m$ be a linear map and let $\alpha\in\bigwedge^p\mathbb{R}^m$. We define the **pull-back** $\varphi^*\alpha\in\bigwedge^p\mathbb{R}^n$ of α by φ by*

$$\big(\varphi^*\alpha\big)(v_1,\dots,v_p)=\alpha\big(\varphi(v_1),\dots,\varphi(v_p)\big)\,,$$

where $v_1,\dots,v_p\in\mathbb{R}^n$.

Remark. The map $\varphi^*:\bigwedge^p\mathbb{R}^m\longrightarrow\bigwedge^p\mathbb{R}^n$ is linear and satisfies also, for α, $\beta\in\bigwedge^p\mathbb{R}^m$,

$$\varphi^*(\alpha\wedge\beta)=\varphi^*\alpha\wedge\varphi^*\beta\,. \tag{3.8}$$

3.1.2 Differential Forms on $\mathbb{R}^n$

The notion of alternating p-forms on $\mathbb{R}^n$ enables us to introduce another important object in differential geometry.

Definition 3.9. *A C^k-**differential** p-**form** on an open set $U\subset\mathbb{R}^n$ is a C^k-map from U into $\bigwedge^p\mathbb{R}^n$. We denote this space by $\Omega_k^p(U)=C^k(U,\bigwedge^p\mathbb{R}^n)$ or simply $\Omega^p(U)$ in the case of C^∞-differential p-forms.*

Example 3.10. Let $f\in C^{k+1}(U,\mathbb{R})$ with its differential defined in Section 1.1. The differential can be seen as the map

$$df\in C^k\big(U,(\mathbb{R}^n)^*\big)=C^k(U,\textstyle\bigwedge^1\mathbb{R}^n)=\Omega_k^1(U)\,.$$

In particular, the differential of the k-th coordinate function $x_k:U\longrightarrow\mathbb{R}$ $(k=1,\dots,n)$ of $\mathbb{R}^n$ reads as $(dx_k)_x\cdot e_i=\delta_{ik}$, $x\in U$; hence, $dx_k\in C^\infty(U,\bigwedge^1\mathbb{R}^n)$ is the constant map $x\longmapsto e_k^*$ on U.

Definition 3.11. *Let $\alpha \in \Omega_k^p(U)$ and $\beta \in \Omega_k^q(U)$. We define the* ***wedge*** *or* ***exterior product*** *$\alpha \wedge \beta \in \Omega_k^{p+q}(U)$ of the differential forms α and β by the map $x \longmapsto \alpha(x) \wedge \beta(x) \in C^k(U, \bigwedge^{p+q} \mathbb{R}^n)$.*

Remark. Clearly, Proposition 3.6 extends to differential forms on $\mathbb{R}^n$.

Since $\{e_{i_1}^* \wedge \ldots \wedge e_{i_p}^*\}_{i_1 < \ldots < i_p}$ is a basis of $\bigwedge^p \mathbb{R}^n$ by Proposition 3.7, we deduce that α is a C^k-differential p-form on $U \subset \mathbb{R}^n$ if and only if there exist C_n^p C^k-functions

$$\begin{aligned} \alpha_I : U &\longrightarrow \mathbb{R}\,, \\ x &\longmapsto \alpha_I(x) := \alpha(x)\big(e_{i_1}, \ldots, e_{i_p}\big)\,, \end{aligned}$$

where $I = \{(i_1, \ldots, i_p) | i_1 < \ldots < i_p\}$, such that (see (3.7))

$$\alpha(x) = \sum_I \alpha_I(x)\, dx_{i_1}(x) \wedge \ldots \wedge dx_{i_p}(x)\,, \tag{3.9}$$

for all $x \in U$. – Note that in the following the argument of a differential form will often be dropped.

In particular, (3.9) implies for Example 3.10 that

$$df = \sum_{k=1}^n \frac{\partial f}{\partial x_k}\, dx_k \in \Omega_k^1(U)\,, \tag{3.10}$$

where $\frac{\partial f}{\partial x_k}$ means the usual partial derivative of $f \in C^k(U, \mathbb{R})$ at x in the direction e_k, i.e., $\frac{\partial f}{\partial x_k}(x) = df_x \cdot e_k$.

Now, we introduce an important operation on differential forms which can be interpreted as extension to the usual differentiation of real valued functions on $\mathbb{R}^n$. – We denote by d the map assigning to a function $f \in C^{k+1}(U, \mathbb{R})$ its differential df, i.e.,

$$d : \Omega_{k+1}^0(U) = C^{k+1}(U, \mathbb{R}) \longrightarrow \Omega_k^1(U)\,.$$

This can be generalized to a mapping, also denoted by d, from $\Omega_{k+1}^p(U)$ into $\Omega_k^{p+1}(U)$, where $0 \le p \le n-1$. The explicit expression of d is given in the following definition.

Definition 3.12. *Let*

$$\alpha = \sum_I \alpha_I\, dx_{i_1} \wedge \ldots \wedge dx_{i_p} \in \Omega_{k+1}^p(U)\,.$$

Then, we define $d\alpha \in \Omega_k^{p+1}(U)$ by

$$d\alpha = \sum_{l=1}^n \sum_I \frac{\partial \alpha_I}{\partial x_l}\, dx_l \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_p}\,. \tag{3.11}$$

The operation d is called the ***exterior derivative****.*

Remark. Note that d is a *linear* operation, i.e., for $\alpha,\ \beta \in \Omega^p(U)$ we have that

$$d(\alpha + \beta) = d\alpha + d\beta\,. \tag{3.12}$$

And comparing the defining equation (3.11) with (3.10), we see that the exterior derivative coincides also on $\Omega^0(U)$ with the usual differentiation.

Moreover, we observe that

$$d(\alpha_I\, dx_{i_1} \wedge \ldots \wedge dx_{i_p}) = \sum_{l=1}^{n} \frac{\partial \alpha_I}{\partial x_l}\, dx_l \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_p}\,,$$

even if $\{i_1, \ldots, i_p\}$ are not ordered in a increasing way. Indeed, if $\sigma \in \mathcal{S}_p$ such that $i_{\sigma(1)} < \ldots < i_{\sigma(p)}$ and $\alpha = \alpha_I\,(-1)^{|\sigma|}\, dx_{i_{\sigma(1)}} \wedge \ldots \wedge dx_{i_{\sigma(p)}}$, then with (3.11) we get

$$\begin{aligned} d(\alpha_I\, dx_{i_1} \wedge \ldots \wedge dx_{i_p}) &= d(\alpha_I\,(-1)^{|\sigma|}\, dx_{i_{\sigma(1)}} \wedge \ldots \wedge dx_{i_{\sigma(p)}}) \\ &= \sum_{l=1}^{n} \frac{\partial}{\partial x_l}\big(\alpha_I\,(-1)^{|\sigma|}\big)\, dx_l \wedge dx_{i_{\sigma(1)}} \wedge \ldots \wedge dx_{i_{\sigma(p)}} \\ &= \sum_{l=1}^{n} \frac{\partial \alpha_I}{\partial x_l}\, dx_l \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_p}\,. \end{aligned}$$

Relation to Vector Analysis in $\mathbb{R}^3$

Now, we want to study the relation between exterior derivative and well-know operations of vector analysis in $\mathbb{R}^3$. – Let $n = 3$. We already know that the exterior derivative of a function corresponds to the gradient (see (3.10)). Moreover, let $\alpha = \alpha_1\, dx_1 + \alpha_2\, dx_2 + \alpha_3\, dx_3 \in \Omega^1(\mathbb{R}^3)$. We compute

$$\begin{aligned} d\alpha = {} & \frac{\partial \alpha_1}{\partial x_2}\, dx_2 \wedge dx_1 + \frac{\partial \alpha_1}{\partial x_3}\, dx_3 \wedge dx_1 \\ & + \frac{\partial \alpha_2}{\partial x_1}\, dx_1 \wedge dx_2 + \frac{\partial \alpha_2}{\partial x_3}\, dx_3 \wedge dx_2 \\ & + \frac{\partial \alpha_3}{\partial x_1}\, dx_1 \wedge dx_3 + \frac{\partial \alpha_3}{\partial x_2}\, dx_2 \wedge dx_3 \\ = {} & \left(\frac{\partial \alpha_2}{\partial x_1} - \frac{\partial \alpha_1}{\partial x_2}\right) dx_1 \wedge dx_2 + \left(\frac{\partial \alpha_1}{\partial x_3} - \frac{\partial \alpha_3}{\partial x_1}\right) dx_3 \wedge dx_1 \\ & + \left(\frac{\partial \alpha_3}{\partial x_2} - \frac{\partial \alpha_2}{\partial x_3}\right) dx_2 \wedge dx_3\,. \end{aligned}$$

If we identify the exterior product $dx_{i+1} \wedge dx_{i-1}$, where the indices are in $\mathbb{Z}/3\mathbb{Z}$, with the basis vectors e_i (i=1,. . . ,3) on $\mathbb{R}^3$, we observe that the exterior derivative $d\alpha$ of a 1-form corresponds to $rot\,\boldsymbol{\alpha}$, with $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ vector field on $\mathbb{R}^3$.

Let now $\beta \in \Omega^2(\mathbb{R}^3)$ with

$$\beta = \beta_1\, dx_2 \wedge dx_3 + \beta_2\, dx_3 \wedge dx_1 + \beta_3\, dx_1 \wedge dx_2 \,.$$

A simple calculation shows that

$$d\beta = \left(\frac{\partial \beta_1}{\partial x_1} + \frac{\partial \beta_2}{\partial x_2} + \frac{\partial \beta_3}{\partial x_3} \right) dx_1 \wedge dx_2 \wedge dx_3 \,.$$

Hence, if we identify $dx_1 \wedge dx_2 \wedge dx_3$ with the scalar 1, we observe that $d\beta$ corresponds to $div\, \boldsymbol{\beta}$, with $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3)$ vector field on $\mathbb{R}^3$. – In summary, in the three-dimensional case the exterior derivative translates to well-known operations of vector analysis.

The "Leibniz-rule" in the following proposition shows that the exterior derivative is a derivation.

Proposition 3.13. *Let $\alpha \in \Omega^p(U)$ and $\beta \in \Omega^q(U)$. Then we have*

$$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^p \alpha \wedge d\beta \,. \tag{3.13}$$

Proof. We first prove (3.13) for

$$\alpha = \alpha_I\, dx_{i_1} \wedge \ldots \wedge dx_{i_p} \,, \qquad \beta = \beta_J\, dx_{j_1} \wedge \ldots \wedge dx_{j_q} \,,$$

where $\alpha_I,\ \beta_J \in C^\infty(U, \mathbb{R})$ for some choice of indices $i_1 < \ldots < i_p$ and $j_1 < \ldots < j_q$. The exterior product reads as

$$\alpha \wedge \beta = \alpha_I\, \beta_J\, dx_{i_1} \wedge \ldots \wedge dx_{i_p} \wedge dx_{j_1} \wedge \ldots \wedge dx_{j_q} \,.$$

We compute the exterior derivative of the wedge product, using several properties of the wedge product (see Proposition 3.6) and the definition of the exterior derivative:

$$\begin{aligned} d(\alpha \wedge \beta) &= \sum_{l=1}^{n} \frac{\partial}{\partial x_l}(\alpha_I\, \beta_J)\, dx_l \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_p} \wedge dx_{j_1} \wedge \ldots \wedge dx_{j_q} \\ &= \left(\sum_{l=1}^{n} \frac{\partial \alpha_I}{\partial x_l}\, dx_l \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_p} \right) \wedge \beta_J\, dx_{j_1} \wedge \ldots \wedge dx_{j_q} \\ &\quad + \sum_{l=1}^{n} \left(\alpha_I\, dx_l \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_p} \right) \wedge \frac{\partial \beta_J}{\partial x_l}\, dx_{j_1} \wedge \ldots \wedge dx_{j_q} \\ &= d\alpha \wedge \beta + \sum_{l=1}^{n} \left(\alpha_I\, dx_{i_1} \wedge \ldots \wedge dx_{i_p} \right) \wedge (-1)^p \frac{\partial \beta_J}{\partial x_l}\, dx_l \wedge dx_{j_1} \wedge \ldots \wedge dx_{j_q} \\ &= d\alpha \wedge \beta + (-1)^p \alpha \wedge d\beta \,. \end{aligned}$$

By linearity of the exterior derivative, we easily establish the proof for general α and β. ☐

Proposition 3.14. *The map*

$$d \circ d : \Omega^p(U) \longrightarrow \Omega^{p+2}(U)\,,$$

where $0 \le p \le n-2$, *is equal to zero.*

Proof. Again it suffices to prove the proposition for $\alpha = \alpha_I\, dx_{i_1} \wedge \ldots \wedge dx_{i_p} \in \Omega^p(U)$ and to conclude by linearity. We know that

$$d\alpha = \sum_{l=1}^{n} \frac{\partial \alpha_I}{\partial x_l}\, dx_l \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_p}\,.$$

Then we compute

$$\begin{aligned} d(d\alpha) &= d\left(\sum_{l=1}^{n} \frac{\partial \alpha_I}{\partial x_l}\, dx_l \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_p}\right) \\ &= \sum_{k=1}^{n}\sum_{l=1}^{n} \frac{\partial^2 \alpha_I}{\partial x_k \partial x_l}\, dx_k \wedge dx_l \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_p}\,. \end{aligned}$$

Since by the Lemma of Schwartz

$$\frac{\partial^2 \alpha_I}{\partial x_k \partial x_l} = \frac{\partial^2 \alpha_I}{\partial x_l \partial x_k}\,,$$

and clearly $dx_k \wedge dx_l = -dx_l \wedge dx_k$, we obtain (the case $k = l$ is trivial)

$$\begin{aligned} d(d\alpha) &= \sum_{k<l} \frac{\partial^2 \alpha_I}{\partial x_k \partial x_l}\, dx_k \wedge dx_l \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_p} \\ &\quad + \sum_{k>l} \frac{\partial^2 \alpha_I}{\partial x_k \partial x_l}\, dx_k \wedge dx_l \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_p} \\ &= \sum_{k<l} \frac{\partial^2 \alpha_I}{\partial x_k \partial x_l}\, dx_k \wedge dx_l \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_p} \\ &\quad - \sum_{k<l} \frac{\partial^2 \alpha_I}{\partial x_k \partial x_l}\, dx_k \wedge dx_l \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_p} = 0\,. \end{aligned}$$

□

We denote by

$$\Omega(U) = \bigoplus_{p=0}^{n} \Omega^p(U) \tag{3.14}$$

the associative, graded and “anti-commutative algebra of differential forms. This algebra is often called *Grassman algebra* or *exterior algebra* of differential forms on U.

Theorem 3.15. *There exists a unique linear map $d : \Omega(U) \longrightarrow \Omega(U)$ such that*

(i) $d : \Omega^p(U) \longrightarrow \Omega^{p+1}(U)$*;*
(ii) on $\Omega^0(U)$ the map d coincides with the differential of functions;
(iii) for all $\alpha \in \Omega^p(U)$ and $\beta \in \Omega^q(U)$, we have

$$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^p \alpha \wedge d\beta\,;$$

(iv) and $d \circ d = 0$.

Proof. □

Next, we want to answer the question how to transport differential forms.

Definition 3.16. *Let $\varphi \in C^{k'}(U, V)$, with $k' \geq k+1$, where U and V open sets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Moreover, let $\alpha \in \Omega^p_k(V)$. We define the* ***pull-back*** *$\varphi^*\alpha \in \Omega^p_k(U)$ of α by φ in the following way:*

$$\varphi^*\alpha_x = (d\varphi_x)^* \alpha_{\varphi(x)}\,, \qquad x \in U\,. \tag{3.15}$$

Remark. By Definition 3.8, we get, for $v_1, \dots, v_p \in \mathbb{R}^n$,

$$\varphi^*\alpha_x(v_1, \dots, v_p) = \alpha_{\varphi(x)}(d\varphi_x \cdot v_1, \dots, d\varphi_x \cdot v_p)\,, \qquad x \in U\,. \tag{3.16}$$

If $\alpha \in \Omega^0(V)$, then obviously $\varphi^*\alpha = \alpha \circ \varphi$.

We prove that $\varphi^*\alpha$ is indeed a C^k-differential p-form on U. – Clearly, we have that $\varphi^*\alpha_x \in \bigwedge^p \mathbb{R}^n$ for all $x \in U$. Since $\{e^*_{i_1} \wedge \ldots \wedge e^*_{i_p}\}_{i_1 < \ldots < i_p}$ is a basis for $\bigwedge^p \mathbb{R}^n$, we can write

$$\varphi^*\alpha_x = \sum_I \varphi^*\alpha_x(e_{i_1}, \dots, e_{i_p})\, e^*_{i_1} \wedge \ldots \wedge e^*_{i_p}\,.$$

And the definition of the pull-back gives

$$\varphi^*\alpha_x = \sum_I \alpha_{\varphi(x)} \left(\frac{\partial \varphi}{\partial x_{i_1}}(x), \dots, \frac{\partial \varphi}{\partial x_{i_p}}(x) \right) e^*_{i_1} \wedge \ldots \wedge e^*_{i_p}\,.$$

Since by assumption $\alpha \in \Omega^p_k(V)$ and $\varphi \in C^{k'}(U, V)$, we get that

$$x \longmapsto \alpha_{\varphi(x)} \left(\frac{\partial \varphi}{\partial x_{i_1}}(x), \dots, \frac{\partial \varphi}{\partial x_{i_p}}(x) \right) \in C^k(U, \mathbb{R})\,,$$

implying $\varphi^*\alpha \in \Omega^p_k(U)$. – Note that the regularity of the map φ determines the regularity of the pull-back.

The properties of the pull-back are summarized in the following proposition.

Proposition 3.17. *Let $\varphi \in C^\infty(U,V)$ with $U \subset \mathbb{R}^n$ open, $V \subset \mathbb{R}^m$ open, and let $\alpha \in \Omega^p(V)$.*

(i) For all $\beta \in \Omega^p(V)$, we have

$$\varphi^*(\alpha+\beta) = \varphi^*\alpha + \varphi^*\beta\,. \tag{3.17}$$

(ii) For all $\beta \in \Omega^q(V)$, we have

$$\varphi^*(\alpha\wedge\beta) = \varphi^*\alpha \wedge \varphi^*\beta\,. \tag{3.18}$$

(iii) For $\psi \in C^\infty(V,W)$ with $W \subset \mathbb{R}^l$ open, we have

$$(\psi\circ\varphi)^*\gamma = \varphi^*(\psi^*\gamma)\,, \tag{3.19}$$

where $\gamma \in \Omega^p(W)$.

(iv) The exterior derivative and the pull-back commute, i.e.,

$$d(\varphi^*\alpha) = \varphi^*(d\alpha)\,. \tag{3.20}$$

Proof. Let $x \in U \subset \mathbb{R}^n$ and $v_1,\ldots,v_p \in \mathbb{R}^n$. *For (i)*: We compute, using the vector space structure on $\bigwedge^p \mathbb{R}^n$,

$$\begin{aligned}\varphi^*(\alpha+\beta)_x(v_1,\ldots,v_p) &= (\alpha+\beta)_{\varphi(x)}(d\varphi_x\cdot v_1,\ldots,d\varphi_x\cdot v_p)\\ &= \alpha_{\varphi(x)}(d\varphi_x\cdot v_1,\ldots,d\varphi_x\cdot v_p)\\ &\quad +\beta_{\varphi(x)}(d\varphi_x\cdot v_1,\ldots,d\varphi_x\cdot v_p)\\ &= \varphi^*\alpha_x(v_1,\ldots,v_p)+\varphi^*\beta_x(v_1,\ldots,v_p)\,.\end{aligned}$$

For (ii): With Definition 3.4 of the exterior product, we obtain

$$\begin{aligned}\varphi^*\big(\alpha\wedge\beta\big)_x(v_1,\ldots,v_{p+q}) &= \big(\alpha\wedge\beta\big)_{\varphi(x)}(d\varphi_x\cdot v_1,\ldots,d\varphi_x\cdot v_{p+q})\\ &= \sum_{\sigma\in\mathcal{S}_{p+q}} \frac{(-1)^{|\sigma|}}{p!q!}\,\alpha_{\varphi(x)}\big(d\varphi_x\cdot v_{\sigma(1)},\ldots,d\varphi_x\cdot v_{\sigma(p)}\big)\\ &\quad \cdot\beta_{\varphi(x)}\big(d\varphi_x\cdot v_{\sigma(p+1)},\ldots,d\varphi_x\cdot v_{\sigma(p+q)}\big)\\ &= \sum_{\sigma\in\mathcal{S}_{p+q}} \frac{(-1)^{|\sigma|}}{p!q!}\varphi^*\alpha_x(v_{\sigma(1)},\ldots,v_{\sigma(p)})\\ &\quad \cdot\varphi^*\beta_x(v_{\sigma(p+1)},\ldots,v_{\sigma(p+q)})\\ &= \varphi^*\alpha_x\wedge\varphi^*\beta_x(v_1,\ldots,v_{p+q})\,.\end{aligned}$$

For (iii): Recall that for all $v \in \mathbb{R}^n$ the following formula for the composition holds (see Proposition 1.4):

$$d(\psi\circ\varphi)_x\cdot v = d\psi_{\varphi(x)}\cdot d\varphi_x\cdot v\,.$$

This implies, for $\gamma \in \Omega^p(W)$,

$$
\begin{aligned}
\big(\psi\circ\varphi\big)^*\gamma_x(v_1,\ldots,v_p) &= \gamma_{\psi\circ\varphi(x)}\big(d(\psi\circ\varphi)_x\cdot v_1,\ldots,d(\psi\circ\varphi)_x\cdot v_p\big)\\
&= \gamma_{\psi(\varphi(x))}\big(d\psi_{\varphi(x)}\cdot(d\varphi_x\cdot v_1),\ldots,d\psi_{\varphi(x)}\cdot(d\varphi_x\cdot v_p)\big)\\
&= \psi^*\gamma_{\varphi(x)}(d\varphi_x\cdot v_1,\ldots,d\varphi_x\cdot v_p)\\
&= \varphi^*\big(\psi^*\gamma\big)_x(v_1,\ldots,v_p)\,.
\end{aligned}
$$

For (iv): First, we prove the statement for functions $f\in\Omega^0(V)$. For $v\in\mathbb{R}^n$, we deduce, using Definition 3.16,

$$
\begin{aligned}
\big(\varphi^*df\big)_x(v) &= df_{\varphi(x)}(d\varphi_x\cdot v)\\
&= d(f\circ\varphi)_x\cdot v = \big(d(\varphi^*f)\big)_x(v)\,.
\end{aligned}
$$

Hence, for functions we have that

$$
\varphi^*(df) = d(\varphi^*f)\,. \tag{3.21}
$$

For the general case, let

$$
\alpha = \sum_J \alpha_J\, dy_{j_1}\wedge\ldots\wedge dy_{j_p}\in\Omega^p(V)\,,
$$

where $(y_1,\ldots,y_m)$ denote the canonical coordinate functions on $\mathbb{R}^m$. Using (3.17) and (3.18), it follows

$$
\varphi^*\alpha = \sum_J \alpha_J\circ\varphi\ \ \varphi^*dy_{j_1}\wedge\ldots\wedge\varphi^*dy_{j_p}\,.
$$

Writing $y_{i_k}\circ\varphi=\varphi_{i_k}:\mathbb{R}^n\longrightarrow\mathbb{R}$, for $k=1,\ldots,p$, it follows from (3.21) that

$$
\varphi^*\alpha = \sum_J \alpha_J\circ\varphi\ \ d\varphi_{j_1}\wedge\ldots\wedge\varphi^*d\varphi_{j_p}\,.
$$

Applying the exterior derivative, we get (recall that $d\circ d=0$)

$$
\begin{aligned}
d(\varphi^*\alpha) &= \sum_J d(\alpha_J\circ\varphi)\wedge d\varphi_{j_1}\wedge\ldots\wedge d\varphi_{j_p}\\
&\overset{(3.10)}{=} \sum_J\sum_{k=1}^n \frac{\partial(\alpha_J\circ\varphi)}{\partial x_k}\,dx_k\wedge d\varphi_{j_1}\wedge\ldots\wedge d\varphi_{j_p}
\end{aligned}
$$

By Proposition 1.4, we obtain

$$
\begin{aligned}
d(\varphi^*\alpha) &= \sum_J\sum_{k=1}^n\sum_{l=1}^m \frac{\partial\alpha_J}{\partial y_l}\big(\varphi(\cdot)\big)\frac{\partial\varphi_l}{\partial x_k}\,dx_k\wedge d\varphi_{j_1}\wedge\ldots\wedge d\varphi_{j_p}\\
&\overset{(3.10)}{=} \sum_J\sum_{l=1}^m \frac{\partial\alpha_J}{\partial y_l}\circ\varphi\ d\varphi_l\wedge d\varphi_{j_1}\wedge\ldots\wedge d\varphi_{j_p}\\
&= \varphi^*\left(\sum_J\sum_{l=1}^m \frac{\partial\alpha_J}{\partial y_l}\,dy_l\wedge dy_{j_1}\wedge\ldots\wedge dy_{j_p}\right)\\
&= \varphi^*(d\alpha)\,,
\end{aligned}
$$

where the definition for the exterior derivative was used in the last line. □

As an application of the proposition, we can compute the pull-back explicitly. – Let

$$\alpha = \sum_J \alpha_J \, dy_{j_1} \wedge \ldots \wedge dy_{j_p} \in \Omega^p(V)\,,$$

where $(y_1, \ldots, y_m)$ denote the canonical coordinate functions on $\mathbb{R}^m$. Then the pull-back becomes, using (3.17) and (3.18),

$$\begin{aligned}\varphi^*\alpha &= \sum_J \varphi^*\big(\alpha_J \, dy_{j_1} \wedge \ldots \wedge dy_{j_p}\big) \\ &= \sum_J \varphi^*\alpha_J \;\; \varphi^* dy_{j_1} \wedge \ldots \wedge \varphi^* dy_{j_p}\end{aligned}$$

Writing again $y_{i_k} \circ \varphi = \varphi_{i_k} : \mathbb{R}^n \longrightarrow \mathbb{R}$, for $k = 1, \ldots, p$, we obtain

$$\begin{aligned}\varphi^*\alpha &= \sum_J \alpha_J \circ \varphi \;\; d(y_{j_1} \circ \varphi) \wedge \ldots \wedge d(y_{j_p} \circ \varphi) \\ &= \sum_J \alpha_J \circ \varphi \;\; d\varphi_{j_1} \wedge \ldots \wedge d\varphi_{j_p}\,. \end{aligned} \tag{3.22}$$

With (3.10), we arrive at

$$\varphi^*\alpha = \sum_J \alpha_J \circ \varphi \sum_{i_1,\ldots,i_p} \frac{\partial \varphi_{j_1}}{\partial x_{i_1}} \cdots \frac{\partial \varphi_{j_p}}{\partial x_{i_p}} \, dx_{i_1} \wedge \ldots \wedge dx_{i_p}\,. \tag{3.23}$$

3.2 Differential Forms on Manifolds

3.2.1 The Cotangent Bundle

In the last section, we have seen skew-symmetric (alternated) p-forms on $\mathbb{R}^n$. Similarly, for an arbitrary finite dimensional vector space E, we define $\bigwedge^p E^*$ to be the alternated p-forms on E. Note that all the results obtained in Section 3.1.1 remain true for this more general setting.

Alternated p-forms on a vector space play an important role in the context of manifolds. – Let M be a m-dimensional C^k-differentiable manifold, $k \geq 1$, and $p \in M$. We already know that the tangent space T_pM at p is a m-dimensional vector space (see Definition 2.35). Denoting its dual space $(T_pM)^*$ by T_p^*M, we denote by

$$\bigwedge^p T_p^* M$$

the alternated p-forms on T_pM. From a "set" point of view the **cotangent bundle** of M is then defined by

$$\bigwedge^p T^*M = \bigcup_{p\in M} \bigwedge^p T_p^*M \,.$$

For $\alpha \in \bigwedge^p T_p^*M$, the projection map to the base point reads as

$$\begin{aligned} \pi : \bigwedge^p T^*M &\longrightarrow M\,, \\ \alpha &\longmapsto p\,. \end{aligned} \tag{3.24}$$

Let (U, x) be a local chart on M. The basis for T_pM , $p \in U$, is then given by $\{\frac{\partial}{\partial x_i}\}_{i=1,\dots,m}$ (see (2.15)). Moreover, let $x_i \in C^k(U, \mathbb{R})$, for $i = 1, \dots, m$, denote the coordinate functions associated to the chart (U, x). As shown in Exercise 2.43 the following formula holds, for all $p \in U$,

$$dx_i(p) \cdot \frac{\partial}{\partial x_j}(p) = \delta_{ij}\,. \tag{3.25}$$

Hence

$$\big\{dx_i(p)\big\}_{i=1,\dots,m} \tag{3.26}$$

is a basis of T_p^*M dual to $\{\frac{\partial}{\partial x_i}\}_{i=1,\dots,m}$ basis of T_pM. Using Proposition 3.7, we deduce that

$$\big\{dx_{i_1}(p) \wedge \ldots \wedge dx_{i_p}(p)\big\}_{i_1<\ldots<i_p} \tag{3.27}$$

is a basis of $\bigwedge^p T_p^*M$.

Definition 3.18. *Let M be a C^k-differentiable manifold of dimension m. A C^{k-1}-**differential** p-**form** on M is a map*

$$\omega : M \longrightarrow \bigwedge^p T^*M$$

such that $\pi \circ \omega = id_M$, and for all local charts (U, x) on M^m there exist C_m^p functions $\omega_I \in C^{k-1}(U, \mathbb{R})$ with $I = \{(i_1, \dots, i_p) | i_1 < \ldots < i_p\}$ satisfying

$$\omega(p) = \sum_I \omega_I(p)\, dx_{i_1}(p) \wedge \ldots \wedge dx_{i_p}(p)\,, \tag{3.28}$$

for all $p \in U$. – We denote the vector space of C^k-differential p-forms on M by $\Omega_k^p(M)$.

Next, we want to construct a differentiable structure on the cotangent bundle $\bigwedge^p T^*M$ in such a way that C^k-differentiable p-forms on M are just C^k-*sections*, i.e., $\omega \in C^k(M, \bigwedge^p T^*M)$ and $\pi \circ \omega = id_M$. – The method will be very similar to the one used for the construction of the differentiable structure on the tangent bundle TM in Section 2.5.3.

Let $(U_i, \varphi_i)_{i\in I}$ be an atlas for M. We set

$$\bigwedge^p T^*U_i = \bigcup_{p\in U_i} \bigwedge^p T_p^*M = \pi^{-1}(U_i)\,,$$

and define

$$\begin{aligned}\Phi_i^p : \bigwedge^p T^*U_i &\longrightarrow \mathbb{R}^m \times \bigwedge^p \mathbb{R}^m\,,\\ \omega &\longmapsto \big(\varphi_i \circ \pi(\omega), \varphi_i^p(\omega)\big)\,,\end{aligned} \tag{3.29}$$

with

$$\varphi_i^p(\omega) = \big((d\varphi_i^{-1})_{\varphi_i(p)}\big)^*\omega\,, \tag{3.30}$$

where $\pi(\omega) = p \in U_i$. More precisely, for $v_1, \ldots, v_p \in \mathbb{R}^m$, we have that

$$\varphi_i^p(\omega)(v_1, \ldots, v_p) = \omega_p\big(d(\varphi_i^{-1})_{\varphi_i(p)} \cdot v_1, \ldots, d(\varphi_i^{-1})_{\varphi_i(p)} \cdot v_p\big) \in \mathbb{R}\,.$$

It is clear that Φ_i^p is a bijection from $\bigwedge^p T^*U_i$ into $\varphi_i(U_i) \times \bigwedge^p \mathbb{R}^m$. We then say that $\Omega \subset \bigwedge^p T^*M$ is *open* if and only if for all $i \in I$ the set $\Phi_i^p(\Omega \cap \bigwedge^p T^*U_i)$ is open in $\mathbb{R}^m \times \bigwedge^p \mathbb{R}^m$.

Proposition 3.19. *These open sets define a (separated) topology on $\bigwedge^p T^*M$ which depends only on the differentiable structure of M, and not on the atlas $(U_i, \varphi_i)_{i\in I}$ representing the fixed differentiable structure on M. Moreover, the maps Φ_i^p, $i \in I$, defined in (3.29) are homeomorphisms for this topology and the projection π is continuous.*

Proposition 3.20. *Let M be a m-dimensional C^k-manifold with $k \geq 2$ and let $(U_i, \varphi_i)_{i\in I}$ be an atlas for M^m. Then $(\bigwedge^p T^*U_i, \Phi_i^p)_{i\in I}$ defines a C^{k-1}-differentiable structure on the cotangent bundle $\bigwedge^p T^*M$, depending only on the differentiable structure on M^m. And $\bigwedge^p T^*M$ is a C^{k-1}-differentiable manifold of dimension $m + C_m^p$ for this differentiable structure. Moreover, the projection π is a C^{k-1}-submersion for this differentiable structure.*

The two last propositions can be proved exactly in the same way as the corresponding Propositions 2.37 and 2.38 for the tangent bundle, once we have identified the transition functions. Therefore, we first prove the following

Lemma 3.21. *Let $(U_i, \varphi_i)_{i\in I}$ and $(V_j, \psi_j)_{j\in J}$ be two equivalent systems of charts for M. Moreover, let $i \in I$ and $j \in J$ such that $U_i \cap V_j \neq \emptyset$. Then on $\Phi_i^p(\bigwedge^p T^*(U_i \cap V_j)) \subset \mathbb{R}^m \times \bigwedge^p \mathbb{R}^m$ the following formula for the transition functions on $\bigwedge^p T^*M$ holds:*

$$\Psi_j^p \circ (\Phi_i^p)^{-1}(x, \beta) = \big(\psi_j \circ \varphi_i^{-1}(x), (\varphi_i \circ \psi_j^{-1})^*\beta\big)\,, \tag{3.31}$$

where $\varphi_i^{-1}(x) = p \in U_i \cap V_j$ and $\beta \in \bigwedge^p \mathbb{R}^m$.

Proof. The defining equation (3.29) gives

$$\Psi_j^p(\omega)(v_1, \ldots, v_p) = \Big(\psi_j \circ \pi(\omega), \omega_p\big(d(\psi_j^{-1})_{\psi(p)} \cdot v_1, \ldots, d(\psi_j^{-1})_{\psi(p)} \cdot v_p\big)\Big)\,,$$

where $\omega \in \bigwedge^p T^*V_j$ and $v_1, \ldots, v_p \in \mathbb{R}^m$. Moreover, let $[\gamma_1], \ldots, [\gamma_p] \in T_pM$ be tangent vectors at p, i.e., $\gamma_l \in C^1([0,1], M)$, for $l = 1, \ldots, p$, are

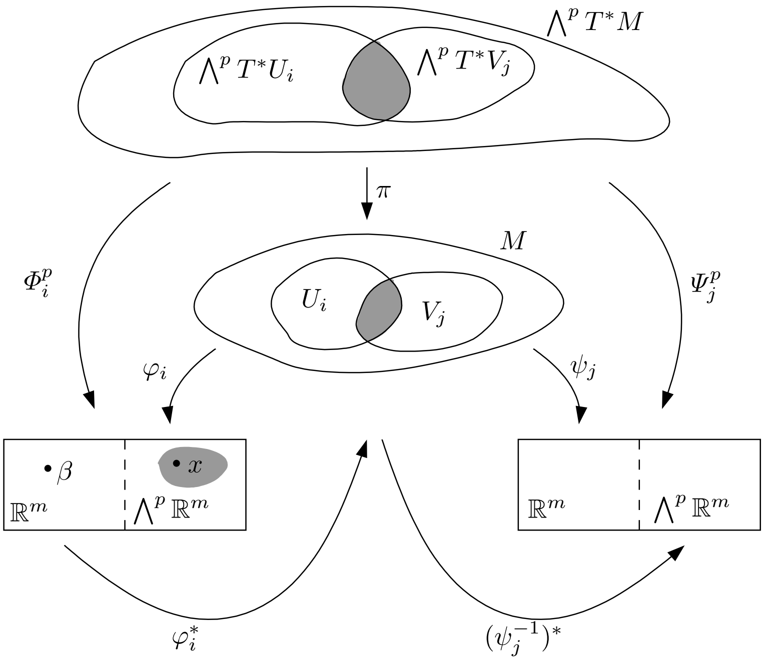

Fig. 3.1. Transition functions for the cotangent bundle.

smooth paths in M with $\gamma_l(0) = \varphi_i^{-1}(x) = p$. Then the alternating p-form $\big((\Phi_i^p)^{-1}(x,\beta)\big)$ on T_pM, acting on $[\gamma_1], \ldots, [\gamma_p]$, equals

$$\big((\Phi_i^p)^{-1}(x,\beta)\big)([\gamma_1], \ldots, [\gamma_p]) = \beta\big(d\varphi_i \cdot [\gamma_1], \ldots, d\varphi_i \cdot [\gamma_p]\big) \in \bigwedge\nolimits^p T_p^*M\,,$$

where $\beta \in \bigwedge^p \mathbb{R}^m$. Hence, for the transition functions we deduce

$$\begin{aligned}
&\Psi_j^p\big((\Phi_i^p)^{-1}(x,\beta)\big)(v_1, \ldots, v_p) \\
&= \Big(\psi_j \circ \pi\big((\Phi_i^p)^{-1}(x,\beta)\big), \big((\Phi_i^p)^{-1}(x,\beta)\big)(d\psi_j^{-1} \cdot v_1, \ldots, d\psi_j^{-1} \cdot v_p)\Big) \\
&= \Big(\psi_j \circ \varphi_i^{-1}(x), \beta\big(d\varphi_i \cdot d\psi_j^{-1} \cdot v_1, \ldots, d\varphi_i \cdot d\psi_j^{-1} \cdot v_p\big)\Big)\,.
\end{aligned}$$

With Definition 3.16 of the pull-back, we then arrive at

$$\begin{aligned}
&\Psi_j^p\big((\Phi_i^p)^{-1}(x,\beta)\big)(v_1, \ldots, v_p) \\
&= \Big(\psi_j \circ \varphi_i^{-1}(x), \beta\big(d(\varphi_i \circ \psi_j^{-1}) \cdot v_1, \ldots, d(\varphi_i \circ \psi_j^{-1}) \cdot v_p\big)\Big) \\
&= \Big(\psi_j \circ \varphi_i^{-1}(x), (\varphi_i \circ \psi_j^{-1})^*\beta(v_1, \ldots, v_p)\Big)\,.
\end{aligned}$$

$\square$

As mentioned before we can prove Propositions 3.19 and 3.20 with this lemma exactly as in the case of the tangent bundle. We only note that from the explicit expression (3.31) for the transition functions, we deduce that they are C^{k-1}-diffeomorphisms (direct consequence of the Definition 3.16 for the pull-back for differential forms on $\mathbb{R}^m$). – These results allow us to see differential forms on a manifold from a different point of view, namely as sections of the cotangent bundle.

Proposition 3.22. *Let M be a $C^{k'}$-differentiable manifold of dimension m, with $k' \geq k+1$. Then a C^k-differential p-form is a C^k-section of $\bigwedge^p T^*M$ for the above defined differential structure.*

Proof. Let $\omega \in \Omega_k^p(M)$ and $(U_i, \varphi_i)_{i\in I}$ an atlas for M. In order to show that $\omega \in C^k(M, \bigwedge^p T^*M)$, it suffices by Definition 2.27 to prove that

a) the map $\omega : M \longrightarrow \bigwedge^p T^*M$ is continuous,
b) and for all $i, j \in I$ with $U_i \cap U_j \neq \emptyset$ the coordinate expression for ω

$$\Phi_j^p \circ \omega \circ \varphi_i^{-1} : \varphi_i\Big(\omega^{-1}\big(\bigwedge\nolimits^p T^*U_j\big) \cap U_i\Big) \subset \mathbb{R}^m \longrightarrow \mathbb{R}^m \times \bigwedge\nolimits^p \mathbb{R}^m$$

is a C^k-map.

For a): We assume that $\Omega \subset \bigwedge^p T^*M$ is open and show that $\omega^{-1}(\Omega)$ is open in M. Since φ_i, $i \in I$, is a bijection, we can write

$$\omega^{-1}(\Omega) = \bigcup_{i\in I} (\varphi_i)^{-1}\varphi_i\big(\omega^{-1}(\Omega) \cap U_i\big)\,, \tag{3.32}$$

and, since Φ_j^p, $j \in I$, is also a bijection,

$$\begin{aligned}
\varphi_i\big(\omega^{-1}(\Omega) \cap U_i\big) &= \varphi_i\left(\omega^{-1}\left(\bigcup_{j\in J} (\Phi_j^p)^{-1}\Phi_j^p\big(\Omega \cap \bigwedge\nolimits^p T^*U_j\big)\right) \cap U_i\right) \\
&= \varphi_i\left(\bigcup_{j\in J} \big(\Phi_j^p \circ \omega\big)^{-1}\big(\Phi_j^p(\Omega \cap \bigwedge\nolimits^p T^*U_j)\big) \cap U_i\right) \\
&= \bigcup_{j\in J} \varphi_i\Big(\big(\Phi_j^p \circ \omega\big)^{-1}\big(\Phi_j^p(\Omega \cap \bigwedge\nolimits^p T^*U_j)\big) \cap U_i\Big)\,.
\end{aligned} \tag{3.33}$$

This is equivalent to

$$\bigcup_{j\in J} \Big(\big(\Phi_j^p \circ \omega \circ \varphi_i^{-1}\big)^{-1}\big(\Phi_j^p(\Omega \cap \bigwedge\nolimits^p T^*U_j)\big)\Big)\,.$$

By definition, the map $\Phi_j^p \circ \omega \circ \varphi_i^{-1}$ is defined on the set $\varphi_i\Big(\omega^{-1}\big(\bigwedge^p T^*U_j\big) \cap U_i\Big) \subset \mathbb{R}^m$ which is *open*.

We postpone the proof of a) and look first at b). By Definition 3.18, we mean by $\omega \in \Omega_k^p(M)$ that there exist for each local chart (U_i, φ_i) functions $\omega_I \in C^k(U_i, \mathbb{R})$ such that

$$\omega = \sum_I \omega_I \, d\varphi_{i_1} \wedge \ldots \wedge d\varphi_{i_p} \,, \tag{3.34}$$

on U_i. Moreover, for $x \in \varphi_i\Big(\omega^{-1}\big(\bigwedge^p T^*U_j\big) \cap U_i\Big)$ and $v_1, \ldots, v_p \in \mathbb{R}^m$ the coordinate expression for ω reads as, using (3.29),

$$\big(\Phi_j^p \circ \omega \circ \varphi_i^{-1}\big)_x(v_1, \ldots, v_p) = \Big(\varphi_j \circ \varphi_i^{-1}(x), \varphi_j^p\big(\omega\big)(v_1, \ldots, v_p)\Big) \,, \tag{3.35}$$

where

$$\varphi_j^p\big(\omega\big)(v_1, \ldots, v_p) = \omega_{\varphi_i^{-1}(x)}\big(d(\varphi_j^{-1})_x \cdot v_1, \ldots, d(\varphi_j^{-1})_x \cdot v_p\big) \,.$$

The first factor of (3.35) is clearly a $C^{k'}$-map, since M is by assumption a $C^{k'}$-differential manifold. Using the representation (3.34), the second factor becomes

$$\sum_I \omega_I \circ \varphi_i^{-1}(x) \left(d\varphi_{i_1} \wedge \ldots \wedge d\varphi_{i_p}\right)_{\varphi_i^{-1}(x)} \big(d(\varphi_j^{-1})_x \cdot v_1, \ldots, d(\varphi_j^{-1})_x \cdot v_p\big) \,.$$

Since for the differentials of the coordinate functions we have $d\varphi_i(p) = e_i^*$ for arbitrary $p \in M$ (see Section 3.1.2), the last expression equals

$$\begin{aligned}
&\sum_I \omega_I \circ \varphi_i^{-1}(x) \, e_{i_1}^* \wedge \ldots \wedge e_{i_p}^* \big(d\varphi_i \cdot d(\varphi_j^{-1})_x \cdot v_1, \ldots, d\varphi_i \cdot d(\varphi_j^{-1})_x \cdot v_p\big) \\
&= \sum_I \omega_I \circ \varphi_i^{-1}(x) \, e_{i_1}^* \wedge \ldots \wedge e_{i_p}^* \big(d(\varphi_i \circ \varphi_j^{-1})_x \cdot v_1, \ldots, d(\varphi_i \circ \varphi_j^{-1})_x \cdot v_p\big) \\
&= \sum_I \omega_I \circ \varphi_i^{-1}(x) \, (\varphi_i \circ \varphi_j^{-1})^* \big(e_{i_1}^* \wedge \ldots \wedge e_{i_p}^*\big)(v_1, \ldots, v_p) \,,
\end{aligned}$$

which is clearly a C^k-map.

Coming back to a), we first note that by definition of the topology on $\bigwedge^p T^*M$ the set $\Phi_j^p(\Omega \cap \bigwedge^p T^*U_j)$ is by assumption open in $\mathbb{R}^m \times \bigwedge^p \mathbb{R}^m$. From this and using (3.33), we then get the openness of $\varphi_i\big(\omega^{-1}(\Omega) \cap U_i\big)$. With (3.32), we deduce that $\omega^{-1}(\Omega)$ is open in M and the continuity of the map ω. $\square$

Remark. The last proposition can be interpreted as alternative definition for differential forms on manifolds. Note also that the proof of it is completely analogous to the one showing the regularity of the tangent map in Section 2.6.

3.2.2 Operations on Differential Forms on Manifolds

At this place, the operations of pull-back and exterior derivative, already defined for differential forms on $\mathbb{R}^m$ in Section 3.1.2, will be extended to differential forms on manifolds.

Definition 3.23. *Let M^m and N^n be two C^∞-differentiable manifolds. Moreover, let $\varphi : M^m \longrightarrow N^n$ be a $C^{k'}$-map, $k' \geq k+1$ and let $\omega \in \Omega_k^p(N)$. The **pull-back** of ω by φ is the following element of $\Omega_k^p(M)$ denoted by $\varphi^*\omega$:*

$$\big(\varphi^*\omega\big)_p(X_1, \ldots, X_p) = \omega_{\varphi(p)}(d\varphi_p \cdot X_1, \ldots, d\varphi_p \cdot X_p)\,, \tag{3.36}$$

where $p \in M$ and $X_1, \ldots, X_p \in T_pM$.

Remark. As in the case of differential forms on Euclidean space (see Definition 3.16) the previous definition can be reformulated as

$$(\varphi^*\omega)_p = (d\varphi_p)^*\omega_{\varphi(p)}\,. \tag{3.37}$$

From the definition, it is clear that the pull-back is in $\bigwedge^p T_p^*M$ for all $p \in M$. However, it remains to show that $\varphi^*\omega \in \Omega_k^p(M)$. – Let (U, x) be a local chart on M. Due to (3.28), we can expand the pull-back on U as

$$\varphi^*\omega = \sum_I (\varphi^*\omega)_I \, dx_{i_1} \wedge \ldots \wedge dx_{i_p}\,,$$

where

$$\big(\varphi^*\omega\big)_I(p) = (\varphi^*\omega)_p \left(\frac{\partial}{\partial x_{i_1}}(p), \ldots, \frac{\partial}{\partial x_{i_p}}(p)\right) \in C^k(U, \mathbb{R})\,,$$

showing that $\varphi^*\omega \in \Omega_k^p(M)$. More precisely, this is a direct consequence of (3.23) which generalizes easily to differential forms on manifolds.

The properties of the pull-back for differential forms on $\mathbb{R}^n$, summarized in Proposition 3.17, remain true for differential forms on manifolds.

Proposition 3.24. *Let $\varphi \in C^\infty(M, N)$ and let $\omega \in \Omega^p(N)$.*

(i) For all $\tilde{\omega} \in \Omega^p(N)$, we have

$$\varphi^*(\omega + \tilde{\omega}) = \varphi^*\omega + \varphi^*\tilde{\omega}\,. \tag{3.38}$$

(ii) For all $\tilde{\omega} \in \Omega^q(N)$, we have

$$\varphi^*(\omega \wedge \tilde{\omega}) = \varphi^*\omega \wedge \varphi^*\tilde{\omega}\,. \tag{3.39}$$

(iii) For $\psi \in C^\infty(N, W)$ with W another differentiable manifold, we have

$$(\psi \circ \varphi)^*\gamma = \varphi^*(\psi^*\gamma)\,, \tag{3.40}$$

where $\gamma \in \Omega^p(W)$.

Definition 3.25. *Let $\omega \in \Omega^p_k(M)$ and (U,x) local chart of a C^∞-differentiable manifold M. We define on U the following element in $\Omega^{p+1}_{k-1}(M)$ denoted by $d\omega$:*

$$d\omega_p = \Big(x^*\, d\big((x^{-1})^*\omega\big)\Big)_p\,, \qquad p \in U\,. \tag{3.41}$$

*The form $d\omega \in \Omega^{p+1}_{k-1}(M)$ is called the **exterior derivative** of $\omega \in \Omega^p_k(M)$.*

We want to show that this definition is independent of the chart (U,x). – Let (U,y) be another chart, i.e., that $y : U \longrightarrow \mathbb{R}^m$ is a C^∞-diffeomorphism from U into $y(U)$. Then (3.40) implies

$$\begin{aligned} y^*\, d\big((y^{-1})^*\omega\big) &= y^*\, d\big((y^{-1})^* x^* (x^{-1})^*\omega\big) \\ &= y^*\, d\big((x \circ y^{-1})^* (x^{-1})^*\omega\big)\,, \end{aligned}$$

where $x \circ y^{-1} : y(U) \subset \mathbb{R}^m \longrightarrow \mathbb{R}^m$ is a C^∞-diffeomorphism. Note that $\alpha := (x^{-1})^*\omega \in \Omega^p\big(y(U)\big)$. Using that the operations d and the pull-back commute for a C^∞-map on Euclidean spaces (see Proposition 3.17), we obtain

$$\begin{aligned} y^*\, d\big((y^{-1})^*\omega\big) &= y^* (x \circ y^{-1})^*\, d\big((x^{-1})^*\omega\big) \\ &= (x \circ y^{-1} \circ y)^*\, d\big((x^{-1})^*\omega\big) \\ &= x^*\, d\big((x^{-1})^*\omega\big)\,. \end{aligned}$$

This shows that the previous definition is independent of the chart. Moreover, the fact that $d\omega \in \Omega^{p+1}_{k-1}(M)$ becomes clear due to the following proposition.

Proposition 3.26. *Let $\omega \in \Omega^p_k(M)$ and let (U,x) be a local chart for the C^∞-differentiable manifold M. Assume that ω has the local form*

$$\omega = \sum_I \omega_I\, dx_{i_1} \wedge \ldots \wedge dx_{i_p}\,.$$

Then for the exterior derivative of ω, we have the following local expression:

$$d\omega = \sum_I \sum_{l=1}^m \frac{\partial \omega_I}{\partial x_l}\, dx_l \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_p}\,. \tag{3.42}$$

Proof. We compute, using the local representation for ω,

$$\begin{aligned} d\omega &= x^*\, d\big((x^{-1})^*\omega\big) \\ &= x^*\, d\left((x^{-1})^*\Big(\sum_I \omega_I\, dx_{i_1} \wedge \ldots \wedge dx_{i_p}\Big)\right) \end{aligned}$$

From the linearity of the pull-back and (3.39), we then deduce

$$d\omega = x^* d\left(\sum_I \omega_I \circ x^{-1}\, \big((x^{-1})^* dx_{i_1}\big) \wedge \ldots \wedge \big((x^{-1})^* dx_{i_p}\big)\right)\,. \tag{3.43}$$

On the other hand, we claim that, for all $x \in x(U) \subset \mathbb{R}^m$,

$$\big((x^{-1})^* dx_k\big)_x = e_k^* \,. \tag{3.44}$$

Writing $v = \sum_{j=1}^m v^j \, e_j \in \mathbb{R}^m$, the linearity of the differential implies

$$(dx^{-1})_x \cdot v = \sum_{j=1}^m v^j \, (dx^{-1})_x \cdot e_j \overset{(2.26)}{=} \sum_{j=1}^m v^j \frac{\partial(x^{-1})}{\partial x_j}(x) \,.$$

This leads to

$$\begin{aligned}\big((x^{-1})^* dx_k\big)_x(v) &= (dx_k)_{x^{-1}(x)}\big((dx^{-1})_x \cdot v\big) \\ &= (dx_k)_{x^{-1}(x)} \left(\sum_{j=1}^m v^j \, \frac{\partial(x^{-1})}{\partial x_j}(x) \right) \,.\end{aligned}$$

The linearity of the differential one-form $dx_k \in \Omega^1(M)$ then gives, for the right-hand side,

$$\sum_{j=1}^m v^j dx_k\big(x^{-1}(x)\big) \left(\frac{\partial(x^{-1})}{\partial x_j}(x) \right) = v_k = e_k^*(v) \,,$$

showing the claim (3.44).

Inserting this in (3.43), we deduce

$$d\omega = x^* d \left(\sum_I \omega_I \circ x^{-1} \, e_{i_1}^* \wedge \ldots \wedge e_{i_p}^* \right) \,.$$

Using the expression (3.11) for the exterior derivative acting on differential forms on $\mathbb{R}^m$, we obtain (denoting by $(y_1, \ldots, y_m)$ the coordinates on $\mathbb{R}^m$)

$$d\omega = x^* \left(\sum_I \sum_{l=1}^m \frac{\partial}{\partial y_l}(\omega_I \circ x^{-1}) \, e_{i_1}^* \wedge \ldots \wedge e_{i_p}^* \right) \,.$$

The properties of the pull-back and (3.44), then imply

$$\begin{aligned} d\omega &= \sum_I \sum_{l=1}^m \frac{\partial}{\partial y_l}(\omega_I \circ x^{-1}) \circ x \; (x^* e_{i_1}^*) \wedge \ldots \wedge (x^* e_{i_p}^*) \\ &= \sum_I \sum_{l=1}^m \frac{\partial \omega_I}{\partial x_l} \, dx_l \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_p} \,, \end{aligned}$$

as claimed in (3.42). □

The properties of the exterior derivative d for differential forms on manifolds coincide with those of the exterior derivative for differential forms on Euclidean space. More precisely, we have the following

Proposition 3.27. *Let $\omega \in \Omega^p(M)$ and d the exterior derivative.*

(i) The map $d : \Omega^0(M) \longrightarrow \Omega^1(M)$ assigns to a function $f \in \Omega^0(M)$ its tangent map $df : TM \longrightarrow \mathbb{R}$.
(ii) For all $\tilde{\omega} \in \Omega^q(M)$ the following "Leibniz-rule" holds:

$$d(\omega \wedge \tilde{\omega}) = d\omega \wedge \tilde{\omega} + (-1)^p \omega \wedge d\tilde{\omega} . \tag{3.45}$$

(iii) We have that $d \circ d = 0$.
(iv) For all $\varphi \in C^\infty(M, N)$ and all $\gamma \in \Omega^p(N)$ the exterior derivative and the pull-back commute:

$$\varphi^*(d\gamma) = d(\varphi^*\gamma) . \tag{3.46}$$

Proof. The proof is strongly based on the analogous results for $\mathbb{R}^m$ in Section 3.1.2. – Let (U, x) be a local chart for M. *For (i)*: Let $f \in \Omega^0(M)$ and its tangent map df defined by (see Definition (2.40) and the following remark)

$$df_p \cdot [\gamma] = \frac{d}{dt}\bigg|_{t=0} \big(f \circ \gamma(t)\big) ,$$

where $[\gamma] \in T_pM$. In the other hand, the one-form df acts on the tangent vector $[\gamma]$ by

$$df\big([\gamma]\big) = \Big(x^*\, d\big((x^{-1})^* f\big)\Big)\big([\gamma]\big) = d\big((x^{-1})^* f\big)\big(dx \cdot [\gamma]\big) .$$

Since $(x^{-1})^* f = f \circ x^{-1}$, we then obtain

$$df\big([\gamma]\big) = d\big(f \circ x^{-1}\big)\big(dx \cdot [\gamma]\big) .$$

Because the exterior derivative of differential one-forms on $\mathbb{R}^m$ coincides with the tangent map, the last expression equals

$$d\big(f \circ x^{-1}\big) \cdot \big(dx \cdot [\gamma]\big) ,$$

and by definition of the tangent map, we get the result:

$$\begin{aligned} df\big([\gamma]\big) &= d\big(f \circ x^{-1}\big) \cdot \frac{d}{dt}\bigg|_{t=0} \big(x \circ \gamma(t)\big) \\ &= \frac{d}{dt}\bigg|_{t=0} \big(f \circ x^{-1} \circ x \circ \gamma(t)\big) \\ &= \frac{d}{dt}\bigg|_{t=0} \big(f \circ \gamma(t)\big) = df \cdot [\gamma] . \end{aligned}$$

For (ii): Let $\omega \in \Omega^p(M)$ and $\tilde{\omega} \in \Omega^q(M)$. Then (3.39) gives

$$d(\omega \wedge \tilde{\omega}) \overset{(3.41)}{=} x^*\, d\big((x^{-1})^*(\omega \wedge \tilde{\omega})\big) = x^*\, d\big((x^{-1})^*\omega \wedge (x^{-1})^*\tilde{\omega}\big) .$$

From the “Leibniz-rule” (3.13) for differential forms on $\mathbb{R}^m$, we obtain

$$d(\omega \wedge \tilde{\omega}) = x^* \Big(d\big((x^{-1})^*\omega\big) \wedge (x^{-1})^*\tilde{\omega} + (-1)^p (x^{-1})^*\omega \wedge d\big((x^{-1})^*\tilde{\omega}\big)\Big) .$$

Using the properties of the pull-back summarized in Proposition 3.24 and (3.20), we arrive at

$$\begin{aligned} d(\omega \wedge \tilde{\omega}) &= x^*\, d\big((x^{-1})^*\omega\big) \wedge \tilde{\omega} + (-1)^p \omega \wedge x^*\, d\big((x^{-1})^*\tilde{\omega}\big) \\ &= d\omega \wedge \tilde{\omega} + (-1)^p \omega \wedge d\tilde{\omega} . \end{aligned}$$

For (iii): Proposition 3.14 shows that

$$\begin{aligned} d(d\omega) &= d\Big(x^*\, d\big((x^{-1})^*\omega\big)\Big) \\ &= x^*\, d\Big((x^{-1})^*\big(x^*\, d\big((x^{-1})^*\omega\big)\big)\Big) \\ &= x^*\, d \circ d\big((x^{-1})^*\omega\big) = 0 . \end{aligned}$$

For (iv): For $\varphi \in C^\infty(M,N)$ and all $\gamma \in \Omega^p(N)$, we have, by definition of the exterior derivative,

$$d(\varphi^*\gamma) = x^*\, d\big((x^{-1})^*\, \varphi^*\gamma\big) .$$

The right-hand side can be written as

$$x^*\, d\big((x^{-1})^*\, \varphi^* x^* (x^{-1})^*\gamma\big) .$$

From (3.40), we then deduce

$$d(\varphi^*\gamma) = x^*\, d\Big(\big(x \circ \varphi \circ (x^{-1})\big)^*\, (x^{-1})^*\gamma\Big) .$$

Since the pull-back and the exterior derivative commute for forms on $\mathbb{R}^n$ (see (3.20)), we get

$$d(\varphi^*\gamma) = x^*\, \big(\big(x \circ \varphi \circ (x^{-1})\big)^*\, d\big((x^{-1})^*\gamma\big) .$$

Using again (3.40), we obtain easily the result

$$\begin{aligned} d(\varphi^*\gamma) &= x^*(x^{-1})^*\varphi^* x^*\, d\big((x^{-1})^*\gamma\big) \\ &= \varphi^* x^*\, d\big((x^{-1})^*\gamma\big) = \varphi^*(d\gamma) . \end{aligned}$$

□

The next operation on differential forms will be very useful in Section 3.4.

Definition 3.28. *Let M^m be a C^k-differentiable manifold, $p \in M^m$, and $\omega \in \Omega^p_{k-1}(M)$. Moreover, let X be a C^{k-1}-vector field on M^m. The **interior product** of ω and X, denoted by $int_X\omega$, is the following element of $\Omega^{p-1}_{k-1}(M)$:*

$$(int_X\omega)_p(X_1, \ldots, X_{p-1}) = \omega_p\big(X(p), X_1, \ldots, X_{p-1}\big) , \qquad (3.47)$$

where $X_1, \ldots, X_{p-1} \in T_pM^m$.

Remark. It is left as an exercise to check that the definition is well-posed (see Exercise 3.30).

Example 3.29. In the case of $p = m$, we compute the interior product explicitly. – Let (U, x) be a local chart for the C^∞-manifold M and let (see (3.28))

$$\omega = dx_1 \wedge \ldots \wedge dx_m \in \Omega^m(M)\,.$$

Moreover, let the C^∞-vector field X for $p \in U$ be given in the local representation (see (2.27))

$$X(p) = \sum_{i=1}^{m} X_i(p)\,\frac{\partial}{\partial x_i}(p)\,,$$

where $X_i \in C^\infty(U, \mathbb{R})$. Using (3.47), we then arrive at

$$\begin{aligned}
\mathrm{int}_X\omega(X_1, \ldots, X_{m-1}) &= dx_1 \wedge \ldots \wedge dx_m \left(\sum_{i=1}^{m} X_i\,\frac{\partial}{\partial x_i}, X_1, \ldots, X_{m-1}\right) \\
&= \sum_{i=1}^{m} X_i\, dx_1 \wedge \ldots \wedge dx_m \left(\frac{\partial}{\partial x_i}, X_1, \ldots, X_{m-1}\right) \\
&= \sum_{i=1}^{m} (-1)^{i-1} X_i\, dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_m\,(X_1, \ldots, X_{m-1})\,,
\end{aligned} \tag{3.48}$$

where the "hat" over dx_i means removing it. Thus, for the interior product of a m-form the following formula holds:

$$\mathrm{int}_X\omega = \sum_{i=1}^{m} (-1)^{i-1} X_i\, dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_m\,. \tag{3.49}$$

Exercise 3.30. Let M^m be a C^k-differentiable manifold. Prove that if $\omega \in \Omega^p_{k-1}(M)$ and if X a C^{k-1}-vector field on M^m, then $\mathrm{int}_X\omega \in \Omega^{p-1}_{k-1}(M)$. Prove it first for M^m being a subset of $\mathbb{R}^m$.

3.3 Restriction of Differential Forms to Submanifolds

Let N^n be a C^k-submanifold of $\mathbb{R}^m$. We already know that $(U_x, \varphi_x)_{x\in N^n}$ defined in Example 2.20 is an atlas for N^n. Moreover, denote the canonical inclusion map by

$$\iota_{N^n} : N^n \hookrightarrow \mathbb{R}^m\,, \tag{3.50}$$

Proposition 3.31. *Let N^n be a n-dimensional C^k-submanifold of $\mathbb{R}^m$. Then the inclusion map ι_{N^n} is a C^k-map from the manifold N^n into $\mathbb{R}^m$.*

Proof. First, we show that ι_{N^n} is continuous. – Let Ω by open in $\mathbb{R}^m$. Then $\iota_{N^n}^{-1}(\Omega) = \Omega \cap N^n$ is clearly open in N^n by definition of the restricted topology.

Let (U, φ) be a (canonical) local chart on N^n with $U \subset \mathbb{R}^m$ open. In order to show that $\iota_{N^n} \in C^k(N^n, \mathbb{R}^m)$, we have to show that

$$id_{\mathbb{R}^m} \circ \iota_{N^n} \circ \varphi^{-1} : \varphi(U \cap N^n) \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

is a C^k-map. It is evident that $id_{\mathbb{R}^m} \circ \iota_{N^n} \circ \varphi^{-1}(x) = \varphi^{-1}(x)$ for $x \in \varphi(U \cap N^n)$ which is an open set in $\mathbb{R}^n$. It follows from $\varphi = \pi \circ f|_{N^n \cap U}$, where f is a straightening map about U and $\pi : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ the canonical projection, that

$$\varphi^{-1} = \big(\pi \circ f|_{N^n \cap U}\big)^{-1} = f^{-1}|_{\mathbb{R}^n}\,.$$

Since by assumption N^n is a C^k-submanifold the straightening map f is a C^k-diffeomorphism and $f^{-1}|_{\mathbb{R}^n}$ is a C^k-map. □

With the help of this proposition the following definition makes sense.

Definition 3.32. *Let N^n be a $C^{k'}$-submanifold of $\mathbb{R}^m$ and let $\omega \in \Omega_k^p(\mathbb{R}^m)$ with $k' \geq k+1$. The* ***restriction*** *of ω to N^n, denoted by $\omega|_{N_n}$, is the following element of $\Omega_k^p(N^n)$:*

$$\begin{aligned}(\omega|_{N_n})_p\big(X_1, \ldots, X_p\big) &= (\iota_{N^n}^* \omega)_p\big(X_1, \ldots, X_p\big) \\ &= \omega_{\iota_{N^n}(p)}\big(d\iota_p \cdot X_1, \ldots, d\iota_p \cdot X_p\big)\,, \end{aligned} \tag{3.51}$$

where $p \in N^n$ and $X_1, \ldots, X_p \in T_p N^n$.

Remark. The inclusion map ι_{N^n} enables us to pass from "paths" $X_1, \ldots, X_p$ (see Definition 2.35) to vectors $d\iota_p \cdot X_1, \ldots, d\iota_p \cdot X_p$ in $\mathbb{R}^m$.

Example 3.33. Let $f \in \Omega^0(\mathbb{R}^m)$. Then by definition $\iota_{N^n}^* f = f \circ \iota_{N^n} = f|_{N^n}$, which is the usual restriction operation for functions. Moreover, using (3.20), the restriction to N^n of the tangent map is the tangent map of the restriction: $\iota_{N^n}^*(df) = d(\iota_{N^n}^* f) = d(f \circ \iota_{N^n})$. – Consider, for example, on $\mathbb{R}^3$ the function $f : x = (x, y, z) \longmapsto (x^2+y^2+z^2)/2$ and the sphere $S^2 = \{x \in \mathbb{R}^3 \,:\, |x|^2 = 1\}$. For the restriction on S^2, it then follows

$$\iota_{S^2}^*(df) = d\left(\frac{(x^2 + y^2 + z^2)}{2} \circ \iota_{S^2}\right) = d(\frac{1}{2}) = 0\,.$$

Restriction of one-forms to Curves in $\mathbb{R}^m$

In the following, we want to study the restriction of 1-forms to C^1-submanifolds with dimension $n = 1$. – Recall that a 1-dimensional C^1-submanifold of $\mathbb{R}^m$ can be seen as regular C^1-curve in $\mathbb{R}^m$.

Definition 3.34. *Let $\Gamma \in C^1([0,1],\mathbb{R}^m)$ be a regular curve in $\mathbb{R}^m$. An **orientation** for Γ is a continuous map $\boldsymbol{t} : \Gamma \longrightarrow \mathbb{R}^m$ such that $\|\boldsymbol{t}\| = 1$ and $\boldsymbol{t}(p) \in T_p\Gamma$, for all $p \in \Gamma$. Moreover, a chart (U,φ) on Γ is said to be **positively oriented** with respect to the orientation $\boldsymbol{t}$ if $0 < d\varphi_p \cdot \boldsymbol{t}(p) \in \mathbb{R}$, for all $p \in U$.*

Definition 3.35. *Let Γ be a regular C^1-curve on $\mathbb{R}^m$ with an orientation $\boldsymbol{t}$ and (U,φ) a positively oriented chart on Γ. The **length form** dl_Γ on U with respect to $\boldsymbol{t}$ is defined by*

$$dl_\Gamma = \|d\varphi^{-1} \cdot e_1\|\, d\varphi \in \Omega^1_0(\Gamma \cap U)\,, \tag{3.52}$$

where $\|\cdot\|$ denotes the usual norm on $\mathbb{R}^m$ and e_1 the canonical basis of $\mathbb{R}$.

For (U,ψ) another positively oriented chart, we compute

$$\|d\psi^{-1} \cdot e_1\|\, d\psi = \|d\psi^{-1} \cdot e_1\|\, d(\psi \circ \varphi^{-1} \circ \varphi) = \|d\psi^{-1} \cdot e_1\|\, d(\psi \circ \varphi^{-1}) \cdot d\varphi\,.$$

Let $v = \lambda\, e_1 \in \mathbb{R}$ with $\lambda \in \mathbb{R}$. Since the charts ψ and φ are both positively oriented with respect to $\boldsymbol{t}$, it follows that

$$d(\psi \circ \varphi^{-1}) \cdot v = \lambda\, d(\psi \circ \varphi^{-1}) \cdot e_1 = \lambda\, |d(\psi \circ \varphi^{-1}) \cdot e_1|\, e_1.$$

Setting $d\varphi(\cdot) = \lambda(\cdot)\, e_1$, we arrive at

$$\begin{aligned} \|d\psi^{-1} \cdot e_1\|\, d\psi &= \|d\psi^{-1} \cdot e_1\|\, |d(\psi \circ \varphi^{-1}) \cdot e_1|\, d\varphi \\ &= \|d\psi^{-1} \cdot d(\psi \circ \varphi^{-1}) \cdot e_1\|\, d\varphi \\ &= \|d\varphi^{-1} \cdot e_1\|\, d\varphi\,, \end{aligned}$$

proving that the definition of the length form makes sense.

Proposition 3.36. *Let Γ be a regular C^1-curve in $\mathbb{R}^m$ with an orientation $\boldsymbol{t} = (t_1,\ldots,t_m)$ and let $\omega = \sum_{i=1}^m \omega_i\, dx_i \in \Omega^1_0(\mathbb{R}^3)$. Then the restriction of ω to Γ equals*

$$\omega|_\Gamma = \iota_\Gamma^*\omega = \sum_{i=1}^m \omega_i t_i\, dl_\Gamma \in \Omega^1_0(\Gamma)\,, \tag{3.53}$$

where $\iota_\Gamma : \Gamma \hookrightarrow \mathbb{R}^m$ is the canonical inclusion.

Proof. Let (U,φ), $U \subset \mathbb{R}^m$ open, be a positively oriented chart for Γ. We first prove the proposition for $\omega = dx_k$, with $k = 1,\ldots,m$. Since $x_k \circ \iota_\Gamma = (\varphi^{-1})_k \circ \varphi : \Gamma \cap U \longrightarrow \mathbb{R}$ and using Example 3.33, it follows that

$$\iota_\Gamma^*\omega = \iota_\Gamma^* dx_k = d(x_k \circ \iota_\Gamma) = d(\varphi^{-1})_k \cdot d\varphi\,.$$

With $v = \lambda\, e_1 \in \mathbb{R}$, we obtain that

$$d(\varphi^{-1}) \cdot v = \lambda\big(d(\varphi^{-1}) \cdot e_1\big) = \lambda\, \|d(\varphi^{-1}) \cdot e_1\|\, \boldsymbol{t}\,,$$

and hence that $d(\varphi^{-1})_k \cdot v = \lambda \, \|d(\varphi^{-1}) \cdot e_1\| \, t_k$. If $d\varphi(\cdot) = \lambda(\cdot)\, e_1$, we thus arrive at

$$\iota_\Gamma^* \omega = d(\varphi^{-1})_k \cdot d\varphi = \|d(\varphi^{-1}) \cdot e_1\| \, t_k \, d\varphi = t_k \, dl_\Gamma \,,$$

where we used Definition 3.35 of the length form. – We conclude using the linearity of the exterior derivative. □

Restriction of two-forms to Surfaces in $\mathbb{R}^3$

In a next step, we want to study the restriction of 2-forms to a surface in $\mathbb{R}^3$. Recall that by definition a surface in $\mathbb{R}^3$ is a 2-dimensional C^1-submanifold of $\mathbb{R}^3$.

Definition 3.37. *Let Σ be a surface in $\mathbb{R}^3$. An* ***orientation*** *for Σ is a continuous map $\boldsymbol{n} : \Sigma \longrightarrow \mathbb{R}^3$ such that $\|\boldsymbol{n}\| = 1$ and $\boldsymbol{n}(p) \perp T_p\Sigma$, for all $p \in \Sigma$. We also call the map $\boldsymbol{n}$ a* ***unit normal*** *vector field to Σ.*

Definition 3.38. *Let Σ be a surface in $\mathbb{R}^3$ with an orientation $\boldsymbol{n}$. The* ***area form*** *dA_Σ with respect to $\boldsymbol{n}$ on U is defined by*

$$dA_\Sigma = \left\langle \frac{\partial}{\partial \varphi_1} \times \frac{\partial}{\partial \varphi_2} \,, \boldsymbol{n} \right\rangle d\varphi_1 \wedge d\varphi_2 \in \Omega_0^2(\Sigma \cap U) \,, \tag{3.54}$$

where (U, φ) is a local chart for Σ, $\langle \cdot \,, \cdot \rangle$ denotes the scalar product in $\mathbb{R}^3$ and $\times$ denotes the usual vector product in $\mathbb{R}^3$.

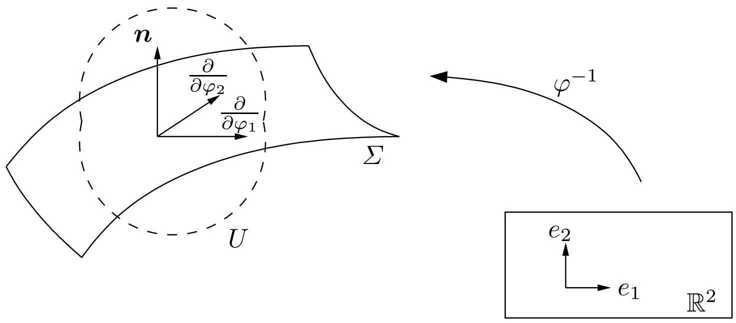

Fig. 3.2. Surface in $\mathbb{R}^3$.

In order to show that this definition is well-posed, it is necessary to check that it is independent of the local chart (U, φ). – Let (V, ψ) be another chart about $p \in \Sigma$. It follows from (3.42) that, for $i = 1, 2$,

$$d\varphi_i = \frac{\partial \varphi_i}{\partial \psi_1} \, d\psi_1 + \frac{\partial \varphi_i}{\partial \psi_2} \, d\psi_2 \,, \tag{3.55}$$

and hence

$$d\varphi_1 \wedge d\varphi_2 = \left(\frac{\partial \varphi_1}{\partial \psi_1} \frac{\partial \varphi_2}{\partial \psi_2} - \frac{\partial \varphi_1}{\partial \psi_2} \frac{\partial \varphi_2}{\partial \psi_1} \right) d\psi_1 \wedge d\psi_2 \,. \tag{3.56}$$

Moreover, let (e_1, e_2) denote the canonical basis on $\mathbb{R}^2$, and we write $\psi^{-1} = \varphi^{-1} \circ \varphi \circ \psi^{-1}$ to get

$$\frac{\partial}{\partial \psi_i} = d\psi^{-1} \cdot e_i = d\varphi^{-1} \big(d(\varphi \circ \psi^{-1}) \cdot e_i \big) \,. \tag{3.57}$$

The term in parenthesis reduces to

$$d(\varphi \circ \psi^{-1}) \cdot e_i = d\varphi \cdot d\psi^{-1} \cdot e_i = d\varphi \cdot \frac{\partial}{\partial \psi_i} = \frac{\partial \varphi}{\partial \psi_i} \,.$$

The last expression can then be written as

$$\frac{\partial \varphi}{\partial \psi_i} = \frac{\partial \varphi_1}{\partial \psi_i} \, e_1 + \frac{\partial \varphi_2}{\partial \psi_i} \, e_2 \,.$$

Inserting this result into (3.57) leads to

$$\begin{aligned} \frac{\partial}{\partial \psi_i} &= \frac{\partial \varphi_1}{\partial \psi_i} \, d\varphi^{-1} \cdot e_1 + \frac{\partial \varphi_2}{\partial \psi_i} \, d\varphi^{-1} \cdot e_2 \\ &= \frac{\partial \varphi_1}{\partial \psi_i} \, \frac{\partial}{\partial \varphi_1} + \frac{\partial \varphi_2}{\partial \psi_i} \, \frac{\partial}{\partial \varphi_2} \,. \end{aligned}$$

From this computation we also obtain

$$\begin{aligned} \left\langle \frac{\partial}{\partial \psi_1} \times \frac{\partial}{\partial \psi_2} \,, \boldsymbol{n} \right\rangle &= \left\langle \left(\frac{\partial \varphi_1}{\partial \psi_1} \, \frac{\partial}{\partial \varphi_1} + \frac{\partial \varphi_2}{\partial \psi_1} \, \frac{\partial}{\partial \varphi_2} \right) \right. \\ &\qquad \left. \times \left(\frac{\partial \varphi_1}{\partial \psi_2} \, \frac{\partial}{\partial \varphi_1} + \frac{\partial \varphi_2}{\partial \psi_2} \, \frac{\partial}{\partial \varphi_2} \right) \,, \boldsymbol{n} \right\rangle \\ &= \left(\frac{\partial \varphi_1}{\partial \psi_1} \frac{\partial \varphi_2}{\partial \psi_2} - \frac{\partial \varphi_1}{\partial \psi_2} \frac{\partial \varphi_2}{\partial \psi_1} \right) \left\langle \frac{\partial}{\partial \varphi_1} \times \frac{\partial}{\partial \varphi_2} \,, \boldsymbol{n} \right\rangle \,. \end{aligned}$$

Combining this result with (3.56), we arrive at

$$\left\langle \frac{\partial}{\partial \varphi_1} \times \frac{\partial}{\partial \varphi_2} \,, \boldsymbol{n} \right\rangle d\varphi_1 \wedge d\varphi_2 = \left\langle \frac{\partial}{\partial \psi_1} \times \frac{\partial}{\partial \psi_2} \,, \boldsymbol{n} \right\rangle d\psi_1 \wedge d\psi_2 \,,$$

showing that the definition of the area form for a surface is independent of the local chart.

Remark. For explicit calculations (see below), it is useful to rewrite (3.54) in the following way:

$$dA_\Sigma = \left\langle \frac{\partial \varphi^{-1}}{\partial x_1} \times \frac{\partial \varphi^{-1}}{\partial x_2} \,, \boldsymbol{n} \right\rangle d\varphi_1 \wedge d\varphi_2 \in \Omega^2_0(\Sigma \cap U) \,, \tag{3.58}$$

where (x_1, x_2) denote the canonical coordinates on $\mathbb{R}^2$. Recall that we have $d\varphi^{-1} \cdot e_i = \frac{\partial \varphi^{-1}}{\partial x_i}$, for $i = 1, 2$.

Example 3.39. In the case of Σ being the *graph* of a function, we define

$$\begin{aligned}\varphi^{-1} : U' \subset \mathbb{R}^2 &\longrightarrow U \cap \Sigma \subset \mathbb{R}^3\,,\\ (x_1, x_2) &\longmapsto \big(x_1, x_2, h(x_1, x_2)\big)\,,\end{aligned} \tag{3.59}$$

where $h : U' \longrightarrow \mathbb{R}$ is a C^1-function with $U' \subset \mathbb{R}^2$ open. Inserting (3.59) into (3.58), the area form dA_Σ with respect to the orientation

$$\boldsymbol{n} = \frac{\dfrac{\partial\varphi^{-1}}{\partial x_1} \times \dfrac{\partial\varphi^{-1}}{\partial x_2}}{\left\| \dfrac{\partial\varphi^{-1}}{\partial x_1} \times \dfrac{\partial\varphi^{-1}}{\partial x_2} \right\|}\,,$$

then reads as

$$dA_\Sigma = \sqrt{1 + \left(\frac{\partial h}{\partial x_1}\right)^2 + \left(\frac{\partial h}{\partial x_2}\right)^2}\, d\varphi_1 \wedge d\varphi_2\,. \tag{3.60}$$

Area Form for the Sphere S^2

We calculate explicitly the area form for the two-sphere S^2. For this purpose, we choose spherical coordinates. – Let the inverse of φ be given by

$$\begin{aligned}\varphi^{-1} : U' \subset \mathbb{R}^2 &\longrightarrow S^2 \subset \mathbb{R}^3\,,\\ (\psi, \theta) &\longmapsto \big(\cos\psi\cos\theta, \sin\psi\cos\theta, \sin\theta\big)\,.\end{aligned} \tag{3.61}$$

where $U' = (-\pi, \pi) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. In other words, we want to calculate the area form dA_{S^2} in the local chart (U, φ) with the inverse of φ defined in (3.61) and $S^2 \cap U = \varphi^{-1}(U')$. Note that φ^{-1} is often called a **(local) parameterization** (see the remark at the end of this section).

By a straightforward computation, it is then easy to check that

$$\left\langle \frac{\partial\varphi^{-1}}{\partial\psi} \times \frac{\partial\varphi^{-1}}{\partial\theta}\,, \boldsymbol{n} \right\rangle = |\cos\theta| = \cos\theta\,, \tag{3.62}$$

where the usual orientation for S^2, given by

$$\boldsymbol{n} = \frac{\dfrac{\partial\varphi^{-1}}{\partial\psi} \times \dfrac{\partial\varphi^{-1}}{\partial\theta}}{\left\| \dfrac{\partial\varphi^{-1}}{\partial\psi} \times \dfrac{\partial\varphi^{-1}}{\partial\theta} \right\|}\,, \tag{3.63}$$

is used. Denoting the coordinate functions of φ also by (ψ, θ) and inserting (3.62) into (3.58), we arrive at

$$dA_{S^2} = \cos\theta\, d\psi \wedge d\theta \in \Omega^2(S^2 \cap U)\,. \tag{3.64}$$

Remark. The notation for the spherical coordinate functions (ψ, θ) of φ is also used for the coordinates on $U' \subset \mathbb{R}^2$. The reason for the choice of this notation will become clear later when we consider the integral of the area form dA_{S^2} (see Example 4.10).

Now, we consider only the northern half-sphere, denoted by S^2_+, which can be interpreted as graph of a function (see Example 3.39). More precisely, we can write

$$S^2_+ = \Big\{ \big(x_1, x_2, h(x_1, x_2)\big) \ : \ x_1^2 + x_2^2 < 1 \Big\} ,$$

where the function h reads as

$$h(x_1, x_2) = \sqrt{1 - x_1^2 - x_2^2} \, . \tag{3.65}$$

Inserting (3.65) into (3.60), we obtain for the area form

$$dA_{S^2_+} = \frac{1}{\sqrt{1 - x_1^2 - x_2^2}} \, d\varphi_1 \wedge d\varphi_2 \, . \tag{3.66}$$

Next, we prove an analogous result to Proposition 3.36 for area forms.

Proposition 3.40. *Let Σ be a surface in $\mathbb{R}^3$ with an orientation $\boldsymbol{n}$ and let*

$$\omega = \omega_1 \, dx_2 \wedge dx_3 + \omega_2 \, dx_3 \wedge dx_1 + \omega_3 \, dx_1 \wedge dx_2 \in \Omega^2_0(\mathbb{R}^3) \, .$$

Then the restriction of ω to Σ equals

$$\omega|_\Sigma = \iota_\Sigma^* \omega = \sum_{i=1}^{3} \omega_i n_i \, dA_\Sigma \in \Omega^2_0(\Sigma) \, , \tag{3.67}$$

where $\boldsymbol{n} = (n_1, n_2, n_3)$ is the unit normal vector field defining the orientation and $\iota_\Sigma : \Sigma \hookrightarrow \mathbb{R}^3$ the canonical inclusion.

Proof. By linearity of the pull-back ι_Σ^*, it suffices to prove the formula for, say, $\omega = dx_1 \wedge dx_2$. – From Proposition 3.17, it follows

$$\begin{aligned} \iota_\Sigma^*(dx_1 \wedge dx_2) &= \iota_\Sigma^* dx_1 \wedge \iota_\Sigma^* dx_2 \\ &= d(\iota_\Sigma^* x_1) \wedge d(\iota_\Sigma^* x_2) \\ &= d(x_1 \circ \iota_\Sigma) \wedge d(x_2 \circ \iota_\Sigma) \, . \end{aligned}$$

If (U, φ) is a local chart on Σ, we also get

$$x_1 \circ \iota_\Sigma = x_1|_\Sigma = (\varphi^{-1})_1 \circ \varphi : \Sigma \cap U \longrightarrow \mathbb{R} \, ,$$

and hence

$$\iota_\Sigma^*(dx_1 \wedge dx_2) = d\big((\varphi^{-1})_1 \circ \varphi\big) \wedge d\big((\varphi^{-1})_2 \circ \varphi\big) \, . \tag{3.68}$$

Clearly, from

$$d(\varphi^{-1})_i = \frac{\partial(\varphi^{-1})_i}{\partial x_1}\,dx_1 + \frac{\partial(\varphi^{-1})_i}{\partial x_2}\,dx_2\,, \qquad i = 1, 2\,,$$

we obtain that

$$\begin{aligned} d\big((\varphi^{-1})_i \circ \varphi\big) &= d(\varphi^{-1})_i \cdot d\varphi \\ &= \frac{\partial(\varphi^{-1})_i}{\partial x_1}\,dx_1 \cdot d\varphi + \frac{\partial(\varphi^{-1})_i}{\partial x_2}\,dx_2 \cdot d\varphi \\ &= \frac{\partial(\varphi^{-1})_i}{\partial x_1}\,d\varphi_1 + \frac{\partial(\varphi^{-1})_i}{\partial x_2}\,d\varphi_2\,. \end{aligned}$$

Inserting this result in (3.68), leads to

$$d\big((\varphi^{-1})_1 \circ \varphi\big) \wedge d\big((\varphi^{-1})_2 \circ \varphi\big) = \left(\frac{\partial(\varphi^{-1})_1}{\partial x_1}\frac{\partial(\varphi^{-1})_2}{\partial x_2} - \frac{\partial(\varphi^{-1})_1}{\partial x_2}\frac{\partial(\varphi^{-1})_2}{\partial x_1}\right) d\varphi_1 \wedge d\varphi_2\,.$$

Moreover, we easily see that

$$\frac{\partial(\varphi^{-1})_1}{\partial x_1}\frac{\partial(\varphi^{-1})_2}{\partial x_2} - \frac{\partial(\varphi^{-1})_1}{\partial x_2}\frac{\partial(\varphi^{-1})_2}{\partial x_1} = \left\langle \frac{\partial\varphi^{-1}}{\partial x_1} \times \frac{\partial\varphi^{-1}}{\partial x_2}\,,\, e_3 \right\rangle .$$

Hence, so far we have shown that

$$\iota_\Sigma^*(dx_1 \wedge dx_2) = \left\langle \frac{\partial\varphi^{-1}}{\partial x_1} \times \frac{\partial\varphi^{-1}}{\partial x_2}\,,\, e_3 \right\rangle d\varphi_1 \wedge d\varphi_2\,. \tag{3.69}$$

We see that $\frac{\partial\varphi^{-1}}{\partial x_1} \times \frac{\partial\varphi^{-1}}{\partial x_2}$ is parallel to $\boldsymbol{n}$ and doesn't vanish. Hence, it follows that

$$\iota_\Sigma^*(dx_1 \wedge dx_2) = \left\langle \frac{\partial\varphi^{-1}}{\partial x_1} \times \frac{\partial\varphi^{-1}}{\partial x_2}\,,\, \boldsymbol{n} \right\rangle \langle e_3\,, \boldsymbol{n}\rangle\, d\varphi_1 \wedge d\varphi_2\,.$$

Since by assumption $\omega_1 = 0$, $\omega_2 = 0$ and $\omega_3 = 1$, this becomes

$$\iota_\Sigma^*(dx_1 \wedge dx_2) = \sum_{i=1}^{3} \omega_i n_i\, dA_\Sigma\,.$$

□

Remark. Recall that by a chart (U, φ) for a surface Σ we mean a C^1-map $\varphi : \Sigma \cap U \longrightarrow \mathbb{R}^2$, where U is an open subset of $\mathbb{R}^3$. Hence, for the inverse we have $\varphi^{-1} : \varphi(\Sigma \cap U) \subset \mathbb{R}^2 \longrightarrow \Sigma \cap U$. Moreover, it follows that

$$d\varphi^{-1} \cdot e_i = \frac{\partial\varphi^{-1}}{\partial x_i}\,,$$

for $i = 1, 2$, are not vectors in $\mathbb{R}^3$. However, considering the map (local parameterization)

$$\iota_\Sigma \circ \varphi^{-1} : \varphi(\Sigma \cap U) \subset \mathbb{R}^2 \longrightarrow \Sigma \cap U \subset \mathbb{R}^3 ,$$

instead of φ^{-1}, we obtain

$$\frac{\partial(\iota_\Sigma \circ \varphi^{-1})}{\partial x_i} = d(\iota_\Sigma \circ \varphi^{-1}) \cdot e_i = d\iota_\Sigma \cdot d\varphi^{-1} \cdot e_i = d\iota_\Sigma \cdot \frac{\partial \varphi^{-1}}{\partial x_i} ,$$

which are indeed vectors in $\mathbb{R}^3$. In this section, there is no difference in the notation for φ^{-1} and its composition with ι_Σ, called later $\tilde{\varphi}^{-1}$ (see Section 5.2).

3.4 Orientation of Manifolds

First, we need the concept of orientation on $\mathbb{R}^n$. – Let $\{e_i\}_{i=1,\ldots,n}$ and $\{e'_i\}_{i=1,\ldots,n}$ denote two bases of $\mathbb{R}^n$. Then they are said to have the *same orientation* if $\det P_{e_i}^{e'_j} > 0$, where P represents the transformation matrix for passing from one basis to the other. One easily checks that this is an equivalence relation on the set of all bases of $\mathbb{R}^n$ and that there are exactly two equivalence classes. These two equivalence classes represent the two *orientations* of $\mathbb{R}^n$.

Definition 3.41. *An* ***orientation*** *of* $\mathbb{R}^n$ *is given by the choice of one equivalence class. And the bases being in this equivalence class are called* ***positively oriented****.*

In the following, we will see that the notion of orientation is strongly related to the choice of a basis for $\bigwedge^n(\mathbb{R}^n)$. Recall that this vector space with basis $e_1^* \wedge \ldots \wedge e_n^*$ has dimension 1 (see Proposition 3.7). – Let $\{e_i\}_{i=1,\ldots,n}$ be a basis of $\mathbb{R}^n$ and $v_1, \ldots, v_n \in \mathbb{R}^n$ such that $v_i = \sum_{j=1}^n v_i^j \, e_j$, for $i = 1, \ldots, n$. We then compute, using (3.6) and the definition of the determinant,

$$\begin{aligned} e_1^* \wedge \ldots \wedge e_n^*(v_1, \ldots, v_n) &= \sum_{\sigma \in \mathcal{S}_n} (-1)^{|\sigma|} v_1^{i_{\sigma(1)}} \ldots v_n^{i_{\sigma(n)}} \\ &= \det(v_1^j, \ldots, v_n^j) . \end{aligned} \tag{3.70}$$

This computation obviously generalizes to an arbitrary basis of $\bigwedge^n(\mathbb{R}^n)$, i.e., a non-vanishing element of $\bigwedge^n(\mathbb{R}^n)$. We denote them by $\alpha_{\mathbb{R}^n}$.

Proposition 3.42. *A non-vanishing* $\alpha_{\mathbb{R}^n} \in \bigwedge^n(\mathbb{R}^n)$ *has the same sign on two bases of* $\mathbb{R}^n$ *if and only if they have the same orientation.*

Proof. This proposition is a direct consequence of (3.70). □

Now, we can reformulate Definition 3.41 in terms of a basis of $\bigwedge^n(\mathbb{R}^n)$.

Definition 3.43. *A choice of a basis* $\alpha_{\mathbb{R}^n} \in \bigwedge^n(\mathbb{R}^n)$ *gives an orientation of* $\mathbb{R}^n$*.*

Remark. Two bases $\alpha_{\mathbb{R}^n}$ and $\tilde{\alpha}_{\mathbb{R}^n}$ of $\bigwedge^n(\mathbb{R}^n)$ determine the same orientation if and only if $\alpha_{\mathbb{R}^n} = \lambda\, \tilde{\alpha}_{\mathbb{R}^n}$ for $0 < \lambda \in \mathbb{R}$.

For later use, we consider now the linear map

$$\begin{aligned} \varphi : \mathbb{R}^n &\longrightarrow \mathbb{R}^n\,, \\ e_i &\longmapsto \varphi(e_i) = \sum_{j=1}^{n} \varphi^j(e_i)\, e_j\,, \end{aligned}$$

and deduce from (3.70) that

$$e_1^* \wedge \ldots \wedge e_n^*\big(\varphi(e_1), \ldots, \varphi(e_n)\big) = \det\big(\varphi^j(e_1), \ldots, \varphi^j(e_n)\big)\,.$$

Moreover, from Definition 3.8 of the pull-back, it follows

$$\begin{aligned} \big(\varphi^*\, e_1^* \wedge \ldots \wedge e_n^*\big)(e_1, \ldots, e_n) &= e_1^* \wedge \ldots \wedge e_n^*\big(\varphi(e_1), \ldots, \varphi(e_n)\big) \\ &= \det\big(\varphi^j(e_1), \ldots, \varphi^j(e_n)\big)\,, \end{aligned}$$

showing that

$$\varphi^*(e_1^* \wedge \ldots \wedge e_n^*) = \det \varphi\, e_1^* \wedge \ldots \wedge e_n^* \tag{3.71}$$

Remark 3.44. From (3.71) and Definition 3.16 for the pull-back of differential forms on $\mathbb{R}^n$, it can be seen that the pull-back of a "maximal" differential form by a map is just multiplication by the determinant of the Jacobian of the map. This important observation will be used at various places in the following and generalizes naturally to differential forms on manifolds.

We come back to the notion of orientation. – Note that the orientation of an arbitrary vector space, in particular of a tangent space of a manifold, can be defined in a completely analogous manner. This enables us to extend the concept of orientation of $\mathbb{R}^n$ to manifolds. – In this section, we assume for simplicity the manifolds to be C^∞.

Definition 3.45. *Let M^m be a m-dimensional manifold. A **volume form** ω_{M^m} on M^m is an element of $\Omega^m(M)$ which never vanishes.*

Definition 3.46. *A manifold is said to be **orientable** if it posseses a volume form. The choice of ω_{M^m} determines the **orientation** of M^m.*

Remark. Translating the results for the orientation of $\mathbb{R}^n$ to the tangent spaces of a manifold, any such ω_{M^m} thus orients each tangent space T_pM in such a way that the orientations of nearby tangent spaces are compatible.

The following proposition, together with Proposition 3.52 below, can be interpreted as second definition of orientability for a manifold.

Proposition 3.47. *Let M^m be a manifold. If M^m is orientable, then there exists an atlas $(U_i, \varphi_i)_{i\in I}$ whose transition functions have positive Jacobian, i.e., for all $p \in \varphi_i(U_i \cap U_j)$, we have that*

$$det\, J_p(\varphi_j \circ \varphi_i^{-1}) > 0\,. \tag{3.72}$$

Such an atlas is called ***positively oriented****.*

Proof. Let ω_{M^m} be the volume form on M^m and let $(U_i, \varphi_i)_{i\in I}$ be an atlas. Then, we write

$$(\varphi_i^{-1})^*\omega_{M^m} = f_i\, dx_1 \wedge \ldots \wedge dx_m \in \Omega^m(\varphi_i(U_i))\,, \tag{3.73}$$

where $f_i : \varphi_i(U_i) \longrightarrow \mathbb{R}$. Since ω_{M^m} is a volume form on M^m, it follows by definition that f_i is a non-vanishing function on $\varphi_i(U_i)$. If f_i is positive, we are done. If f_i is negative, however, we have to modify the diffeomorphism φ_i by a negative isometry $\sigma : \mathbb{R}^m \longrightarrow \mathbb{R}^m$, i.e., $\det\sigma = -1$, to obtain (see Remark 3.44)

$$\begin{aligned}\big((\sigma\circ\varphi_i)^{-1}\big)^*\omega_{M^m} &= (\varphi_i^{-1}\circ\sigma^{-1})^*\omega_{M^m}\\ &= (\sigma^{-1})^*\big(f_i\, dx_1 \wedge \ldots \wedge dx_m\big)\\ &= -f_i\circ\sigma^{-1}\, dx_1 \wedge \ldots \wedge dx_m\,.\end{aligned}$$

Thus, we can construct an atlas $(U_i, \varphi_i)_{i\in I}$ such that $f_i > 0$, for all $i \in I$.

Considering transition functions on $\varphi_i(U_i \cap U_j)$, we now write

$$\begin{aligned}(\varphi_i^{-1})^*\omega_{M^m} &= (\varphi_i^{-1})^*\varphi_j^*(\varphi_j^{-1})^*\omega_{M^m}\\ &= (\varphi_j\circ\varphi_i^{-1})^*\big((\varphi_j^{-1})^*\omega_{M^m}\big)\,.\end{aligned}$$

Using (3.73), it then follows

$$(\varphi_i^{-1})^*\omega_{M^m} = (\varphi_j\circ\varphi_i^{-1})^*\big(f_j\, dx_1 \wedge \ldots \wedge dx_m\big)\,.$$

From the explicit expression for the pull-back given in (3.23) and standard properties of differential forms on $\mathbb{R}^m$, we deduce[1]

$$\begin{aligned}&(\varphi_i^{-1})^*\omega_{M^m} = f_j\big((\varphi_j\circ\varphi_i^{-1})\big)(\varphi_j\circ\varphi_i^{-1})^*\, dx_1 \wedge \ldots \wedge dx_m\\ &= f_j\big((\varphi_j\circ\varphi_i^{-1})\big)\sum_{i_1,\ldots,i_m}\frac{\partial(\varphi_j\circ\varphi_i^{-1})_1}{\partial x_{i_1}}\cdots\frac{\partial(\varphi_j\circ\varphi_i^{-1})_m}{\partial x_{i_m}}\, dx_{i_1}\wedge\ldots\wedge dx_{i_m}\\ &= f_j\big((\varphi_j\circ\varphi_i^{-1})\big)\sum_{\sigma\in\mathcal{S}_m}\frac{\partial(\varphi_j\circ\varphi_i^{-1})_1}{\partial x_{i_{\sigma(1)}}}\cdots\frac{\partial(\varphi_j\circ\varphi_i^{-1})_m}{\partial x_{i_{\sigma(m)}}}\, dx_{i_{\sigma(1)}}\wedge\ldots\wedge dx_{i_{\sigma(m)}}\\ &= f_j\big((\varphi_j\circ\varphi_i^{-1})\big)\sum_{\sigma\in\mathcal{S}_m}\frac{\partial(\varphi_j\circ\varphi_i^{-1})_1}{\partial x_{i_{\sigma(1)}}}\cdots\frac{\partial(\varphi_j\circ\varphi_i^{-1})_m}{\partial x_{i_{\sigma(m)}}}\,(-1)^{|\sigma|}\, dx_1\wedge\ldots\wedge dx_m\\ &= f_j\big((\varphi_j\circ\varphi_i^{-1})\big)\det J(\varphi_j\circ\varphi_i^{-1})\, dx_1\wedge\ldots\wedge dx_m\,,\end{aligned} \tag{3.74}$$

[1] We give an explicite calculation for a direct consequence of Remark 3.44.

where the definition of the determinant is used in the last line. Inserting (3.73) with $f_i > 0$, for all $i \in I$, in the left-hand side of the last equation, then shows by comparison that $\det J(\varphi_j \circ \varphi_i^{-1}) > 0$ on $\varphi_i(U_i \cap U_j)$, for every $i, j \in I$. $\square$

Remark. Note that being positively oriented as defined in (3.72) gives an equivalence relation on the set of all systems of charts on M^m with exactly two equivalence classes. And the choice of one of these two equivalence classes determines the orientation of the manifold.

In the following, some more definitions and results will be given in order to show the converse of the last proposition.

Definition 3.48. *Let X be a topological space for the topology τ. We say that τ is **separable** if there exists $\mathcal{B} \subset \tau$ such that $\mathcal{B}$ is countable and every element of the topology τ can be written as an union of elements of $\mathcal{B}$.*

This naturally leads to the definition of topological *separable* manifolds which will be always considered from now on.

Definition 3.49. *Let X be a topological space. A **locally finite covering** for X consists of a family of open sets $(U_i)_{i \in I}$ such that*

a) it covers the topological space, i.e., $X \subset \bigcup_{i \in I} U_i$;
b) for all $x \in X$, there exists U_x open containing x such that it intersects only finitely many U_i, $i \in I$.

Definition 3.50. *A **partition of unity** on a manifold M^m is a family $(U_i, \theta_i)_{i \in I}$, where $(U_i)_{i \in I}$ is a locally finite covering for M^m and the functions $\theta_i \in C^\infty(M^m, \mathbb{R})$ for all $i \in I$ are chosen such that*

a) they are non-negative, i.e., $\theta_i(p) \geq 0$ for all $p \in M^m$;
b) their support is contained in U_i, i.e., $\operatorname{supp} \theta_i \subset U_i$;
c) and for all $p \in M^m$ the following equation holds:

$$1 \equiv \sum_{i \in I} \theta_i(p) \,. \tag{3.75}$$

Remark. Note that from b) and the fact that the covering is locally finite, we obtain that for every $p \in M^m$ only a *finite* number of functions θ_i are non-zero in p. This implies that (3.75) is well-defined.

Now, we are ready to state an important result for separable manifolds which will be essential in the proof of the converse of Proposition 3.52.

Proposition 3.51. *Let M^m be a separable manifold. Then there exists a partition of unity on M^m. Moreover, let $(U_j)_{j \in J}$ be a covering of M^m. Then there exist $I \subset J$ and a family $(\theta_i)_{i \in I}$ of C^∞-functions on M^m such that $(U_i, \theta_i)_{i \in I}$ is a partition of unity for M^m. Such a partition of unity is called **subordinated** to $(U_j)_{j \in J}$.*

Proof. For the proof we refer to []. □

Remark. It is important to note that the initial covering $(U_j)_{j\in J}$ is *not* assumed to be locally finite.

We now arrive at the converse of Proposition 3.47 holding only for separable manifolds.

Proposition 3.52. *Let M^m be a separable manifold and assume that there exists a positively oriented atlas on M^m. Then the manifold M^m is orientable.*

Proof. Let $(U_j,\varphi_j)_{j\in J}$ be a positively oriented atlas. Proposition 3.51 then shows that there exist $I\subset J$ and $\theta_i\in C^\infty(M,\mathbb{R})$ such that $(U_i,\theta_i)_{i\in I}$ is a subordinated partition of unity for M^m. It is easy to see that $(U_i,\varphi_i)_{i\in I}$ is still a positively oriented atlas.

Next, we define

$$\omega_{M^m}=\sum_{i\in I}\theta_i\varphi_i^*\,dx_1\wedge\ldots\wedge dx_m\in\Omega^m(M)\,.$$

Note that because of Definition 3.50 this sum makes sense. Moreover, we obtain, using Proposition 3.24, for the pull-back of ω_{M^m} by another diffeomorphism φ_j (see also (3.74)),

$$\begin{aligned}(\varphi_j^{-1})^*\omega_{M^m}&=\sum_{i\in I}\theta_i\circ\varphi_j^{-1}\,(\varphi_i\circ\varphi_j^{-1})^*\,dx_1\wedge\ldots\wedge dx_m\\&=\sum_{i\in I}\theta_i\circ\varphi_j^{-1}\det J(\varphi_i\circ\varphi_j^{-1})\,dx_1\wedge\ldots\wedge dx_m\,.\end{aligned}$$

From a) in Definition 3.75 and the fact that $(U_i,\varphi_i)_{i\in I}$ is positively oriented, we then deduce that $\omega_{M^m}\in\Omega^m(M)$ does not vanish, implying that it is a volume form for M^m. Hence, by Definition (3.46) the orientability of M^m follows. □

Example 3.53. The constant m-form

$$\omega_{\mathbb{R}^m}:=dx_1\wedge\ldots\wedge dx_m$$

gives a volume form for $\mathbb{R}^m$.

Example 3.54. We want to show that the sphere S^n is orientable. – Let

$$\Omega=\sum_{i=1}^{n+1}(-1)^{i-1}x_i\,dx_1\wedge\ldots\wedge\widehat{dx_i}\wedge\ldots\wedge dx_{n+1}\in\Omega^n(\mathbb{R}^{n+1})\,,\qquad(3.76)$$

and let

$$\omega_{S^n}=\iota_{S^n}^*\Omega\,.$$

We want to show that ω_{S^n} is a volume form on S^n.

Let $p \in S^n$ and let $X_1, \dots, X_n$ be a basis for T_pS^n. In order to prove that ω_{S^n} is a volume form, it suffices to show that

$$\omega_{S^n}(p)\big(X_1, \dots, X_n\big) = \Omega(p)\big(d\iota_p \cdot X_1, \dots, d\iota_p \cdot X_n\big) \neq 0\,. \tag{3.77}$$

For the function $r = \sqrt{\sum_{i=1}^{n+1} |x_i|^2}$ on $\mathbb{R}^{n+1}$, we compute

$$\begin{aligned} r\,dr \wedge \Omega &= \sum_{k=1}^{n+1} x_k\,dx_k \wedge \left(\sum_{i=1}^{n+1} (-1)^{i-1} x_i\,dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_{n+1} \right) \\ &= \sum_{i=1}^{n+1} x_i^2\,dx_1 \wedge \ldots \wedge dx_{n+1} \\ &= r^2\,dx_1 \wedge \ldots \wedge dx_{n+1}\,. \end{aligned} \tag{3.78}$$

Moreover, we observe that if $\bar{X}_k = d\iota_{S^n} \cdot X_k \in \mathbb{R}^{n+1}$, for $k = 1, \dots, n$, and if $\boldsymbol{n}$ denote the unit normal to S^n at p, then $(\boldsymbol{n}, \bar{X}_1, \dots, \bar{X}_n)$ is a basis for $\mathbb{R}^{n+1}$. From (3.78), we then get that

$$r\,dr \wedge \Omega(\boldsymbol{n}, \bar{X}_1, \dots, \bar{X}_n) = r^2\,dx_1 \wedge \ldots \wedge dx_{n+1}(\boldsymbol{n}, \bar{X}_1, \dots, \bar{X}_n) \neq 0\,.$$

On the other hand, since $r|_{S^n} = 1 = \text{const}$,

$$r\,dr \cdot \bar{X}_k = r\,dr \cdot (d\iota_{S^n} \cdot X_k) = r\,d(r \circ \iota_{S^n}) \cdot X_k = 0\,.$$

Therefore, for $p \in S^n$ we arrive at, using the definition of the wedge product,

$$\begin{aligned} 0 \neq r(p)\,dr_p \wedge \Omega(p)\big(\boldsymbol{n}, \bar{X}_1, \dots, \bar{X}_n\big) &= r(p)\,dr_p \cdot \boldsymbol{n}\,\Omega(p)\big(\bar{X}_1, \dots, \bar{X}_n\big) \\ = r(p)\,\Omega(p)\big(\bar{X}_1, \dots, \bar{X}_n\big) &= \omega_{S^n}(p)\big(X_1, \dots, X_n\big)\,. \end{aligned}$$

This shows that ω_{S^n} is indeed a volume form on S^n.

Remark. For later use, we note that the volume form $\omega_{S^n} = \iota_{S^n}^* \Omega$ for S^n equals

$$\iota_{S^n}^* \big(\mathrm{int}_X dx_1 \wedge \ldots \wedge dx_{n+1}\big)\,, \tag{3.79}$$

where X now denotes the vector field $(x_1, \dots, x_n)$ on $\mathbb{R}^{n+1} \setminus \{0\}$. This is direct consequence of the formula (3.49) for the interior product.

Remark. We have shown in the previous example that the restriction $\Omega|_{S^2}$ of

$$\Omega = x_1\,dx_2 \wedge dx_3 + x_2\,dx_3 \wedge dx_1 + x_3\,dx_1 \wedge dx_2 \in \Omega^2(\mathbb{R}^3)$$

to S^2 gives a volume form ω_{S^2}. Passing to spherical coordinates on $\mathbb{R}^3$ and restricting them to the two-sphere S^2, we get, by a straightforward computation,

$$\omega_{S^2} = \cos\theta\,d\psi \wedge d\theta \in \Omega^2(S^2 \cap U)\,. \tag{3.80}$$

As expected, we see that this equals (3.64).

Consider now *codimension* one submanifolds N^n, i.e., n-dimensional submanifolds of $\mathbb{R}^{n+1}$. Generalizing Definition 3.37, we say that N^n is *oriented* if there exists $\boldsymbol{n} \in C^\infty(N^n, \mathbb{R}^{n+1})$ such that $\|\boldsymbol{n}\| = 1$ and $\boldsymbol{n}(p) \perp T_pN^n$ for all $p \in N^n$. One can easily check that this is equivalent to the existence of $\boldsymbol{n} \in C^\infty(N^n, \mathbb{R}^{n+1})$ such that $\boldsymbol{n}(x) \notin T_pN^n$. On the other hand, we already know from Example 2.20 that a submanifold can also be interpreted as an abstract manifold. Hence, there is a second notion of orientation for N^n given by Definition 3.46. The following proposition, however, shows that the two notions of orientation for a codimension one submanifold agree.

Proposition 3.55. *Let N^n be a codimension one submanifold of $\mathbb{R}^{n+1}$. Then N^n is orientable (in the sense of an abstract manifold) if and only if there exists $\boldsymbol{n} \in C^\infty(N^n, \mathbb{R}^{n+1})$ such that $\boldsymbol{n}(x) \notin T_pN^n$.*

For the proof of this proposition, we will need the following lemma which is motivated by (3.79).

Lemma 3.56. *Let N^n be a codimension one submanifold of $\mathbb{R}^{n+1}$ and assume that there exists $\boldsymbol{n} \in C^\infty(N^n, \mathbb{R}^{n+1})$ such that $\boldsymbol{n}(x) \notin T_pN^n$. Moreover, let $\iota_{N^n} : N^n \hookrightarrow \mathbb{R}^{n+1}$ denote the canonical inclusion. Then*

$$\omega_{N^n} = \iota^*_{N^n}\big(int_{\boldsymbol{n}}\, dx_1 \wedge \ldots \wedge dx_{n+1}\big) \in \Omega^n(N^n)\,, \tag{3.81}$$

gives a volume form for N^n.

Proof. Let $p \in N^n$ and let $X_1, \ldots, X_n$ be a basis of T_pN^n. Moreover, for $k = 1, \ldots, n$, we define[2]

$$\bar{X}_k := d\iota_{N^n} \cdot X_k\,. \tag{3.82}$$

By assumption, it is then clear that $(\boldsymbol{n}, \bar{X}_1, \ldots, \bar{X}_n)$ forms a basis of $\mathbb{R}^{n+1}$.

Using Definition 3.23 of the pull-back, we deduce

$$\iota^*_{N^n}\big(\mathrm{int}_{\boldsymbol{n}}\, dx_1 \wedge \ldots \wedge dx_{n+1}\big)(X_1, \ldots, X_n) = \mathrm{int}_{\boldsymbol{n}}\, dx_1 \wedge \ldots \wedge dx_{n+1}(\bar{X}_1, \ldots, \bar{X}_n)\,.$$

And by Definition 3.28 of the interior product, it follows

$$\begin{aligned} \mathrm{int}_{\boldsymbol{n}}\, dx_1 \wedge \ldots \wedge dx_{n+1}(\bar{X}_1, \ldots, \bar{X}_n) &= dx_1 \wedge \ldots \wedge dx_{n+1}(\boldsymbol{n}, \bar{X}_1, \ldots, \bar{X}_n) \\ &= \det(\boldsymbol{n}, \bar{X}_1, \ldots, \bar{X}_n). \end{aligned} \tag{3.83}$$

Since $(\boldsymbol{n}, \bar{X}_1, \ldots, \bar{X}_n)$ forms a basis of $\mathbb{R}^{n+1}$, it is clear that ω_{N^n} never vanishes and is thus a volume form for N^n. □

Now, we are ready to give a proof of Proposition 3.55.

[2] Note that only due to the inclusion map ι_{N^n} the manifold N^n can be interpreted as submanifold. Hence, starting with “paths” $X_1, \ldots, X_n$ (see Definition 2.35), we obtain vectors $\bar{X}_1, \ldots, \bar{X}_n$ in $\mathbb{R}^{n+1}$.

Proof. Assume that there exists $\boldsymbol{n} \in C^\infty(N^n, \mathbb{R}^{n+1})$ such that $\boldsymbol{n}(x) \notin T_pN^n$. Because of Lemma 3.56, the n-form ω_{N^n} defines a volume form for N^n, and hence N^n is orientable.

Conversely, assume that N^n is orientable for the volume form $\omega_{N^n} \in \Omega^n(N^n)$. – Let $p \in N^n$ and $\bar{X}_1, \ldots, \bar{X}_n \in \mathbb{R}^{n+1}$ be a basis of the hyperplane tangent to N^n at p. Moreover, let $X_1, \ldots, X_n$ denote the corresponding elements of T_pN^n (see (3.82)). Then we define $\bar{\omega}_p \in \bigwedge^n(\mathbb{R}^{n+1})$ by

$$\bar{\omega}_p(\bar{X}_1, \ldots, \bar{X}_n) = (\omega_{N^n})_p\big(X_1, \ldots, X_n\big)\,,$$

and

$$\bar{\omega}_p(\bar{Y}_1, \ldots, \bar{Y}_n) = 0\,,$$

if there exists $1 \le j \le n$ such that $\bar{Y}_j \in (T_pN^n)^\perp$. This defines $\bar{\omega}_p$ in a unique way. As a direct consequence, since

$$\big\{dx_1(p) \wedge \ldots \wedge \widehat{dx_k}(p) \wedge \ldots \wedge dx_{n+1}(p)\big\}_{k=1,\ldots,n+1}$$

is basis of $\bigwedge^n(\mathbb{R}^{n+1})$, there exist unique numbers $\alpha_1(p), \ldots, \alpha_{n+1}(p) \in \mathbb{R}$ such that

$$\bar{\omega}_p = \sum_{k=1}^{n+1} \alpha_k(p)\, dx_1(p) \wedge \ldots \wedge \widehat{dx_k}(p) \wedge \ldots \wedge dx_{n+1}(p)\,. \tag{3.84}$$

Repeating this procedure for all $p \in N^n$, we get the existence of C^∞-functions $\alpha_1, \ldots, \alpha_{n+1} : N^n \longrightarrow \mathbb{R}$ such that (the fact that $\alpha_1, \ldots, \alpha_{n+1}$ are C^∞-functions comes from the assumption $\omega_{N^n} \in \Omega^n(N^n)$)

$$\bar{\omega} = \sum_{k=1}^{n+1} \alpha_k\, dx_1 \wedge \ldots \wedge \widehat{dx_k} \wedge \ldots \wedge dx_{n+1} \in \Omega^n(\mathbb{R}^{n+1})\,. \tag{3.85}$$

Next, we define

$$\boldsymbol{n} = \sum_{k=1}^{n+1} (-1)^{1-k} \alpha_k\, e_k \in C^\infty(N^n, \mathbb{R}^{n+1})\,, \tag{3.86}$$

where $\{e_k\}_{k=1,\ldots,n+1}$ denotes the canonical basis of $\mathbb{R}^{n+1}$. If we can show that $\boldsymbol{n}(p) \notin T_pN^n$ for all $p \in N^n$, we are done.

First, (3.85) and (3.86) show that

$$\begin{aligned}\mathrm{int}_{\boldsymbol{n}} dx_1 \wedge \ldots \wedge dx_{n+1}(\bar{X}_1, \ldots, \bar{X}_n) &= \bar{\omega}_p(\bar{X}_1, \ldots, \bar{X}_n)\\ &= (\omega_{N^n})_p\big(X_1, \ldots, X_n\big)\,.\end{aligned}$$

Since ω_{N^n} never vanishes by assumption, we deduce, using (3.83),

$$\det(\boldsymbol{n}, \bar{X}_1, \ldots, \bar{X}_n) \neq 0\,.$$

Thus $(\boldsymbol{n}, \bar{X}_1, \ldots, \bar{X}_n)$ is a basis of $\mathbb{R}^{n+1}$. Moreover, since by assumption $\bar{X}_1, \ldots, \bar{X}_n \in \mathbb{R}^{n+1}$ forms a basis of the hyperplane tangent to N^n at p, it follows that $\boldsymbol{n}(p) \notin T_pN^n$. This completes the proof. $\square$

4 Integration on Manifolds

4.1 Integration of n-forms in $\mathbb{R}^n$

First, we recall an important result for the integration of functions on $\mathbb{R}^n$.

Theorem 4.1 (Change of Variable Formula). *Let $U \subset \mathbb{R}^n$ open and ϕ a C^k-diffeomorphism, $k \geq 1$, from U into $\phi(U)$. Then if $f \in L^1(\phi(U))$ it follows that $(f \circ \phi)|det\, J\phi| \in L^1(U)$ and moreover the following formula holds:*

$$\int_U \big(f \circ \phi\big)(x)\big|det\, J_x\phi\big|\, d\mathcal{L}(x) = \int_{\phi(U)} f(y)\, d\mathcal{L}(y)\,. \tag{4.1}$$

We define now the integral of differential n-forms on $\mathbb{R}^n$. As shown in Proposition 3.7, we have that dim $\bigwedge^n \mathbb{R}^n = 1$.

Definition 4.2. *Let $\alpha = f\, dx_1 \wedge \ldots \wedge dx_n$ be a differential n-form on $U \subset \mathbb{R}^n$ open. Moreover, assume that $f \in L^1(U)$. Then the **integral** of α on U is defined to be the number*

$$\int_U \alpha = \int_U f(x)\, d\mathcal{L}^n(x)\,, \tag{4.2}$$

where $d\mathcal{L}^n$ is the Lebesgue measure on $\mathbb{R}^n$. We write $\alpha \in \Omega^n_{L^1}(U)$.

Proposition 4.3. *Let $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a C^k-map, $k \geq 1$, and $U \subset \mathbb{R}^n$ open such that ϕ is a diffeomorphism from U into $\phi(U)$. Moreover, assume that ϕ is positive, i.e., $det\, J_x\phi > 0$, for all $x \in U$. Then for $\alpha \in \Omega^n_{L^1}(\phi(U))$ the following formula holds:*

$$\int_{\phi(U)} \alpha = \int_U \phi^*\alpha\,. \tag{4.3}$$

Proof. Let $\alpha = f\, dy_1 \wedge \ldots \wedge dy_n \in \Omega^n_{L^1}(\phi(U))$. Since the pull-back of α by ϕ is just multiplication by the determinant of the Jacobian of ϕ (see Remark 3.44), we obtain

$$\phi^*\alpha = f \circ \phi \det J\phi\; dx_1 \wedge \ldots \wedge dx_n\,. \tag{4.4}$$

By the definition of the integral, this gives

$$\int_U \phi^*\alpha = \int_U \big(f \circ \phi\big)(x) \det J_x\phi \, d\mathcal{L}^n(x)\,.$$

Since $\det J\phi > 0$ by assumption, we can apply the change of variable formula (4.1):

$$\int_U \big(f \circ \phi\big)(x) \det J_x\phi \, d\mathcal{L}^n(x) = \int_{\phi(U)} f(y)\, d\mathcal{L}^n(y) = \int_{\phi(U)} \alpha\,.$$

□

Remark. If the diffeomorphism ϕ is assumed to be *negative*, i.e., $\det J\phi < 0$, then we see that

$$\int_{\phi(U)} \alpha = - \int_U \phi^*\alpha\,.$$

Instead of saying that diffeomorphisms are positive or negative, they are often called *orientation preserving*, respectively, *orientation reversing* diffeomorphisms. Assuming that the n-form α in (4.4) is replaced by a volume form for $\mathbb{R}^n$, this terminology can be easily understood (see also the remark after Definition 3.43).

4.2 Integration of n-forms on a n-dimensional Manifold

In the whole chapter, we always assume that M^m denotes a m-dimensional C^∞-differentiable (separable) manifold.

Definition 4.4. *Let M^m be an oriented manifold and let $\omega = f\, dx_1 \wedge \ldots \wedge dx_m$ be a compactly supported m-form on M^m. Moreover, let $(U_i, \varphi_i)_{i\in I}$ be a positively oriented atlas and $(U_i, \theta_i)_{i\in I}$ a partition of unity subordinated to it. Assume also that $\theta_i\, f \circ \varphi_i^{-1} \in L^1\big(\varphi_i(U_i)\big)$, for all $i \in I$. Then, we define the* ***integral*** *of ω on M^m by the number*

$$\int_{M^m} \omega = \sum_{i\in I} \int_{\varphi_i(U_i)} (\varphi_i^{-1})^* \,(\theta_i\, \omega)\,. \tag{4.5}$$

And we write $\omega \in \Omega^m_{L^1}(M)$.

Remark. From the definition, it is easy to see that the map

$$\omega \longmapsto \int_{M^m} \omega$$

is linear. Note also that $(\varphi_i^{-1})^*\,(\theta_i\, \omega) \in \Omega^m_{L^1}(\varphi_i(U_i))$.

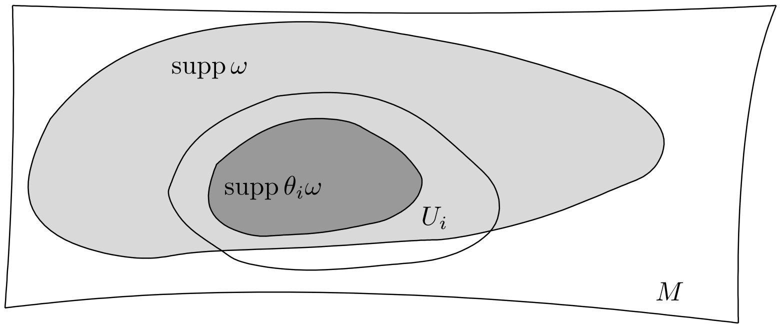

Fig. 4.1. Integration on manifolds.

In order to show that the last definition makes sense, we claim in a first step that for only finitely many $i \in I$ the expression

$$\int_{\varphi_i(U_i)} (\varphi_i^{-1})^* \, (\theta_i \, \omega)$$

is non-zero. – Let $K \subset M^m$ denote the compact support of $\omega \in \Omega^m_{L^1}(M)$. Since $(U_i)_{i\in I}$ is assumed to be a locally finite covering, there exists by Definition 3.49 for every $p \in M^m$ an open subset U_p of M^m containing p such that it intersects only finitely many U_i, $i \in I$. On the other hand, since $K \subset \bigcup_{p\in K} U_p$ compact, we can extract a finite covering $U_{p_1}, \ldots, U_{p_k}$. Clearly, each of these open sets then intersects only finitely many U_i, $i \in I$. Hence, the support K of ω is covered by only finitely many of the U_i. Recalling that $\operatorname{supp}\theta_i \subset U_i$, for all $i \in I$ (see Definition 3.50), the claim follows.

In a second step, we show that the definition of the integral of ω on M^m is independent of the choice of the *positively* oriented atlas and the choice of the partition of unity subordinated to the atlas. – Let $(V_j, \psi_j)_{j\in J}$ be another positively oriented atlas with the same orientation given by $(U_i, \varphi_i)_{i\in I}$, i.e., on $\varphi_i(U_i \cap V_j)$, we have that $\det J(\psi_j \circ \varphi_i^{-1}) > 0$. Moreover, let $(V_j, \eta_j)_{j\in J}$ be a partition of unity subordinated to $(V_j)_{j\in J}$. Then, we can write

$$\int_{\varphi_i(U_i)} (\varphi_i^{-1})^*(\theta_i\omega) = \sum_{j\in J} \int_{\varphi_i(U_i\cap V_j)} (\varphi_i^{-1})^*(\theta_i\eta_j\omega)\,,$$

since $1 = \sum_{j\in J} \eta_j$ on M^m by definition. Next, we apply the formula (4.3) to the differential m-form $(\varphi_i^{-1})^*(\theta_i\eta_j\omega)$ on $\mathbb{R}^m$ and the diffeomorphism

$$\varphi_i \circ \psi_j^{-1} : \psi_j(U_i \cap V_j) \subset \mathbb{R}^m \longrightarrow \varphi_i(U_i \cap V_j) \subset \mathbb{R}^m$$

to get

$$\begin{aligned}\int_{\varphi_i(U_i\cap V_j)}(\varphi_i^{-1})^*(\theta_i\eta_j\omega) &= \int_{\psi_j\circ\varphi_i^{-1}\left(\varphi_i(U_i\cap V_j)\right)}(\varphi_i\circ\psi_j^{-1})^*\big((\varphi_i^{-1})^*(\theta_i\eta_j\omega)\big)\\ &= \int_{\psi_j(U_i\cap V_j)}(\psi_j^{-1})^*(\theta_i\eta_j\omega)\,. \qquad (4.6)\end{aligned}$$

And for the integral of ω on M^m we thus have, using again (3.75),

$$\begin{aligned}\sum_{i\in I}\int_{\varphi_i(U_i)}(\varphi_i^{-1})^*\,(\theta_i\,\omega) &= \sum_{j\in J}\sum_{i\in I}\int_{\psi_j(U_i\cap V_j)}(\psi_j^{-1})^*(\theta_i\eta_j\omega)\\ &= \sum_{j\in J}\int_{\psi_j(V_j)}(\psi_j^{-1})^*(\eta_j\omega)\,.\end{aligned}$$

In summary, we have shown that the integral on a manifold is well-defined. – Note that without fixing an orientation for M^m, the definition of the integral of ω is *not* well-posed. More precisely, looking at (4.6) the sign of the integral changes with the change of orientation.

Remark. Let M^m be an oriented manifold and $K\subset M^m$ compact. Then for $\omega\in\Omega^m_{L^1}(M)$ not necessarily compactly supported, we write

$$\int_K\omega = \int_{M^m}\chi_K\omega\,, \qquad (4.7)$$

where χ_K is the characteristic function of K. Clearly, $\chi_K\omega$ is compactly supported and we can use Definition 4.4 for the right-hand side of the last equation.

Next, we show that Proposition 4.3 holds also for integration on manifolds.

Proposition 4.5. *Let M^m and N^m be two oriented manifolds of same dimension m. Moreover, let $\phi : M^m \longrightarrow N^m$ be a positive C^∞-diffeomorphism, i.e., the image by ϕ of a positive oriented atlas on M^m is a positive oriented atlas on N^m. Then for $\omega\in\Omega^m_{L^1}(M)$ the following formula holds:*

$$\int_{N^m}\omega = \int_{M^m}\phi^*\omega\,. \qquad (4.8)$$

Proof. Let $(U_i,\varphi_i)_{i\in I}$ be a positively oriented atlas of M^m and let $(U_i,\theta_i)_{i\in I}$ be a partition of unity subordinated to it. Then we have

$$\int_{M^m}\phi^*\omega = \sum_{i\in I}\int_{\varphi_i(U_i)}(\varphi_i^{-1})^*\big(\theta_i(\phi^*\omega)\big)\,.$$

Using the properties of the pull-back summarized in Proposition 3.24, the right-hand side of the last equation can also be written as

$$\begin{aligned}
&\sum_{i\in I}\int_{\varphi_i(U_i)}(\varphi_i^{-1})^*\phi^*\big((\theta_i\circ\phi^{-1})\,\omega\big)\\
&=\sum_{i\in I}\int_{\varphi_i(U_i)}(\phi\circ\varphi_i^{-1})^*\big((\theta_i\circ\phi^{-1})\,\omega\big)\\
&=\sum_{i\in I}\int_{\varphi_i(U_i)}\big((\varphi_i\circ\phi^{-1})^{-1}\big)^*\big((\theta_i\circ\phi^{-1})\,\omega\big)\,.
\end{aligned}$$

Hence, we arrive at

$$\int_{M^m}\phi^*\omega=\sum_{i\in I}\int_{\varphi_i\circ\phi^{-1}\left(\phi(U_i)\right)}\big((\varphi_i\circ\phi^{-1})^{-1}\big)^*\big((\theta_i\circ\phi^{-1})\,\omega\big)\,.$$

It is to check that $\big(\phi(U_i),\varphi_i\circ\phi^{-1}\big)_{i\in I}$ is an atlas for N^m. It is also positively oriented, since by assumption ϕ is a positive diffeomorphism. Moreover, the family $\big(\phi(U_i),\eta_i\big)_{i\in I}$ with $\eta_i:=\theta_i\circ\phi^{-1}\in C^\infty(N^m,\mathbb{R})$ defines a subordinated partition of unity. Using Definition 4.4, we then conclude that

$$\int_{M^m}\phi^*\omega=\int_{N^m}\omega\,.$$

□

We generalize the last proposition in the sense that $\phi: M^m \longrightarrow N^m$ needs *not* necessarily to be a diffeomorphism. – Let M^m and N^m be two compact oriented manifolds of same dimension m and let ω be a m-form on N^m. Moreover, let ϕ be an *arbitrary* C^1-map from M^m to N^m. The next theorem, however, then shows that $\int_{N^m}\omega$ and $\int_{M^m}\phi^*\omega$ are related.

Theorem 4.6 (Topological degree). *Let M^m and N^m be two compact oriented manifolds of same dimension m and let $\phi\in C^1(M^m,N^m)$. Then, there exists an integer $deg\,\phi$, called* ***topological degree*** *of ϕ, such that for any $\omega\in\Omega^m_{L^1}(N)$, we have*

$$\int_{M^m}\phi^*\omega = deg\,\phi\int_{N^m}\omega\,. \tag{4.9}$$

Remark. Roughly speaking the topological degree of a map ϕ can be interpreted as the number of times $\phi(M^m)$ covers the manifold N^m.

Example 4.7. Consider the case $M^m=N^m=S^1$ (see Fig. 4.2).

Definition 4.8. *Let ϕ and $\tilde\phi\in C^1(M^m,N^m)$. We say that ϕ is* ***homotopically equivalent*** *to $\tilde\phi$, if there exists*

$$H\in C^0([0,1]\times M^m,N^m)$$

such that $\phi(\cdot)=H(0,\cdot)$ and $\tilde\phi(\cdot)=H(1,\cdot)$.

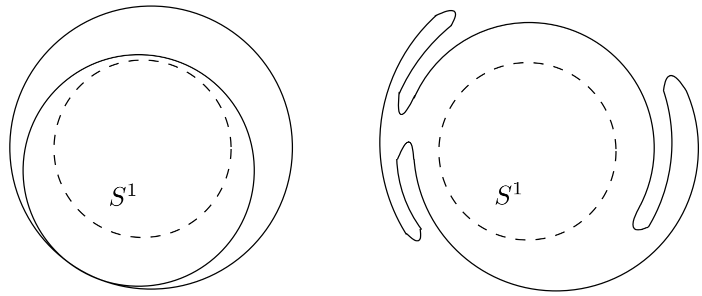

Fig. 4.2. Topological degree of two different maps ϕ and $\tilde{\phi}$.

It is easy to check that homotopically equivalent is an equivalence relation. Moreover, one can show that the topological degree of a map only depends on its homotopy class (see []). In other words, the topological degree depends only on the C^0-structure of the map. Thus differential forms, being more than only C^0-objects, are used to determine C^0-objects.

Conversely, assume that $\deg \phi = \deg \tilde{\phi}$. How are the maps ϕ and $\tilde{\phi}$ related?

Theorem 4.9. *Let ϕ and $\tilde{\phi} \in C^1(M^m, N^m)$, where $N^m = S^m$ and M^m oriented and compact. Then $\deg \phi = \deg \tilde{\phi}$ implies that ϕ and $\tilde{\phi}$ are homotopically equivalent.*

4.2.1 Integration and Volume of a Surface in $\mathbb{R}^3$

Let Σ be a surface in $\mathbb{R}^3$, i.e., a 2-dimensional C^1-submanifold of $\mathbb{R}^3$. Moreover, let (U, φ) denote a local chart for Σ. Then, we consider a cube $C' \subset \varphi(U \cap \Sigma) \subset \mathbb{R}^2$ and divide it in little pieces of side length $1/n$, where $n \in \mathbb{N}$. More precisely, assuming that $C' = [0,1] \times [0,1]$, we write

$$C' = \sum_{k,l=0}^{n-1} \left[\frac{k}{n}, \frac{k+1}{n}\right] \times \left[\frac{l}{n}, \frac{l+1}{n}\right] = \sum_{k,l=0}^{n-1} C'_{k,l}\,.$$

The preimages of the little cubes $C'_{k,l}$ are denoted by $U_{k,l} := \varphi^{-1}(C'_{k,l})$, and the equation $\varphi(p_{k,l}) = \left(\frac{k}{n}, \frac{l}{n}\right)$ defines the points $p_{k,l} \in U_{k,l}$.

We want study how to define the volume (area) of the surface Σ with the help of this grid. This is best done by some elementary geometric considerations. - The inverse map of the C^1-diffeomorphism $\varphi : U \cap \Sigma \subset \mathbb{R}^3 \longrightarrow \varphi(U \cap \Sigma) \subset \mathbb{R}^2$, U open in $\mathbb{R}^3$, is denoted by

$$\varphi^{-1} : \varphi(U \cap \Sigma) \longrightarrow U \cap \Sigma\,.$$

Recall that the two tangent vectors[1]

[1] In a different notation the generating tangent vectors for $T_{p_{k,l}}\Sigma$ can also be written as $\frac{\partial}{\partial \varphi_1}(p_{k,l})$ and $\frac{\partial}{\partial \varphi_2}(p_{k,l})$.

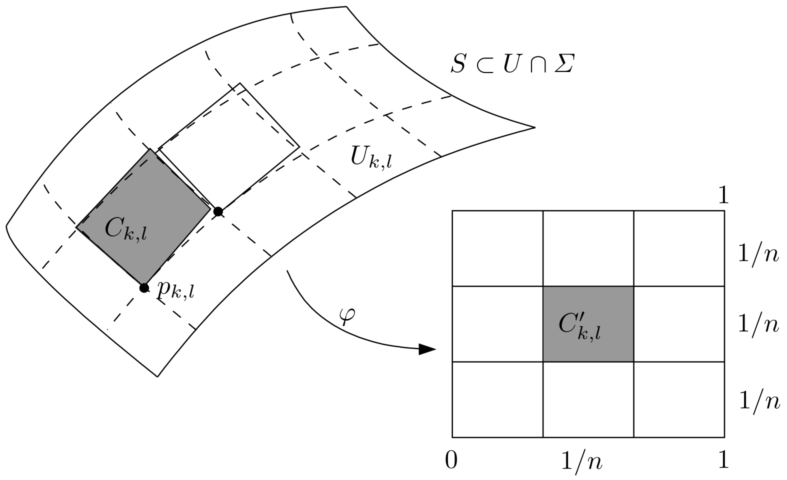

Fig. 4.3. Putting a grid on a surface.

$$\left\{ \frac{\partial \varphi^{-1}}{\partial x_i} \big(\varphi(p_{k,l}) \big) \right\}_{i=1,2} ,$$

where (x_1, x_2) denote the canonical coordinates on $\mathbb{R}^2$, generate the tangent space $T_{p_{k,l}}\Sigma$ of Σ at $p_{k,l}$. Then, for all $0 \leq k, l \leq n-1$, the parallelograms given by the vectors

$$\frac{1}{n} \frac{\partial \varphi^{-1}}{\partial x_1} \big(\varphi(p_{k,l}) \big) \qquad \text{and} \qquad \frac{1}{n} \frac{\partial \varphi^{-1}}{\partial x_2} \big(\varphi(p_{k,l}) \big)$$

can be interpreted as linear approximation of the little pieces $U_{k,l}$ of the surface. These parallelograms, denoted by $C_{k,l}$, have the well-known volume

$$\operatorname{Vol} C_{k,l} = \frac{1}{n^2} \left\| \frac{\partial \varphi^{-1}}{\partial x_1} \big(\varphi(p_{k,l}) \big) \times \frac{\partial \varphi^{-1}}{\partial x_2} \big(\varphi(p_{k,l}) \big) \right\| . \tag{4.10}$$

Hence, the volume of $S := \bigcup_{k,l=0}^{n-1} U_{k,l} \subset U \cap \Sigma$ can be approximatively defined by

$$\sum_{k,l=0}^{n-1} \operatorname{Vol} C_{k,l} = \sum_{k,l=0}^{n-1} \frac{1}{n^2} \left\| \frac{\partial \varphi^{-1}}{\partial x_1} \big(\varphi(p_{k,l}) \big) \times \frac{\partial \varphi^{-1}}{\partial x_2} \big(\varphi(p_{k,l}) \big) \right\| . \tag{4.11}$$

Intuitively, it is clear that the approximation becomes better for large n, and we thus define

$$\operatorname{Vol} S := \lim_{n \longrightarrow \infty} \sum_{k,l=0}^{n-1} \operatorname{Vol} C_{k,l}^n . \tag{4.12}$$

Here we write $C_{k,l}^n$ for the little parallelograms to emphasize that they depend on n.

With the area form dA_Σ defined in (3.54), the volume of S can be written differently. – First, note that the unit normal vector field to Σ at the point $p_{k,l}$ (see Definition 3.37), reads as

$$\boldsymbol{n}(p_{k,l}) = \frac{\frac{\partial\varphi^{-1}}{\partial x_1}\big(\varphi(p_{k,l})\big) \times \frac{\partial\varphi^{-1}}{\partial x_2}\big(\varphi(p_{k,l})\big)}{\left\|\frac{\partial\varphi^{-1}}{\partial x_1}\big(\varphi(p_{k,l})\big) \times \frac{\partial\varphi^{-1}}{\partial x_2}\big(\varphi(p_{k,l})\big)\right\|}\,.$$

This implies for the area form, expressed in the local chart (U,φ),

$$dA_\Sigma = \left\|\frac{\partial\varphi^{-1}}{\partial x_1} \times \frac{\partial\varphi^{-1}}{\partial x_2}\right\| d\varphi_1 \wedge d\varphi_2 \in \Omega_0^2(U\cap\Sigma)\,. \tag{4.13}$$

It follows that

$$\begin{aligned}(dA_\Sigma)_{p_{k,l}}\left(\frac{1}{n}\frac{\partial}{\partial\varphi_1}(p_{k,l}), \frac{1}{n}\frac{\partial}{\partial\varphi_2}(p_{k,l})\right) &= \frac{1}{n^2}\left\|\frac{\partial\varphi^{-1}}{\partial x_1}\big(\varphi(p_{k,l})\big) \times \frac{\partial\varphi^{-1}}{\partial x_2}\big(\varphi(p_{k,l})\big)\right\| \\ &= \operatorname{Vol} C_{k,l}\,.\end{aligned}$$

Hence, the volume of the parallelogram $C_{k,l}$ can also be expressed in terms of the area form $(dA_\Sigma)_{p_{k,l}}$. Thus for the approximation (4.11) of the volume of S, we get

$$\sum_{k,l=0}^{n-1}\operatorname{Vol} C_{k,l} = \sum_{k,l=0}^{n-1}\frac{1}{n^2}\,(dA_\Sigma)_{p_{k,l}}\left(\frac{\partial}{\partial\varphi_1}(p_{k,l}), \frac{\partial}{\partial\varphi_2}(p_{k,l})\right)\,,$$

or equivalently

$$\sum_{k,l=0}^{n-1}\operatorname{Vol} C_{k,l} = \sum_{k,l=0}^{n-1}\int_{C'_{k,l}}\left\|\frac{\partial\varphi^{-1}}{\partial x_1}\big(\varphi(p_{k,l})\big) \times \frac{\partial\varphi^{-1}}{\partial x_2}\big(\varphi(p_{k,l})\big)\right\| dx_1 dx_2\,. \tag{4.14}$$

In a next step, we show that integration of the area form (4.13) over S indeed leads to the volume of S defined in (4.12). By definition of the integral, we have that

$$\int_{U_{k,l}} dA_\Sigma = \int_{\varphi(U_{k,l})}\left\|\frac{\partial\varphi^{-1}}{\partial x_1}(x) \times \frac{\partial\varphi^{-1}}{\partial x_2}(x)\right\| dx_1 dx_2\,.$$

From the continuity of the area form, it then follows

$$\begin{aligned}&\left|\int_{U_{k,l}} dA_\Sigma - \operatorname{Vol} C_{k,l}\right| \\ &= \left|\int_{U_{k,l}} dA_\Sigma - \int_{C'_{k,l}}\left\|\frac{\partial\varphi^{-1}}{\partial x_1}\big(\varphi(p_{k,l})\big) \times \frac{\partial\varphi^{-1}}{\partial x_2}\big(\varphi(p_{k,l})\big)\right\|\right| \\ &\le \varepsilon_{k,l}^n \cdot \operatorname{Vol} C'_{k,l}\,,\end{aligned} \tag{4.15}$$

where

$$\varepsilon_{k,l}^n := \sup_{x\in C'_{k,l}} \left| \left\| \frac{\partial\varphi^{-1}}{\partial x_1}(x) \times \frac{\partial\varphi^{-1}}{\partial x_2}(x) \right\| - \left\| \frac{\partial\varphi^{-1}}{\partial x_1}\big(\varphi(p_{k,l})\big) \times \frac{\partial\varphi^{-1}}{\partial x_2}\big(\varphi(p_{k,l})\big) \right\| \right| .$$

Taking the sum over k and l in (4.15) and using (4.14), we arrive at

$$\left| \int_S dA_\Sigma - \sum_{k,l=0}^{n-1} \operatorname{Vol} C_{k,l} \right| \leq \max_{0\leq k,l\leq n-1} \varepsilon_{k,l}^n \cdot \sum_{k,l=0}^{n-1} \operatorname{Vol} C'_{k,l} . \tag{4.16}$$

Since $\varepsilon_{k,l}^n$ goes to zero for $n \longrightarrow \infty$, taking the limit in (4.16) implies

$$\int_S dA_\Sigma = \operatorname{Vol} S , \tag{4.17}$$

where we used (4.12). So integration of the area form indeed gives the volume (area) of the surface.

Example 4.10 (Volume of S^2). We want to calculate the integral of the area form dA_{S^2} expressed in spherical coordinates (see (3.64)). Since the spherical coordinates cover up to a set of measure zero the whole two-sphere S^2, it is possible to use only a single chart for the integration. Hence, we have

$$\begin{aligned} \int_{S^2} dA_{S^2} &= \int_{-\pi}^{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos\theta \, d\psi d\theta \\ &= 2\pi \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos\theta \, d\theta = 4\pi . \end{aligned} \tag{4.18}$$

Considering only the northern half-sphere S^2_+ and the corresponding volume form $dA_{S^2_+}$ (see (3.66)), the same result for the volume of the sphere can be obtained (by a more precise calculation).

4.3 Stokes' Theorem

Let $f \in C^1([0,1],\mathbb{R})$. Then, from the well-known "Fundamental Theorem of Calculus" or integration by parts formula for functions, it follows that

$$\int_0^1 f'(t)\, dt = f(1) - f(0) . \tag{4.19}$$

More generally, let $df \in \Omega_1^1([0,1])$. Then by Definition 4.2, we have

$$\int_{[0,1]} df = \int_{[0,1]} f'(t)\, dt .$$

With (9.92) and Example 3.33, this can also be written as

$$\int_{[0,1]} df = f(1) - f(0) = \int_{\partial[0,1]} f|_{\partial[0,1]} = \int_{\partial[0,1]} \iota^*_{\partial[0,1]} f\,,$$

where $\partial[0,1] = \{0_-\} \cup \{1_+\}$ denotes the boundary of $[0,1] \subset \mathbb{R}$. It will turn out that the last equation is a special case of Stokes' theorem. Hence, this important theorem can be interpreted as an integration by parts formula for forms on manifolds.

Before stating Stokes' theorem, we have to study the boundary $\partial\Omega = \bar{\Omega}\backslash\Omega$ of an open subset of a manifold in a detailed manner.

Definition 4.11. *An **oriented domain** Ω of an oriented manifold M^m is an open subset of M^m such that, for every $p \in \partial\Omega = \bar{\Omega} \setminus \Omega$, there exists a local chart (U, φ) of M^m, $p \in U$ open, with*

(i) $\varphi(p) = 0$, *i.e., the chart is centered at* p;
(ii) $\varphi(\Omega \cap U) = \big(\mathbb{R}_+ \setminus \{0\}\big) \times \mathbb{R}^{m-1} \cap \varphi(U)$;
(iii) the map φ *is a positive diffeomorphism from* U *into* $\varphi(U)$.

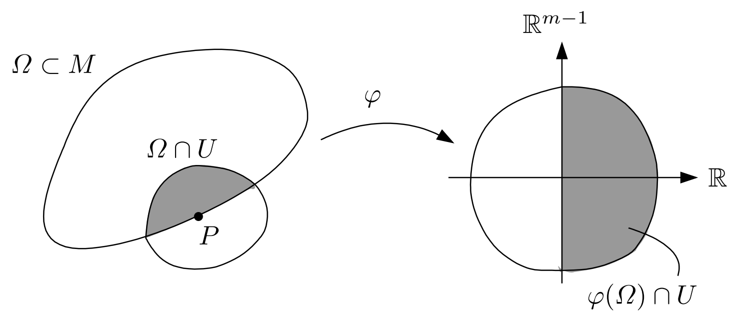

Fig. 4.4. Domain of a manifold.

Proposition 4.12. *Let Ω be an oriented domain of an orientable manifold M^m. Then $\partial\Omega$ is an oriented $(m-1)$-dimensional submanifold of M^m.*

Proof. Let $p \in \partial\Omega = \bar{\Omega} \setminus \Omega$ and let (U, φ) denote the local chart given by Definition 4.11. Since φ is a bijection, we get

$$\varphi\big((\bar{\Omega} \setminus \Omega) \cap U\big) = \varphi(\bar{\Omega} \cap U) \setminus \varphi(\Omega \cap U)\,,$$

and since φ is a homeomorphism,

$$\varphi(\bar{\Omega} \cap U) = \overline{\varphi(\Omega \cap U)} \cap \varphi(U)\,.$$

These two observations then imply that

$$\begin{aligned}\varphi(\partial\Omega \cap U) &= \left(\overline{\varphi(\Omega \cap U)} \cap \varphi(U)\right) \setminus \varphi(\Omega \cap U) \\ &= \mathbb{R}_+ \times \mathbb{R}^{m-1} \cap \varphi(U) \setminus \left(\mathbb{R}_+ \setminus \{0\}\right) \times \mathbb{R}^{m-1} \cap \varphi(U) \\ &= \{0\} \times \mathbb{R}^{m-1} \cap \varphi(U)\,,\end{aligned}$$

showing that $\partial\Omega$ is a $(m-1)$-dimensional submanifold of M^m. – Note that submanifolds of manifolds are also manifolds as in the case of submanifolds of Euclidean space.

In order to show the orientability of the manifold $\partial\Omega$, it suffices by Proposition 3.52 to construct a positively oriented atlas on $\partial\Omega$. Then we are done.

Let (φ, U) and (ψ, V) be two local charts about $p \in \partial\Omega$ given by Definition 4.11. We then denote the transition function by

$$\psi\circ\varphi^{-1} = (f_1, \bar\psi\circ\bar\varphi^{-1}) : \varphi(U\cap V) \subset \mathbb{R}_+\times\mathbb{R}^{m-1} \longrightarrow \psi(U\cap V) \subset \mathbb{R}_+\times\mathbb{R}^{m-1}\,,$$

where $\bar\varphi := \varphi|_{\partial\Omega}$ and $\bar\psi := \psi|_{\partial\Omega}$. Since

$$\varphi(U \cap V \cap \partial\Omega) \subset \{0\} \times \mathbb{R}^{m-1} \cap \varphi(U \cap V)\,,$$

and

$$\psi(U \cap V \cap \partial\Omega) \subset \{0\} \times \mathbb{R}^{m-1} \cap \psi(U \cap V)\,,$$

we have for all $x \in \{0\} \times \mathbb{R}^{m-1} \cap \varphi(U \cap V)$ that $f_1(x) = 0$. This implies that

$$\frac{\partial f_1}{\partial x_i}(x) = 0\,, \qquad i = 2, \ldots, m\,.$$

Moreover, one easily checks that $\lambda_1 := \frac{\partial f_1}{\partial x_1}(0) \geq 0$ at the point $\varphi(p) = \psi(p) = 0$. Thus we can write for the Jacobian at the point $0 \in \mathbb{R}^m$ of the transition function

$$J_0(\psi \circ \varphi^{-1}) = \begin{pmatrix} \lambda_1 & 0 \\ * & J_0(\bar\psi \circ \bar\varphi^{-1}) \end{pmatrix}.$$

It follows that

$$\det J_0(\psi \circ \varphi^{-1}) = \lambda_1 \det J_0(\bar\psi \circ \bar\varphi^{-1})\,.$$

By assumption $\det J_0(\psi \circ \varphi^{-1}) > 0$ showing that $\det J_0(\bar\psi \circ \bar\varphi^{-1}) > 0$. This proves the proposition. □

At this stage, we are ready to formulate Stokes' theorem.

Theorem 4.13 (Stokes' Theorem). *Let M^m be an oriented manifold and let Ω be an oriented domain of M^m such that $\bar\Omega$ is compact. Then for every $\omega \in \Omega_1^{m-1}(M)$ the following formula holds:*

$$\int_\Omega d\omega = \int_{\partial\Omega} \iota^*_{\partial\Omega}\omega\,, \tag{4.20}$$

where $\iota : \partial\Omega \hookrightarrow M^m$ denotes the canonical inclusion.

Proof. To every $p \in \bar{\Omega}$ we assign a local chart (U_p, φ_p) such that φ_p is a positive diffeomorphism and such that (see also Fig. 4.4)

a) if $p \in \Omega$, then $U_p \subset \Omega$ and $\varphi_p(U_p) \subset B_1(0) \subset \mathbb{R}^m$;
b) if $p \in \partial\Omega$, then $\varphi_p(U_p \cap \Omega) = \big(\mathbb{R}_+ \backslash \{0\}\big) \times \mathbb{R}^{m+1} \cap B_1(0)$ and $\varphi_p(U_p \cap \partial\Omega) = \{0\} \times \mathbb{R}^{m+1} \cap B_1(0)$.

Obviously, we have that $\bar{\Omega} \subset \bigcup_{p \in \bar{\Omega}} U_p$ and let U_0 be a subset of the complement of $\bar{\Omega}$ such that $U_0 \cup \big(\bigcup_{p\in\bar{\Omega}} U_p\big)$ covers M^m. Due to the compactness of $\bar{\Omega}$, we can extract a finite covering, i.e., the open sets $U_0, U_1 := U_{p_1}, \ldots, U_N := U_{p_N}$ cover M^m. Moreover, let $\theta_0, \theta_1, \ldots, \theta_N$ denote a partition of unity subordinated to this covering.

By linearity of the exterior derivative d and the integral, we then have (see also (3.75))

$$\int_\Omega d\omega = \int_\Omega d\left(\sum_{i=1}^N \theta_i \omega\right) = \sum_{i=1}^N \int_\Omega d(\theta_i\omega)\,. \tag{4.21}$$

The sum is finite and note that $d(\theta_i\omega) \in \Omega_0^{m-1}(M)$ is compactly supported in U_i. Using Definition 4.4 for the integration on manifolds, it follows for $i = 1, \ldots, N$ that

$$\begin{aligned}\int_\Omega d(\theta_i\omega) &= \int_{M^m} \chi_\Omega d(\theta_i\omega) \\ &= \int_{\varphi_i(U_i)} (\varphi_i^{-1})^* \big(\chi_\Omega d(\theta_i\omega)\big) \\ &= \int_{\varphi_i(U_i)} (\chi_\Omega \circ \varphi_i^{-1})(\varphi_i^{-1})^* \big(d(\theta_i\omega)\big)\,,\end{aligned}$$

where χ_Ω is the characteristic function of Ω and the $\varphi_i := \varphi_{p_i}$ verify a) and b). Clearly, $\chi_\Omega \circ \varphi_i^{-1} = \chi_{\varphi_i(U_i\cap\Omega)}$, we can thus write

$$\int_\Omega d(\theta_i\omega) = \int_{\varphi_i(U_i)} \chi_{\varphi_i(U_i\cap\Omega)} d\big((\varphi_i^{-1})^*(\theta_i\omega)\big)\,, \tag{4.22}$$

where we also used that d and the pull-back commute. .

If we denote by $U_{Q+1}, \ldots, U_N$ the open sets of M^m which intersect $\partial\Omega$, (4.21) can be rewritten as

$$\int_\Omega d\omega = \sum_{i=1}^Q \int_\Omega d(\theta_i\omega) + \sum_{i=Q+1}^N \int_\Omega d(\theta_i\omega)\,. \tag{4.23}$$

For $i = 1, \ldots, Q$, (4.22) then implies

$$\int_\Omega d(\theta_i\omega) = \int_{B_1(0)} d\big((\varphi_i^{-1})^*(\theta_i\omega)\big) = \int_{\mathbb{R}^m} d\big((\varphi_i^{-1})^*(\theta_i\omega)\big)\,. \tag{4.24}$$

Note that $(\varphi_i^{-1})^*(\theta_i\omega) \in \Omega_1^{m-1}(\mathbb{R}^m)$ is compactly supported in $\varphi_i(U_i) \subset B_1(0)$. From Lemma 4.14 below, we then deduce that the right-hand side of (4.24) vanishes.

For $i = Q+1, \ldots, N$, we have

$$\int_\Omega d(\theta_i\omega) \overset{(4.22)}{=} \int_{\left(\mathbb{R}_+\setminus\{0\}\right)\times\mathbb{R}^{m-1}} d\big((\varphi_i^{-1})^*(\theta_i\omega)\big)\,.$$

From Lemma 4.15 below, we then obtain

$$\begin{aligned}\int_\Omega d(\theta_i\omega) &= \int_{\{0\}\times\mathbb{R}^{m-1}} \iota^*_{\{0\}\times\mathbb{R}^{m-1}}\big((\varphi_i^{-1})^*(\theta_i\omega)\big)\\ &= \int_{\partial\Omega\cap U_i} \iota^*_{\partial\Omega}\theta_i\omega = \int_{\partial\Omega}\iota^*_{\partial\Omega}\theta_i\omega\,. \end{aligned}\tag{4.25}$$

Combining (4.23) with (4.24) and (4.25), we arrive at

$$\int_\Omega d\omega = \sum_{i=Q+1}^{N}\int_{\partial\Omega}\iota^*_{\partial\Omega}\theta_i\omega = \int_{\partial\Omega}\iota^*_{\partial\Omega}\omega\,.$$

□

Lemma 4.14. *Let $\beta \in \Omega_1^{m-1}(\mathbb{R}^m)$ compactly supported. Then, we have that*

$$\int_{\mathbb{R}^m} d\beta = 0\,.$$

Proof. Let $\beta \in \Omega_1^{m-1}(\mathbb{R}^m)$ be given in the following representation

$$\beta = \sum_{k=1}^{m}(-1)^{k-1}\beta_k\,dx_1\wedge\ldots\wedge\widehat{dx_k}\wedge\ldots\wedge dx_m\,.\tag{4.26}$$

Then, using (3.42), it is easy to check that

$$d\beta = \operatorname{div} X_\beta\,dx_1\wedge\ldots\wedge dx_m\,,\tag{4.27}$$

where $X_\beta = (\beta_1,\ldots,\beta_m)$ is a C^1-vector field on $\mathbb{R}^m$. From Definition 4.2, we then deduce that

$$\int_{\mathbb{R}^m} d\beta = \int_{\mathbb{R}^m}\operatorname{div} X_\beta(x)\,d\mathcal{L}^m(x)\,.$$

Since β is compactly supported, we can assume that $\operatorname{supp}\beta \subset B_R(0) \subset \mathbb{R}^m$ for some $R > 0$. Hence, we get (note that $\operatorname{div} X_\beta \in L^1(\mathbb{R}^m)$)

$$\int_{\mathbb{R}^m}\operatorname{div} X_\beta(x)\,d\mathcal{L}^m(x) = \int_{B_R(0)}\operatorname{div} X_\beta(x)\,d\mathcal{L}^m(x) = \int_{[-R,R]^m}\operatorname{div} X_\beta(x)\,d\mathcal{L}^m(x)$$

From (4.27) and since integration is a linear operation, it follows

$$\int_{\mathbb{R}^m} d\beta = \sum_{k=1}^{m} \int_{[-R,R]^m} \frac{\partial \beta_k}{\partial x_k}(x)\, d\mathcal{L}^m(x)\,.$$

For every $1 \leq k \leq m$, the theorem of Fubini implies that

$$\int_{[-R,R]^m} \frac{\partial \beta_k}{\partial x_k}(x)\, d\mathcal{L}^m(x) = \int_{[-R,R]^{m-1}} \left(\int_{[-R,R]} \frac{\partial \beta_k}{\partial x_k}(x)\, dx_k \right) dx_1 \dots \hat{dx}_k \dots dx_m\,,$$

where a different notation is used for the Lebesgue measure on the right-hand side. Because $\operatorname{supp}\beta \subset B_R(0)$, we have

$$\int_{[-R,R]} \frac{\partial \beta_k}{\partial x_k}(x)\, dx_k = \beta_k(R) - \beta_k(-R) = 0\,.$$

This gives the result. □

Lemma 4.15. *Let $\beta \in \Omega_1^{m-1}(\mathbb{R}^m)$ compactly supported. Then, we have that*

$$\int_{\mathbb{R}_+ \times \mathbb{R}^{m-1}} d\beta = \int_{\{0\} \times \mathbb{R}^{m-1}} \iota^*_{\{0\} \times \mathbb{R}^{m-1}} \beta\,.$$

Proof. With the same notations as in the proof of the previous lemma, we can write

$$\begin{aligned} \int_{\mathbb{R}_+ \times \mathbb{R}^{m-1}} d\beta &= \int_{x_1=0}^{R} \int_{[-R,R]^{m-1}} \operatorname{div} X_\beta(x)\, d\mathcal{L}^m(x) \\ &= \int_{x_1=0}^{R} \int_{[-R,R]^{m-1}} \operatorname{div} X_\beta(x)\, dx_1 dx_2 \dots dx_m\,. \end{aligned}$$

Using the theorem of Fubini, the right-hand side of the last equation becomes

$$\begin{aligned} &\int_{x_1=0}^{R} \left(\int_{[-R,R]^{m-1}} \frac{\partial \beta_1}{\partial x_1}(x)\, dx_2 \dots dx_m \right) dx_1 \\ &+ \sum_{k=2}^{m} \int_{x_1=0}^{R} \left[\int_{[-R,R]^{m-1}} \frac{\partial \beta_k}{\partial x_k}(x)\, dx_2 \dots dx_m \right] dx_1\,. \end{aligned} \tag{4.28}$$

The integrals in brackets vanish by an analogous argument as in the proof of the previous lemma.

For the first term of (4.28) we obtain, from the fact that $\operatorname{supp}\beta \subset B_R(0)$ by assumption,

$$\int_{x_1=0}^{R} \left(\int_{[-R,R]^{m-1}} \frac{\partial \beta_1}{\partial x_1}(x)\, dx_2 \dots dx_m \right) dx_1 = \int_{[-R,R]^{m-1}} \beta_1(0, x_2, \dots, x_m)\, dx_2 \dots dx_m\,.$$

In summary, we have shown that

$$\begin{aligned}\int_{\mathbb{R}_+\times\mathbb{R}^{m-1}} d\beta &= \int_{[-R,R]^{m-1}} \beta_1(0,x_2,\ldots,x_m)\,dx_2\ldots dx_m \\ &\overset{(4.2)}{=} \int_{\mathbb{R}^{m-1}} \beta_1(0,x_2,\ldots,x_m)\,dx_2\wedge\ldots\wedge dx_m\,. \end{aligned} \tag{4.29}$$

On the other hand, it is straightforward to observe that, for $k=2,\ldots,m$,

$$\iota^*_{\{0\}\times\mathbb{R}^{m-1}} dx_1\wedge\ldots\wedge\widehat{dx_k}\wedge\ldots dx_m = 0\,.$$

This implies, recalling (4.26),

$$\iota^*_{\{0\}\times\mathbb{R}^{m-1}}\beta = \beta_1(0,x_2,\ldots,x_m)\,dx_2\wedge\ldots dx_m\,.$$

Inserting the last expression in (4.29), then gives the result. □

Application of Stokes' Theorem

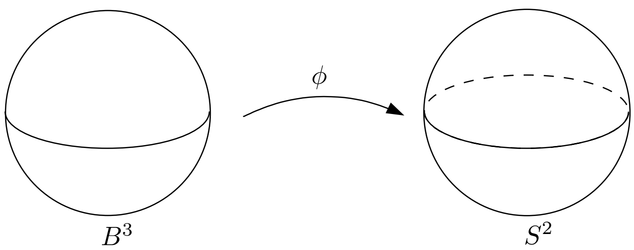

Fig. 4.5. Map from B^3 to $S^2=\partial B^3$.

We want to answer the question if there exists a smooth map $\phi: B^3 \longrightarrow S^2$ such that $\phi(x)=x$ on $S^2=\partial B^3$. Roughly speaking, we ask if it is possible to compress in a smooth way B^3 to its boundary leaving the boundary itself unchanged. – Obviously, this can be done by the map $x\longmapsto \frac{x}{|x|}$, $x\in B^3$, *without* considering the origin.

The answer to the question before is given by the following theorem.

Theorem 4.16 (Brouwer). *Let*

$$\phi: B^m \longrightarrow S^{m-1}=\partial B^m$$

be a C^1-map. Then the restriction of ϕ to ∂B^m can not be the identity on ∂B^m.

Proof. We prove the theorem by contradiction. – First note that

$$\det J_x\phi = 0\,, \qquad \text{for all } x \in B^m\,, \tag{4.30}$$

being a direct consequence of the Local Inversion Theorem 1.10. More precisely, without (4.30), this theorem would give the existence of a diffeomorphism from U open containing $x \in B^3$ into V open containing $\phi(x) \in S^2$.

Let $dx_1 \wedge \ldots \wedge dx_m \in \Omega^m(S^{m-1})$. Since the pull-back of this differential form equals multiplication by $\det J\phi$, it follows (see also (3.74))

$$0 = \phi^* dx_1 \wedge \ldots \wedge dx_m \in \Omega^m(B^m)\,,$$

and also that

$$\int_{B^m} \phi^* dx_1 \wedge \ldots \wedge dx_m = 0\,.$$

Denoting by $(\phi_1, \ldots, \phi_m)$ the coordinate functions of ϕ, we obtain, with Proposition 3.24 and (3.46),

$$\int_{B^m} d\phi_1 \wedge \ldots \wedge d\phi_m = 0\,.$$

From $d\phi_1 \wedge \ldots \wedge d\phi_m = d(\phi_1\, d\phi_2 \wedge \ldots \wedge d\phi_m)$ and Stokes' Theorem 4.13, we deduce that

$$\begin{aligned} 0 &= \int_{B^m} d(\phi_1\, d\phi_2 \wedge \ldots \wedge d\phi_m) \\ &= \int_{\partial B^m} \iota^*_{\partial B^m}(\phi_1\, d\phi_2 \wedge \ldots \wedge d\phi_m)\,. \end{aligned} \tag{4.31}$$

Moreover, using again Proposition 3.24 and (3.46), we get

$$\iota^*_{\partial B^m}(\phi_1\, d\phi_2 \wedge \ldots \wedge d\phi_m) = \phi_1 \circ \iota_{\partial B^m}\, d(\iota^*_{\partial B^m}\phi_2) \wedge \ldots \wedge d(\iota^*_{\partial B^m}\phi_m)\,.$$

Since by assumption $\phi|_{\partial B^m} = id|_{\partial B^m}$, it follows that

$$\iota^*_{\partial B^m}(\phi_1\, d\phi_2 \wedge \ldots \wedge d\phi_m) = x_1\, dx_2 \wedge \ldots \wedge dx_m\,.$$

Inserting this result in (4.31), leads to

$$0 = \int_{\partial B^m} x_1\, dx_2 \wedge \ldots \wedge dx_m\,.$$

Using again Stokes' theorem, we obtain the contradiction:

$$0 = \int_{B^m} dx_1 \wedge \ldots \wedge dx_m = \int_{B^m} 1\, d\mathcal{L}^m = \mathrm{Vol}(B^m)\,.$$

□

Particular Cases of Stokes' Theorem

1. Let $\Omega \subset \mathbb{R}^m = M^m$ be an oriented domain of $\mathbb{R}^m$ such that $\bar{\Omega}$ is compact. Moreover, let $\alpha \in \Omega^{m-1}(\Omega)$ be given by the representation

$$\alpha = \sum_{k=1}^{m} (-1)^{k-1} \alpha_k \, dx_1 \wedge \ldots \wedge \widehat{dx_k} \wedge \ldots \wedge dx_m \, .$$

Then, we obtain (see the proof of Lemma 4.14)

$$d\alpha = \operatorname{div} X_\alpha \, dx_1 \wedge \ldots \wedge dx_m \, , \tag{4.32}$$

where $X_\alpha = (\alpha_1, \ldots, \alpha_m)$ is a C^∞-vector field on $\Omega \subset \mathbb{R}^m$.

On the other hand, extending Proposition 3.40 to m dimensions, we get

$$\iota^*_{\partial\Omega}\alpha = \langle X_\alpha, \boldsymbol{n} \rangle \, dA_{\partial\Omega} \, , \tag{4.33}$$

where $\langle \cdot, \cdot \rangle$ denotes the usual scalar product on $\mathbb{R}^m$ and the map $\boldsymbol{n} \in C^\infty(\Omega, \mathbb{R}^m)$ denotes the orientation for the $(m-1)$-dimensional submanifold $\partial\Omega$ of $\mathbb{R}^m$ (see Proposition 4.12 and Proposition 3.55).

Inserting (4.32) and (4.33) into Stokes' Theorem 4.13, we end up with Gauß formula:

$$\int_\Omega \operatorname{div} X_\alpha \, dx_1 \wedge \ldots \wedge dx_m = \int_{\partial\Omega} \langle X_\alpha, \boldsymbol{n} \rangle \, dA_{\partial\Omega} \, . \tag{4.34}$$

2. We consider a surface Σ in $\mathbb{R}^3$ and let $\Omega \subset \Sigma$ be a oriented domain of Σ. Moreover, let $\beta = \beta_1 \, dx_1 + \beta_2 \, dx_2 + \beta_3 \, dx_3 \in \Omega^1(\mathbb{R}^3)$ and $\alpha = \beta|_\Sigma = \iota^*_\Sigma \beta$ be the restriction of the one-form β to Σ. In Section 3.1.2, we have shown that the exterior derivative $d\beta$ corresponds to $\mathit{rot}\, X_\beta$, where $X_\beta = (\beta_1, \beta_2, \beta_3)$ is a C^∞-vector field on $\mathbb{R}^3$. From Proposition 3.40, we then obtain

$$\int_\Omega d\alpha = \int_\Omega d(\iota^*_\Sigma \beta) = \int_\Omega \iota^*_\Sigma d\beta = \int_\Omega \langle \mathit{rot}\, X_\beta, \boldsymbol{n} \rangle \, dA_\Sigma \, . \tag{4.35}$$

On the other hand, denoting the one-dimensional boundary $\partial\Omega$ of Ω by Γ, we have

$$\int_{\partial\Omega} \iota^*_{\partial\Omega}\alpha = \int_\Gamma \iota^*_\Gamma \beta = \int_\Gamma \langle X_\beta, \boldsymbol{t} \rangle \, dl_\Gamma \, , \tag{4.36}$$

where we used Proposition 3.36. Recall that dl_Γ denotes the length form on Γ with respect to the orientation $\boldsymbol{t}$ (see Definition 3.35). Putting (4.35) and (4.36) together, Stokes' theorem has the well-known form:

$$\int_\Omega \langle \mathit{rot}\, X_\beta, \boldsymbol{n} \rangle \, dA_\Sigma = \int_\Gamma \langle X_\beta, \boldsymbol{t} \rangle \, dl_\Gamma \, . \tag{4.37}$$

5 Riemannian Geometry of Curves and Surfaces in $\mathbb{R}^3$

5.1 Local and Global Geometry of Curves in $\mathbb{R}^3$

Definition 5.1. *Let $k \geq 1$. A* ***parameterized*** C^k***-curve*** *in $\mathbb{R}^3$ is a C^k-map γ from an interval $I \subset \mathbb{R}$ into $\mathbb{R}^3$.*

Remark. If the interval I is *not* open the map γ can be seen as restriction of a C^k-map defined on an open interval containing I.

Example 5.2. For simplicity, we take examples of curves with images lying in $\mathbb{R}^2$. – The map $\gamma : \mathbb{R} \longrightarrow \mathbb{R}^2$, $t \longmapsto (t, |t|)$ is *not* a parameterized C^1-curve, since $|t|$ is not differentiable at $t = 0$. However, the map $\gamma : \mathbb{R} \longrightarrow \mathbb{R}^2$, $t \longmapsto (t^3, t^2)$ defines a parameterized C^1-curve. Note that $\dot{\gamma}(0) = (0, 0)$, showing that γ is not regular in the sense of the following definition.

Definition 5.3. *A* ***regular*** *curve is a parameterized C^1-curve γ such that $\dot{\gamma}(t) \neq 0$, for all $t \in I$. We call $\mathbf{t}(t) := \dot{\gamma}(t) \in \mathbb{R}^3$ the* ***tangent vector*** *of γ at t. Moreover, the* ***arc length*** *of the curve γ from a fixed point $\gamma(t_0)$, $t_0 \in I$, is defined by*

$$s(t) = \int_{t_0}^{t} \|\dot{\gamma}(\tau)\| \, d\tau \,. \tag{5.1}$$

Note that $s : I \subset \mathbb{R} \longrightarrow \mathbb{R}_+$ is a strictly increasing function. – If $\|\dot{\gamma}(t)\| = 1$, for all $t \in I$, then clearly $s(t) = t - t_0$ for the arc length, and we say that γ is *parameterized by arc length.*

Lemma 5.4. *Let γ be a regular parameterized curve in $\mathbb{R}^3$. Then γ can be parameterized by arc length.*

Proof. Let $[a, b] = I \subset \mathbb{R}$ and let L denote the total length of the regular parameterized curve γ, i.e., $L := s(b) = \int_a^b \|\dot{\gamma}(\tau)\| \, d\tau$. From (5.1), we deduce that, for all $t \in [a, b]$,

$$\frac{ds}{dt} = \|\dot{\gamma}(t)\| \neq 0 \,.$$

Hence by the Local Inversion Theorem 1.10, the C^1-inverse $s^{-1} : [0, L] \longrightarrow [a, b]$ exists. Defining $\tilde{\gamma}(\sigma) := \gamma\big(s^{-1}(\sigma)\big)$, for $\sigma \in [0, L]$, we obtain a curve parameterized by arc length. Indeed, we see that

$$\left\| \frac{d\tilde{\gamma}}{d\sigma} \right\| = \left\| \dot{\gamma}\big(s^{-1}(\sigma)\big) \frac{ds^{-1}}{d\sigma} \right\| = 1 \,.$$

□

In the following, we denote by $\gamma(s)$ curves being parameterized by arc length.

Definition 5.5. *Let γ be a C^2-curve parameterized by arc length. We define the* ***curvature*** *of γ at $s \in I$ by the number $k(s) := \|\ddot{\gamma}(s)\|$.*

We easily see that if $k(s) \equiv 0$, then the curve γ is a straight line. Moreover, it is intuitively clear, that the curvature at point measures the deviation of the curve from the tangent vector at this point, also called *bending* of the curve.

Assuming that $\ddot{\gamma}(s) \neq 0$, for a point $s \in I$, we introduce

$$\boldsymbol{n}(s) = \frac{\ddot{\gamma}(s)}{\|\ddot{\gamma}(s)\|} \,, \tag{5.2}$$

the unit **normal vector** to γ at s. We call the plane in $\mathbb{R}^3$ generated by the vectors $\boldsymbol{t}(s)$ and $\boldsymbol{n}(s)$ the **osculating plane** to γ at s. Moreover, we define

$$\boldsymbol{b}(s) := \boldsymbol{t}(s) \times \boldsymbol{n}(s) \,, \tag{5.3}$$

where $\times$ denotes the usual vector product in $\mathbb{R}^3$.

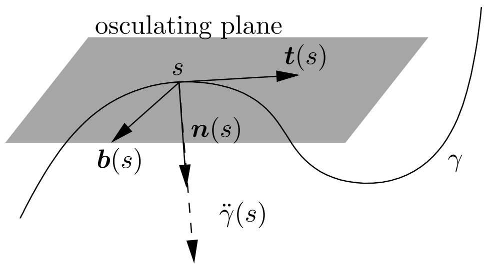

Fig. 5.1. Curve in $\mathbb{R}^3$.

Next, we would like to measure the tendency of a curve to escape from the osculating plane, also called *twisting* of the curve. For this purpose, we consider the derivative of $\boldsymbol{b}(s)$ and compute

$$\dot{\boldsymbol{b}}(s) = \dot{\boldsymbol{t}}(s) \times \boldsymbol{n}(s) + \boldsymbol{t}(s) \times \dot{\boldsymbol{n}}(s) \overset{(5.2)}{=} \boldsymbol{t}(s) \times \dot{\boldsymbol{n}}(s) \,.$$

Thus, we deduce that $\dot{\boldsymbol{b}}(s)$ is parallel to $\boldsymbol{n}$ and the following definition makes sense.

Definition 5.6. *Let γ be a C^3-curve parameterized by arc length. We define the **torsion** of γ at $s \in I$ by the number $\tau(s)$ such that $\dot{\boldsymbol{b}}(s) = \tau(s)\,\boldsymbol{n}(s)$.*

Note that the torsion indeed measures the twisting of the curve and we state without proof the following

Proposition 5.7. *Let γ be a C^3-curve parameterized by arc length. Then γ lies in a plane of $\mathbb{R}^3$, i.e., is a **planar curve**, if and only if $\tau(s) \equiv 0$.*

Remark. Obviously, planar curves can be interpreted as maps from an interval of $\mathbb{R}$ into $\mathbb{R}^2$ (see also Example 5.2).

Definition 5.8. *Let γ be a C^3-curve parameterized by arc length. If $\ddot{\gamma}(s) \neq 0$, for all $s \in I$, then we call γ a **Frenet curve**. Moreover, the **Frenet trihedron** is formed by the orthonormal frame $\{\boldsymbol{t}(s), \boldsymbol{n}(s), \boldsymbol{b}(s)\}$.*

For a Frenet curve, we have by definition $\dot{\boldsymbol{b}}(s) = \tau(s)\,\boldsymbol{n}(s)$ and we deduce directly from (5.2) and Definition 5.5 that $\dot{\boldsymbol{t}}(s) = k(s)\,\boldsymbol{n}(s)$. Then, it follows

$$
\begin{aligned}
\dot{\boldsymbol{n}}(s) &= \dot{\boldsymbol{b}}(s) \times \boldsymbol{t}(s) + \boldsymbol{b}(s) \times \dot{\boldsymbol{t}}(s) \\
&= \tau(s)\,\boldsymbol{n}(s) \times \boldsymbol{t}(s) + k(s)\,\boldsymbol{b}(s) \times \boldsymbol{n}(s) \\
&= -\tau(s)\,\boldsymbol{b}(s) - k(s)\,\boldsymbol{t}(s)\,.
\end{aligned}
$$

In summary, we end up with the following system of ordinary differential equations called **Frenet equations**:

$$
\begin{cases}
\dot{\boldsymbol{t}}(s) &= k(s)\,\boldsymbol{n}(s) \\
\dot{\boldsymbol{n}}(s) &= -\tau(s)\,\boldsymbol{b}(s) - k(s)\,\boldsymbol{t}(s) \\
\dot{\boldsymbol{b}}(s) &= \tau(s)\,\boldsymbol{n}(s)\,.
\end{cases}
\tag{5.4}
$$

Theorem 5.9 (Fundamental Theorem of the Local Theory of Curves). *Let I be an interval of $\mathbb{R}$ and let $k : I \longrightarrow \mathbb{R}_+ \setminus \{0\}$, $\tau : I \longrightarrow \mathbb{R}$ be two given C^∞-functions. Moreover, assume that for a fixed $s_0 \in I$ a point $x \in \mathbb{R}^3$ and an orthonormal frame $\{e_1^0, e_2^0, e_3^0\}$ of $\mathbb{R}^3$ are given. Then there exists a unique C^∞-Frenet curve $\gamma : I \longrightarrow \mathbb{R}^3$ solution of the Frenet equations (5.4) such that*

(i) $\gamma(s_0) = x$;
(ii) *the Frenet trihedron of γ at s_0 is $\{e_1^0, e_2^0, e_3^0\}$;*
(iii) *the curvature and torsion of γ are, respectively, k and τ.*

This important theorem closes the part concerning the *local* theory of curves in $\mathbb{R}^3$. – Now, we come to the *global* theory of curves in $\mathbb{R}^3$ dealing with closed and simple (planar) curves in the sense of the following

Definition 5.10. *A regular parameterized C^k-curve $\gamma : [a,b] \subset \mathbb{R} \longrightarrow \mathbb{R}^n$, $n = 2, 3$, is called **closed** if γ and all its derivatives agree at a and b, i.e., $\gamma^{(l)}(a) = \gamma^{(l)}(b)$, for all $l = 1, \ldots, k$. Moreover, a regular parameterized C^k-curve γ is said to be **simple** if $\gamma|_{[a,b)}$ is injective.*

Theorem 5.11 (Jordan). *Let $\gamma : [a,b] \longrightarrow \mathbb{R}^2$ be a closed and simple planar curve. Then there exists a bounded domain Ω of $\mathbb{R}^2$ such that*

$$\partial\Omega = \bar{\Omega} \setminus \Omega = \gamma\big([a,b]\big)\,.$$

Proof. □

Jordan's Theorem then motivates the following two (equivalent) questions:

a) What shape does the domain Ω with given area A must have, in order to minimize the length L of its boundary γ?
b) What shape does the curve γ with given length L must have, in order to maximize the area A of the enclosed domain Ω?

As expected, it turns out that the optimal shape is given by the circle in $\mathbb{R}^2$. More precisely, we have the following important global theorem.

Theorem 5.12 (Isoperimetric Inequality). *Let γ be a closed and simple planar C^1-curve with total length L. Moreover, let A be the area of the enclosed domain Ω. Then, we have that*

$$L^2 \geq 4\pi\, A\,, \tag{5.5}$$

with equality if and only if γ is a circle in $\mathbb{R}^2$.

Proof. Let $\gamma : [0,L] \longrightarrow \mathbb{R}^2$ be a closed and simple planar curve parameterized by arc length such that $\gamma\big([0,L]\big) = \partial\Omega$ for a domain Ω of $\mathbb{R}^2$ (see Lemma 5.4 and Jordan's Theorem 5.11).

We denote by $\tilde{\Omega}$ the *convex hull* of Ω, i.e, the intersection of all convex subsets of $\mathbb{R}^2$ containing Ω. Then, we see that $A(\tilde{\Omega}) \geq A(\Omega)$ and $L(\tilde{\gamma}) \leq L(\gamma)$, where $\tilde{\gamma}$ denotes the closed and simple planar curve enclosing the convex hull $\tilde{\Omega}$. Hence, if (5.5) holds for $\tilde{\Omega}$, we get

$$A(\Omega) \leq A(\tilde{\Omega}) \leq \frac{1}{4\pi}\, L(\tilde{\gamma}) \leq \frac{1}{4\pi}\, L(\gamma),\,,$$

showing that it suffices to prove the isoperimetric inequality for the case when Ω is a *convex* domain of $\mathbb{R}^2$. – Moreover, because of the invariance of (5.5) under dilations, we can assume that $L = 2\pi$.

Next, we denote by Δ^+ and Δ^- the two half-planes of $\mathbb{R}^2$ such that $\Delta^+ \cup \Delta^- = \mathbb{R}^2$ and

$$\begin{aligned} L(\partial\Omega \cap \Delta^+) = L(\partial\Omega \cap \Delta^-) = \pi\,, \\ A(\Omega \cap \Delta^+) \geq A(\Omega \cap \Delta^-)\,. \end{aligned}$$

Then, we choose coordinates (x,y) on $\mathbb{R}^2$ in such a way that

$$\Delta^+ = \big\{(x,y) \in \mathbb{R}^2 \;:\; y \geq 0\big\}\,.$$

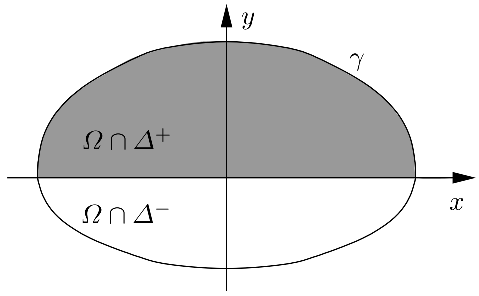

Fig. 5.2. Setting for the proof of the isoperimetric inequality.

For the area A of the convex domain Ω, we have

$$A(\Omega) \leq 2\,A(\Omega \cap \Delta^+) = 2\mathcal{L}^2(\Omega \cap \Delta^+) = 2\int_{\Omega \cap \Delta^+} dx \wedge dy\,, \tag{5.6}$$

where we used Definition 4.2 for the integration of differential forms on Euclidean space. Writing $dx \wedge dy = d(-y\,dx)$, Stokes' Theorem 4.13 implies that

$$2\int_{\Omega \cap \Delta^+} dx \wedge dy = 2\int_{\partial(\Omega \cap \Delta^+)} \iota^*_{\partial(\Omega \cap \Delta^+)}(-y\,dx)\,.$$

Since $\partial(\Omega \cap \Delta^+) = (\Omega \cap \partial\Delta^+) \cup (\partial\Omega \cap \Delta^+)$, the last integral becomes (note that $y = 0$ on $\partial\Delta^+$)

$$2\int_{\partial(\Omega \cap \Delta^+)} \iota^*_{\partial(\Omega \cap \Delta^+)}(-y\,dx) = -2\int_{\partial\Omega \cap \Delta^+} y\,dx\,,$$

where $y\,dx$ now denotes the restricted one-form to $\partial\Omega \cap \Delta^+$, i.e., $\iota^*_{\partial\Omega \cap \Delta^+}(y\,dx) = y\,dx|_{\partial\Omega \cap \Delta^+}$.

By convexity of the domain Ω it is now possible to parameterize $\partial\Omega \cap \Delta^+$ by the curve $\gamma|_{[0,\pi]}$, i.e., $\gamma\big([0,\pi]\big) = \partial\Omega \cap \Delta^+$. Note that in the case of Ω being *not* convex, the length of the curve parameterizing $\partial\Omega \cap \Delta^+$ would by larger than π, because the different components of $\partial\Omega \cap \Delta^+$ need to be connected. – So writing $y\big(\gamma(s)\big) = y(s)$ and $x\big(\gamma(s)\big) = x(s)$ for $s \in [0,\pi]$, we obtain from Proposition 4.3 that

$$-2\int_{\partial\Omega \cap \Delta^+} y\,dx = -2\int_{[0,\pi]} y(s)\dot{x}(s)\,ds\,. \tag{5.7}$$

Since γ is parameterized by arc length, i.e., $1 = \|\dot{\gamma}(s)\|^2 = \dot{x}^2(s) + \dot{y}^2(s)$, we see that

$$-2\,y(s)\dot{x}(s) \leq y^2(s) + \dot{x}^2(s) = y^2(s) + 1 - \dot{y}^2(s)\,.$$

Inserting this into (5.7) and recalling (5.6), we arrive at

$$A(\Omega) \leq \int_0^\pi y^2(s) + 1 - \dot{y}^2(s)\, ds\,. \tag{5.8}$$

Since γ is a C^1-curve by assumption, we can apply Poincaré's Lemma 5.13 (see below) to deduce that

$$A(\Omega) \leq \int_0^\pi \dot{y}^2(s) + 1 - \dot{y}^2(s)\, ds \leq \int_0^\pi 1\, ds = \pi = \frac{L^2}{4\pi}\,.$$

As a direct consequence of (5.6) and Poincaré's Lemma 5.13, equality holds if and only if $A(\Omega \cap \Delta^+) = A(\Omega \cap \Delta^-)$ and $\partial\Omega \cap \Delta^+$ is a half-circle. This concludes the proof of the isoperimetric inequality. □

Lemma 5.13 (Poincaré's Lemma). *Let $y \in C^1\big([0,\pi], \mathbb{R}\big)$ such that $y(0) = y(\pi) = 0$. Then, we have*

$$\int_0^\pi y^2(s)\, ds \leq \int_0^\pi \dot{y}^2(s)\, ds\,, \tag{5.9}$$

with equality if and only if $\gamma(s) = c\, \sin(s)$ for a constant $c \in \mathbb{R}$.

Proof. □

In the case of planar regular curves γ parameterized by arc length, it is possible to assign a sign to its curvature (compare with Definition 5.5). For this purpose, let $\boldsymbol{n}(s)$ denote the unit (normal) vector in $\mathbb{R}^2$ such that $\{\dot{\gamma}(s), \boldsymbol{n}(s)\}$ is a positive oriented basis of $\mathbb{R}^2$. Then, we get the following

Definition 5.14. *Let γ be a planar regular C^2-curve parameterized by arc length. We define the **signed curvature** $k(s)$ of γ at s by $\ddot{\gamma}(s) = k(s)\, \boldsymbol{n}(s)$.*

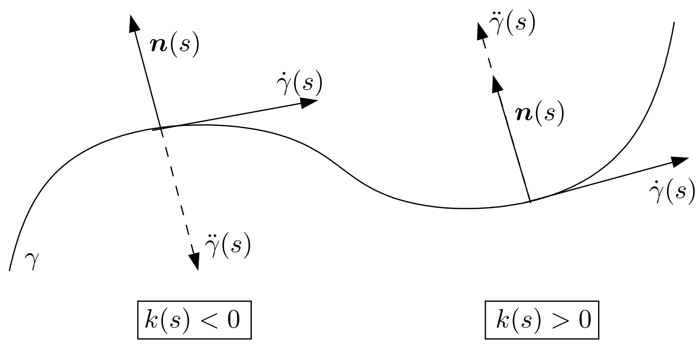

Fig. 5.3. Signed curvature of a planar curve.

Introducing polar coordinates in $\mathbb{R}^2$, the unit tangent vector of γ at s can be written as $\dot{\gamma}(s) = \big(\cos\theta(s), \sin\theta(s)\big)$. Hence, it follows that $\ddot{\gamma}(s) = \dot{\theta}(s)\big(-\sin\theta(s), \cos\theta(s)\big)$. Since

$$\bigl\{\bigl(\cos\theta(s),\sin\theta(s)\bigr),\bigl(-\sin\theta(s),\cos\theta(s)\bigr)\bigr\}$$

is a positive oriented basis of $\mathbb{R}^2$, we can take $\boldsymbol{n}(s) = \bigl(-\sin\theta(s),\cos\theta(s)\bigr)$ and obtain $k(s) = \dot\theta(s)$ for the signed curvature. If in addition the curve γ is closed with length L (see Definition 5.10), it follows that

$$\theta(0) \equiv \theta(L) \mod 2\pi\,.$$

Thus, we can define the *rotation index* of γ, denoted by $\operatorname{Ind}\gamma$, as being the number in $\mathbb{Z}$ satisfying

$$\int_0^L k(s)\,ds = \theta(L) - \theta(0) = 2\pi \operatorname{Ind}\gamma\,. \tag{5.10}$$

Remark. The rotation index can be interpreted as topological degree of $\dot\gamma$ seen as map from S^1 into S^1 (see Section 4.2).

Theorem 5.15. *Let $\gamma : [0,L] \longrightarrow \mathbb{R}^2$ be a closed and simple planar curve parameterized by arc length. Then, we have $Ind\gamma = \pm 1$, i.e.,*

$$\int_0^L k(s)\,ds = \pm 2\pi\,.$$

Proof. We can also always assume that $\gamma\bigl([0,L]\bigr) \subset \mathbb{R}\times\mathbb{R}_+$ and $\gamma(0) = \gamma(L) = 0$, moreover that $\dot\gamma(0) = \dot\gamma(L) = e_1$ (see Fig. 5.4).

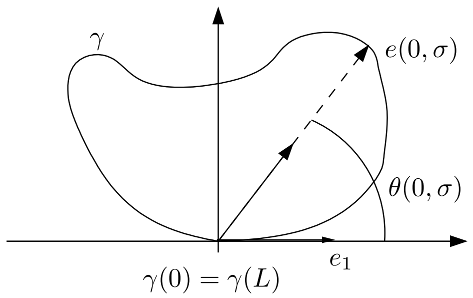

Fig. 5.4. Coordinates for the proof of Theorem 5.15.

Then we define

$$A = \bigl\{(s,\sigma) \in \mathbb{R}^2 \ :\ 0 \le s \le \sigma \le L\bigr\}\,.$$

and the map $e : A \longrightarrow S^1 \subset \mathbb{R}^2$ by

$$e(s,\sigma) = \begin{cases} \dfrac{\gamma(\sigma)-\gamma(s)}{\|\gamma(\sigma)-\gamma(s)\|} & \text{for } s \ne \sigma \text{ and } (s,\sigma) \ne (0,L) \\[2ex] \dot\gamma(\sigma) & \text{for } s = \sigma \\[2ex] -\dot\gamma(0) = -e_1 & \text{for } (s,\sigma) = (0,L)\,. \end{cases}$$

By assumption the planar curve γ is simple, i.e., for all $s, \sigma \in [0, L)$ with $s \neq \sigma$, it follows that $\gamma(s) \neq \gamma(\sigma)$. Hence, the map e is well-defined and its continuity can also easily be checked. Moreover, one sees that $e(\sigma, \sigma)$ gives the unit tangent vector of γ at σ.

It is left as an exercise to show the existence of a continuous function $\theta : A \longrightarrow \mathbb{R}$ such that $e(s, \sigma) = \big(\cos\theta(s, \sigma), \sin\theta(s, \sigma)\big)$ and $\theta(0, 0) = 0$. In a next step, we define the function $\theta(\sigma) := \theta(\sigma, \sigma)$. The tangent vector of γ at σ then reads as $\dot\gamma(\sigma) = e(\sigma, \sigma) = \big(\cos\theta(\sigma), \sin\theta(\sigma)\big)$.

For the signed curvature we then obtain (see (5.10))

$$\begin{aligned}\int_0^L k(\sigma)\, d\sigma &= \int_0^L \dot\theta(\sigma)\, d\sigma = \theta(L) - \theta(0) \\ &= \theta(L, L) - \theta(0, 0)\,.\end{aligned}$$

On the other hand, we have

$$\begin{aligned}\theta(0, L) - \theta(0, 0) &= \pi - 0 = \pi\,, \\ \theta(L, L) - \theta(0, L) &= 2\pi - \pi = \pi\,.\end{aligned}$$

Thus, we can write

$$\int_0^L k(\sigma)\, d\sigma = \big(\theta(L, L) - \theta(0, L)\big) + \big(\theta(0, L) - \theta(0, 0)\big) = 2\pi\,.$$

Note that we obtain the opposite sign in the last equation if we assume that $\dot\gamma(0) = \dot\gamma(L) = -e_1$. □

As a direct consequence of the previous theorem, we have that

$$\int_0^L |k(s)|\, ds \geq 2\pi\,, \tag{5.11}$$

with equality if and only if the signed curvature of the closed and simple planar curve γ does not change the sign. This condition is equivalent for the curve to be *convex*. – In addition to the derived results for closed and simple planar curves, the next theorem says that a certain amount of curvature is needed to close simple (space) curves in $\mathbb{R}^3$.

Theorem 5.16 (Fenchel's Theorem). *For every closed and simple curve $\gamma : [0, L] \longrightarrow \mathbb{R}^3$ parameterized by arc length, the following inequality for the so-called **total curvature** holds:*

$$\int_0^L k(s)\, ds \geq 2\pi\,, \tag{5.12}$$

with equality if and only if γ is a planar convex curve.

Introducing the concept of *knotted* curves in $\mathbb{R}^3$, we obtain a sharpening of Fenchel's Theorem.

Definition 5.17. *A closed and simple C^0-curve $\gamma : S^1 \longrightarrow \mathbb{R}^3$ is called **un-knotted** if there exists a homotopy $H : S^1 \times [0,1] \longrightarrow \mathbb{R}^3$ such that*

$$H(S^1 \times \{0\}) = S^1 ,$$
$$H(S^1 \times \{1\}) = \gamma ,$$

and $H(S^1 \times \{t\}) = \gamma_t \subset \mathbb{R}^3$ is homeomorphic to S^1.

Intuitively, this means that an un-knotted space curve can be deformed continuously into the circle S^1 with all intermediate positions homeomorphic to S^1. For the total curvature of *knotted* curves, i.e., curves which do not satisfy Definition 5.17, we then have the following result.

Theorem 5.18 (Fary-Milnor). *For every knotted (space) curve $\gamma : [0, L] \longrightarrow \mathbb{R}^3$ parameterized by arc length, the following inequality for the total curvature holds:*

$$\int_0^L k(s)\, ds \geq 4\pi . \tag{5.13}$$

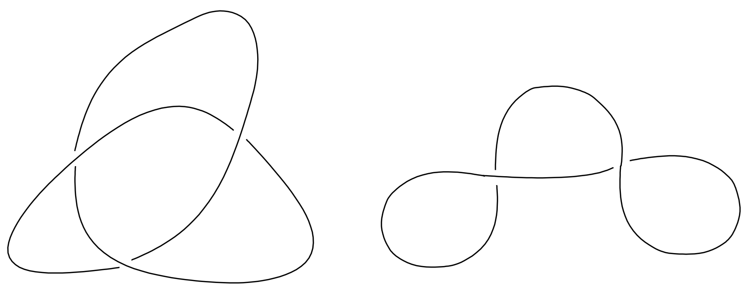

Fig. 5.5. Knotted and un-knotted space curves.

5.2 Local and Global Geometry of Surfaces in $\mathbb{R}^3$

Definition 5.19. *We define a C^k-**surface** of $\mathbb{R}^3$ to be a two-dimensional C^k-submanifold of $\mathbb{R}^3$.*

Remark. In the following, we will only consider C^∞-surfaces.

Let Σ be a surface in $\mathbb{R}^3$ and $p \in \Sigma$. The canonical scalar product of $\mathbb{R}^3$ induces on each tangent space $T_p\Sigma$ a scalar product, denoted by $\langle \cdot, \cdot \rangle_p$. More precisely, if $X_1, X_2 \in T_p\Sigma \subset \mathbb{R}^3$, then $\langle X_1, X_2 \rangle_p$ equals the scalar product

$\langle X_1, X_2\rangle$ of X_1 and X_2 in $\mathbb{R}^3$. Note that we make no difference in the notation for an element of the tangent space $T_p\Sigma$ and its image by the map $d\iota_\Sigma$, where ι_Σ denotes as usual the canonical inclusion of Σ in $\mathbb{R}^3$. – Recall also that a scalar product is a symmetric bilinear form.

Definition 5.20. *Let Σ be a surface in $\mathbb{R}^3$ and $p \in \Sigma$. The restriction of the canonical scalar product in $\mathbb{R}^3$ to $T_p\Sigma$, denoted by $\langle\cdot,\cdot\rangle_p$ or also $I_p(\cdot,\cdot)$, defines the* ***first fundamental form*** *of Σ at p.*

Let (U, φ) be a local chart for the surface Σ about $p \in U \subset \Sigma$ in the sense of Example 2.20. In this chart, we can write $X, Y \in T_p\Sigma$ as

$$X = \sum_{i=1}^{2} X_i \frac{\partial}{\partial\varphi_i}(p)\,, \qquad Y = \sum_{i=1}^{2} Y_i \frac{\partial}{\partial\varphi_i}(p)\,.$$

The first fundamental form I_p then reads as

$$\begin{aligned} I_p(X,Y) &= \langle X, Y\rangle_p && (5.14)\\ &= X_1Y_1 \left\langle \frac{\partial}{\partial\varphi_1}(p), \frac{\partial}{\partial\varphi_1}(p)\right\rangle_p + X_1Y_2 \left\langle \frac{\partial}{\partial\varphi_1}(p), \frac{\partial}{\partial\varphi_2}(p)\right\rangle_p \\ &+ X_2Y_1 \left\langle \frac{\partial}{\partial\varphi_2}(p), \frac{\partial}{\partial\varphi_1}(p)\right\rangle_p + X_2Y_2 \left\langle \frac{\partial}{\partial\varphi_2}(p), \frac{\partial}{\partial\varphi_2}(p)\right\rangle_p . \\ & && (5.15)\end{aligned}$$

Introducing for the coefficients of the first fundamental form in the chart (U, φ) the notations

$$\begin{aligned} E(p) &= \left\langle \frac{\partial}{\partial\varphi_1}(p), \frac{\partial}{\partial\varphi_1}(p)\right\rangle_p , \qquad F(p) = \left\langle \frac{\partial}{\partial\varphi_1}(p), \frac{\partial}{\partial\varphi_2}(p)\right\rangle_p ,\\ G(p) &= \left\langle \frac{\partial}{\partial\varphi_2}(p), \frac{\partial}{\partial\varphi_2}(p)\right\rangle_p , && (5.16)\end{aligned}$$

Equation (5.14) reduces to

$$I_p(X,Y) = X_1Y_1\,E(p) + X_1Y_2\,F(p) + X_2Y_1\,F(p) + X_2Y_2\,G(p)\,. \qquad (5.17)$$

Next, let $\gamma \in C^1([0,1], \Sigma \cap U)$ be a regular parameterized curve in Σ. If we write $\gamma(t) = \varphi^{-1}\big(c_1(t), c_2(t)\big)$ for the curve in the chart (U, φ), then it follows

$$\dot\gamma(t) = \dot c_1(t)\,\frac{\partial\varphi^{-1}}{\partial x_1}\big(\varphi(\gamma(t))\big) + \dot c_2(t)\,\frac{\partial\varphi^{-1}}{\partial x_2}\big(\varphi(\gamma(t))\big)\,.$$

By Definition 5.3 the arc length of γ is given by

$$s(t) = \int_0^t \|\dot\gamma(\tau)\|\,d\tau\,,$$

and using the notations introduced in (5.16), it follows

$$s(t) = \int_0^t \sqrt{\dot c_1^2(\tau)\, E\big(\gamma(\tau)\big) + 2\, \dot c_1(\tau)\dot c_2(\tau)\, F\big(\gamma(\tau)\big) + \dot c_2^2(\tau)\, G\big(\gamma(\tau)\big)}\,. \quad (5.18)$$

Hence, the importance of the first fundamental form comes from the fact that if it is given, then metric questions on a surface can be treated without further references to the ambient space $\mathbb{R}^3$.

Example 5.21 (First Fundamental Form for the Sphere S^2).

We have already seen in Section 3.3, that Σ is orientable if there exists a C^∞-map $\boldsymbol{n} : \Sigma \longrightarrow \mathbb{R}^3$ such that $\|\boldsymbol{n}(p)\| = 1$ and $\boldsymbol{n}(p) \perp T_p\Sigma$, for all $p \in \Sigma$, and an orientation for Σ is given by the choice of such a map $\boldsymbol{n}$. An orientation $\boldsymbol{n}$ can be constructed in the following way: Let $\iota_\Sigma : \Sigma \hookrightarrow \mathbb{R}^3$ denote the canonical inclusion and let (U, φ) be a local chart about p. Then we consider the C^∞-map, called *local parameterization* of Σ,

$$\tilde\varphi^{-1} := \iota_\Sigma \circ \varphi^{-1} : \varphi(\Sigma \cap U) \subset \mathbb{R}^2 \longrightarrow \Sigma \cap U \subset \mathbb{R}^3\,. \quad (5.19)$$

Since φ is a local diffeomorphism about p, we deduce that $d\tilde\varphi^{-1}_{\varphi(p)} \cdot e_i = \frac{\partial \tilde\varphi^{-1}}{\partial x_i}\big(\varphi(p)\big)$, for $i = 1, 2$, are two linearly independent vectors of $T_p\Sigma \subset \mathbb{R}^3$. Moreover, the map

$$\boldsymbol{n} = \frac{\dfrac{\partial \tilde\varphi^{-1}}{\partial x_1} \times \dfrac{\partial \tilde\varphi^{-1}}{\partial x_2}}{\left\| \dfrac{\partial \tilde\varphi^{-1}}{\partial x_1} \times \dfrac{\partial \tilde\varphi^{-1}}{\partial x_2} \right\|} \quad (5.20)$$

gives an orientation for Σ. (At this place, the remark at the end of Section 3.3 can be useful, see also (3.63).)

Remark. At this place, it is important to note that in terms of a local parameterization the first fundamental form reads as

$$\left\langle \frac{\partial}{\partial \varphi_i}(p), \frac{\partial}{\partial \varphi_j}(p) \right\rangle_p = \left\langle \frac{\partial \tilde\varphi^{-1}}{\partial x_i}\big(\varphi(p)\big), \frac{\partial \tilde\varphi^{-1}}{\partial x_j}\big(\varphi(p)\big) \right\rangle . \quad (5.21)$$

For later use, we note that, for $X, Y \in T_p\Sigma$, this can also be written as

$$\langle X, Y \rangle_p = \langle d\iota_\Sigma \cdot X, d\iota_\Sigma \cdot Y \rangle. \quad (5.22)$$

Definition 5.22. *Let Σ be a surface in $\mathbb{R}^3$ with orientation $\boldsymbol{n}$. We define the **Gauß map $\boldsymbol{n}$** of Σ by*

$$\begin{aligned} \boldsymbol{n} : \Sigma &\longrightarrow S^2\,, \\ p &\longmapsto \boldsymbol{n}(p)\,, \end{aligned} \quad (5.23)$$

Since $T_p\Sigma$ and $T_{\boldsymbol{n}(p)}S^2$ are parallel to each other, the tangent map $d\boldsymbol{n}_p$ to the Gauß map can be seen as linear map on $T_p\Sigma$. Moreover, it has an important property.

Proposition 5.23. *For any point $p \in \Sigma$, the tangent map $d\boldsymbol{n} : T\Sigma \longrightarrow TS^2$ of the Gauß map is self-adjoint with respect to the first fundamental form, i.e, for all $X, Y \in T_p\Sigma$, we have*

$$\langle d\boldsymbol{n}_p \cdot X, Y\rangle_p = \langle X, d\boldsymbol{n}_p \cdot Y\rangle_p \,. \tag{5.24}$$

Proof. Let (U, φ) be a chart about $p \in \Sigma$ and let $\tilde{\varphi}^{-1} = \iota_\Sigma \circ \varphi^{-1}$ be the local parameterization defined in (5.19). Moreover, let $\gamma, \bar{\gamma} \in C^1([-1,1], \Sigma \cap U)$ be two curves in Σ (seen as curves in $\mathbb{R}^3$) such that $\gamma(0) = \bar{\gamma}(0) = p$ and (see Definition 2.32)

$$X = \frac{d}{dt}\bigg|_{t=0} \gamma(t) \in T_p\Sigma \,, \qquad Y = \frac{d}{dt}\bigg|_{t=0} \bar{\gamma}(t) \in T_p\Sigma \,. \tag{5.25}$$

From Definition 2.40, we see that the tangent map $d\boldsymbol{n}_p$ of the Gauß map acts on X like

$$d\boldsymbol{n}_p \cdot X = \frac{d}{dt}\bigg|_{t=0} \big(\boldsymbol{n} \circ \gamma\big)(t) \,. \tag{5.26}$$

Writing $\big(\varphi \circ \gamma\big)(t) = c_1(t)\, e_1 + c_2(t)\, e_2 \in \mathbb{R}^2$, $t \in [-1,1]$, for the coordinate expression of γ, we then obtain

$$\frac{d}{dt}\big(\varphi \circ \gamma\big)(t) = \dot{c}_1(t)\, e_1 + \dot{c}_2(t)\, e_2 \,.$$

Thus for the map

$$\boldsymbol{n} \circ \gamma = \boldsymbol{n} \circ \tilde{\varphi}^{-1} \circ (\varphi \circ \gamma) =: \tilde{\boldsymbol{n}} \circ (\varphi \circ \gamma) : [-1,1] \subset \mathbb{R} \longrightarrow S^2 \subset \mathbb{R}^3 \,, \tag{5.27}$$

it follows, using the chain rule,

$$\begin{aligned} \frac{d}{dt}\big(\boldsymbol{n} \circ \gamma\big)(t) &= d\tilde{\boldsymbol{n}}_{\varphi(\gamma(t))} \cdot \big(\dot{c}_1(t)\, e_1 + \dot{c}_2(t)\, e_2\big) \\ &= \dot{c}_1(t)\, \frac{\partial \tilde{\boldsymbol{n}}}{\partial x_1}\big(\varphi(\gamma(t))\big) + \dot{c}_2(t)\, \frac{\partial \tilde{\boldsymbol{n}}}{\partial x_2}\big(\varphi(\gamma(t))\big) \,. \end{aligned}$$

Inserting this result in (5.26), we arrive at

$$d\boldsymbol{n}_p \cdot X = \dot{c}_1(0)\, \frac{\partial \tilde{\boldsymbol{n}}}{\partial x_1}\big(\varphi(p)\big) + \dot{c}_2(0)\, \frac{\partial \tilde{\boldsymbol{n}}}{\partial x_2}\big(\varphi(p)\big) \,. \tag{5.28}$$

In an analogous manner, we also get

$$d\boldsymbol{n}_p \cdot Y = \dot{\bar{c}}_1(0)\, \frac{\partial \tilde{\boldsymbol{n}}}{\partial x_1}\big(\varphi(p)\big) + \dot{\bar{c}}_2(0)\, \frac{\partial \tilde{\boldsymbol{n}}}{\partial x_2}\big(\varphi(p)\big) \,. \tag{5.29}$$

Moreover, we deduce similarly from (5.25) that

$$\begin{aligned} X &= \frac{d}{dt}\bigg|_{t=0} \big(\tilde{\varphi}^{-1} \circ \varphi \circ \gamma\big)(t) = \dot{c}_1(0)\,\frac{\partial \tilde{\varphi}^{-1}}{\partial x_1}\big(\varphi(p)\big) + \dot{c}_2(0)\,\frac{\partial \tilde{\varphi}^{-1}}{\partial x_2}\big(\varphi(p)\big)\,. \\ Y &= \frac{d}{dt}\bigg|_{t=0} \big(\tilde{\varphi}^{-1} \circ \varphi \circ \bar{\gamma}\big)(t) = \dot{\bar{c}}_1(0)\,\frac{\partial \tilde{\varphi}^{-1}}{\partial x_1}\big(\varphi(p)\big) + \dot{\bar{c}}_2(0)\,\frac{\partial \tilde{\varphi}^{-1}}{\partial x_2}\big(\varphi(p)\big)\,. \end{aligned} \tag{5.30}$$

From (5.28), respectively (5.29), and (5.30), we then get for the scalar products:

$$\begin{aligned} \langle d\boldsymbol{n}_p \cdot X, Y\rangle &= \bigg\langle \dot{c}_1(0)\,\frac{\partial \tilde{\boldsymbol{n}}}{\partial x_1}\big(\varphi(p)\big) + \dot{c}_2(0)\,\frac{\partial \tilde{\boldsymbol{n}}}{\partial x_2}\big(\varphi(p)\big)\,, \\ &\qquad \dot{\bar{c}}_1(0)\,\frac{\partial \tilde{\varphi}^{-1}}{\partial x_1}\big(\varphi(p)\big) + \dot{\bar{c}}_2(0)\,\frac{\partial \tilde{\varphi}^{-1}}{\partial x_2}\big(\varphi(p)\big)\bigg\rangle\,. \\ \langle X, d\boldsymbol{n}_p \cdot Y\rangle &= \bigg\langle \dot{c}_1(0)\,\frac{\partial \tilde{\varphi}^{-1}}{\partial x_1}\big(\varphi(p)\big) + \dot{c}_2(0)\,\frac{\partial \tilde{\varphi}^{-1}}{\partial x_2}\big(\varphi(p)\big)\,, \\ &\qquad \dot{\bar{c}}_1(0)\,\frac{\partial \tilde{\boldsymbol{n}}}{\partial x_1}\big(\varphi(p)\big) + \dot{\bar{c}}_2(0)\,\frac{\partial \tilde{\boldsymbol{n}}}{\partial x_2}\big(\varphi(p)\big)\bigg\rangle\,. \end{aligned}$$

For the difference of the last two expressions, it follows

$$\begin{aligned} &\langle d\boldsymbol{n}_p \cdot X, Y\rangle - \langle X, d\boldsymbol{n}_p \cdot Y\rangle \\ &= \dot{c}_1(0)\dot{\bar{c}}_2(0)\left(\bigg\langle \frac{\partial \tilde{\boldsymbol{n}}}{\partial x_1}\big(\varphi(p)\big), \frac{\partial \tilde{\varphi}^{-1}}{\partial x_2}\big(\varphi(p)\big)\bigg\rangle - \bigg\langle \frac{\partial \tilde{\varphi}^{-1}}{\partial x_1}\big(\varphi(p)\big), \frac{\partial \tilde{\boldsymbol{n}}}{\partial x_2}\big(\varphi(p)\big)\bigg\rangle\right) \\ &\quad + \dot{c}_2(0)\dot{\bar{c}}_1(0)\left(\bigg\langle \frac{\partial \tilde{\boldsymbol{n}}}{\partial x_2}\big(\varphi(p)\big), \frac{\partial \tilde{\varphi}^{-1}}{\partial x_1}\big(\varphi(p)\big)\bigg\rangle - \bigg\langle \frac{\partial \tilde{\varphi}^{-1}}{\partial x_2}\big(\varphi(p)\big), \frac{\partial \tilde{\boldsymbol{n}}}{\partial x_1}\big(\varphi(p)\big)\bigg\rangle\right)\,. \end{aligned} \tag{5.31}$$

Since the unit normal vector is perpendicular to the tangent space of the surface (see (5.20)), i.e.,

$$\bigg\langle \tilde{\boldsymbol{n}}\big(\varphi(p)\big), \frac{\partial \tilde{\varphi}^{-1}}{\partial x_1}\big(\varphi(p)\big)\bigg\rangle = 0\,,$$

we obtain by differentiating with respect to x_2 that

$$\bigg\langle \frac{\partial \tilde{\boldsymbol{n}}}{\partial x_2}\big(\varphi(p)\big), \frac{\partial \tilde{\varphi}^{-1}}{\partial x_1}\big(\varphi(p)\big)\bigg\rangle + \bigg\langle \tilde{\boldsymbol{n}}\big(\varphi(p)\big), \frac{\partial^2 \tilde{\varphi}^{-1}}{\partial x_2 \partial x_1}\big(\varphi(p)\big)\bigg\rangle = 0\,. \tag{5.32}$$

Similarly, we also get

$$\bigg\langle \frac{\partial \tilde{\boldsymbol{n}}}{\partial x_1}\big(\varphi(p)\big), \frac{\partial \tilde{\varphi}^{-1}}{\partial x_2}\big(\varphi(p)\big)\bigg\rangle + \bigg\langle \tilde{\boldsymbol{n}}\big(\varphi(p)\big), \frac{\partial^2 \tilde{\varphi}^{-1}}{\partial x_1 \partial x_2}\big(\varphi(p)\big)\bigg\rangle = 0\,. \tag{5.33}$$

And (5.32)−(5.33), then implies

$$\left\langle \frac{\partial \tilde{\boldsymbol{n}}}{\partial x_2}(\varphi(p)), \frac{\partial \tilde{\varphi}^{-1}}{\partial x_1}(\varphi(p)) \right\rangle - \left\langle \frac{\partial \tilde{\boldsymbol{n}}}{\partial x_1}(\varphi(p)), \frac{\partial \tilde{\varphi}^{-1}}{\partial x_2}(\varphi(p)) \right\rangle = 0\,.$$

After inserting this in (5.31), we get the result (5.24). □

Definition 5.24. *Let Σ be a surface in $\mathbb{R}^3$ with Gauß map $\boldsymbol{n}$. The* ***second fundamental form*** *II_p of Σ at $p \in \Sigma$ is the symmetric bilinear form on $T_p\Sigma$ defined by*

$$II_p(X,Y) = -\langle d\boldsymbol{n}_p \cdot X, Y\rangle_p\,, \qquad X, Y \in T_p\Sigma\,. \tag{5.34}$$

Remark. From Proposition 5.23, we deduce

$$II_p(X,Y) = -\langle d\boldsymbol{n}_p \cdot X, Y\rangle_p = -\langle X, d\boldsymbol{n}_p \cdot Y\rangle_p = -\langle d\boldsymbol{n}_p \cdot Y, X\rangle_p = II_p(Y,X)\,,$$

showing the the second fundamental form II_p is indeed symmetric.

Definition 5.25. *Let Σ be a surface in $\mathbb{R}^3$ with orientation $\boldsymbol{n}$ and let $\gamma \in C^2([-1,1],\Sigma)$ be a regular curve such that $\gamma(0) = p \in \Sigma$. Assume also that $\ddot{\gamma}(0) \neq 0$ and write $\cos\theta = \langle \boldsymbol{n}(p), \boldsymbol{n}_\gamma(0)\rangle$, where $\boldsymbol{n}_\gamma(0)$ denotes the normal vector of γ at 0, i.e., $\ddot{\gamma}(0) = k(0)\,\boldsymbol{n}_\gamma(0)$. Then we define the* ***normal curvature*** *$k_{\boldsymbol{n}}(p)$ of γ in Σ at p by*

$$k_{\boldsymbol{n}}(p) := k(0)\,\cos\theta = k(0)\,\langle \boldsymbol{n}(p), \boldsymbol{n}_\gamma(0)\rangle\,. \tag{5.35}$$

In other words, the normal curvature $k_{\boldsymbol{n}}$ is the length of the projection of the vector $k\,\boldsymbol{n}_\gamma$ on the normal $\boldsymbol{n}$ of the surface (see Fig. 5.6).

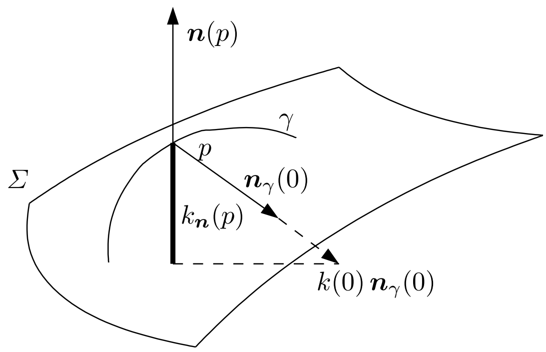

Fig. 5.6. Normal curvature of a curve in a surface.

To give an interpretation of the second fundamental form II_p, we consider a regular curve $\gamma \in C^2([-1,1],\Sigma)$ parameterized by arc length (seen as a curve in $\mathbb{R}^3$) such that $\gamma(0) = p \in \Sigma$. Denoting by $\boldsymbol{n}(s)$ the restriction of $\boldsymbol{n}$ to the curve $\gamma(s)$, we deduce that $\langle \boldsymbol{n}(s), \dot{\gamma}(s)\rangle = 0$. Hence, it follows

$$\langle \dot{\boldsymbol{n}}(s), \dot\gamma(s)\rangle + \langle \boldsymbol{n}(s), \ddot\gamma(s)\rangle = 0\,. \tag{5.36}$$

Note also that $\boldsymbol{dn}_p \cdot \dot\gamma(s) = \dot{\boldsymbol{n}}(s)$. Then we obtain for the second fundamental form

$$\begin{aligned} II_p\big(\dot\gamma(0), \dot\gamma(0)\big) &= -\langle \boldsymbol{dn}_p \cdot \dot\gamma(0), \dot\gamma(0)\rangle \\ &= -\langle \dot{\boldsymbol{n}}(0), \dot\gamma(0)\rangle \overset{(5.36)}{=} \langle \boldsymbol{n}(0), \ddot\gamma(0)\rangle \\ &= \langle \boldsymbol{n}(p), k(0)\,\boldsymbol{n}_\gamma(0)\rangle = k_{\boldsymbol{n}}(p)\,. \end{aligned}$$

Thus the value of the second fundamental form II_p for a unit vector $\dot\gamma(0) \in T_p\Sigma$ equals the normal curvature of the regular curve γ parameterized by arc length passing through p with tangent vector $\dot\gamma(0)$ at p. Moreover, we have

Proposition 5.26 (Meusnier). *The value of the second fundamental form II_p for a unit vector $X \in T_p\Sigma$ equals the normal curvature of any regular curve in Σ passing through p with tangent vector X at p.*

We come back to the linear map $\boldsymbol{dn}_p$. – Due to Proposition 5.23, the map $\boldsymbol{dn}_p$ is self-adjoint with respect to the first fundamental form $\langle\cdot,\cdot\rangle_p$. Hence, from a well-known result of linear algebra, there exists for all $p \in T_p\Sigma$ an orthonormal basis $\{e_1, e_2\}$ of $T_p\Sigma$ such that

$$\boldsymbol{dn}_p \cdot e_1 = -k_1\, e_1\,, \qquad \boldsymbol{dn}_p \cdot e_2 = -k_2\, e_2\,.$$

Assuming that $k_2 \geq k_1$ for the eigenvalues, we have also that $II_p(X,X) = -\langle \boldsymbol{dn}_p \cdot X, X\rangle \in [k_1, k_2]$, for all $X \in T_p\Sigma$ with $\|X\| = 1$.

Definition 5.27. *The maximum normal curvature k_2 and the minimum normal curvature k_1 are called the **principal curvatures** at $p \in \Sigma$. Moreover, the corresponding eigenvectors e_1 and e_2 are called the **principal directions** at p.*

Example 5.28.

Definition 5.29. *If for $p \in \Sigma$, we have that $k_1 = k_2$, then p is called an **umbilical point** of Σ.*

We prove now the interesting fact that the only surfaces made up entirely of umbilical points are essentially spheres and planes.

Theorem 5.30. *Let Σ be a connected surface in $\mathbb{R}^3$. Assume also that all points of Σ are umbilical points. Then the surface Σ is either contained in a sphere or in a plane.*

Proof. Let (U, φ) be a chart about $p \in \Sigma$. We first prove the theorem for $\Sigma|_U$.

Since each $p \in U$ is an umbilical point by assumption, we have that, for all $X \in T_p\Sigma$,

$$d\boldsymbol{n}_p \cdot X = \lambda(p)\, X\,, \tag{5.37}$$

where $\lambda : U \longrightarrow \mathbb{R}$ is a real function.

We show that λ is constant on U. – For this purpose, we write

$$X = X_1 \frac{\partial}{\partial \varphi_1} + X_2 \frac{\partial}{\partial \varphi_2} \in T_p\Sigma\,,$$

or, seeing X as a vector in $\mathbb{R}^3$,

$$X = X_1 \frac{\partial \tilde{\varphi}^{-1}}{\partial x_1}\big(\varphi(p)\big) + X_2 \frac{\partial \tilde{\varphi}^{-1}}{\partial x_2}\big(\varphi(p)\big)\,,$$

where the map $\tilde{\varphi}^{-1} = \iota_\Sigma \circ \varphi^{-1}$ was defined in (5.19). (Recall that $d\iota_\Sigma \cdot \frac{\partial}{\partial \varphi_i} = \frac{\partial \tilde{\varphi}^{-1}}{\partial x_i}$, for $i = 1, 2$.) By linearity, Equation (5.37) then becomes

$$\begin{aligned} &X_1\, d\boldsymbol{n}_p \cdot \frac{\partial \tilde{\varphi}^{-1}}{\partial x_1}\big(\varphi(p)\big) + X_2\, d\boldsymbol{n}_p \cdot \frac{\partial \tilde{\varphi}^{-1}}{\partial x_2}\big(\varphi(p)\big) \\ &= \lambda(p) \left(X_1 \frac{\partial \tilde{\varphi}^{-1}}{\partial x_1}\big(\varphi(p)\big) + X_2 \frac{\partial \tilde{\varphi}^{-1}}{\partial x_2}\big(\varphi(p)\big) \right). \end{aligned}$$

Defining $\tilde{\boldsymbol{n}} := \boldsymbol{n} \circ \tilde{\varphi}^{-1} : \varphi(\Sigma \cap U) \subset \mathbb{R}^2 \longrightarrow S^2 \subset \mathbb{R}^3$ (see also (5.27)), we get

$$X_1 \frac{\partial \tilde{\boldsymbol{n}}}{\partial x_1}\big(\varphi(p)\big) + X_2 \frac{\partial \tilde{\boldsymbol{n}}}{\partial x_2}\big(\varphi(p)\big) = \lambda(p) \left(X_1 \frac{\partial \tilde{\varphi}^{-1}}{\partial x_1}\big(\varphi(p)\big) + X_2 \frac{\partial \tilde{\varphi}^{-1}}{\partial x_2}\big(\varphi(p)\big) \right).$$

Hence, since X is arbitrary, it follows that

$$\begin{aligned} \frac{\partial \tilde{\boldsymbol{n}}}{\partial x_1}\big(\varphi(p)\big) &= \lambda(p) \frac{\partial \tilde{\varphi}^{-1}}{\partial x_1}\big(\varphi(p)\big) \\ \frac{\partial \tilde{\boldsymbol{n}}}{\partial x_2}\big(\varphi(p)\big) &= \lambda(p) \frac{\partial \tilde{\varphi}^{-1}}{\partial x_2}\big(\varphi(p)\big)\,. \end{aligned} \tag{5.38}$$

Differentiating the first equation with respect to x^2, the second one with respect to x^1 and subtracting the resulting equations, we obtain (check that $\lambda \in C^1(U, \mathbb{R})$)

$$\frac{\partial \lambda}{\partial x_2}(p)\, \frac{\partial \tilde{\varphi}^{-1}}{\partial x_1}\big(\varphi(p)\big) - \frac{\partial \lambda}{\partial x_1}(p)\, \frac{\partial \tilde{\varphi}^{-1}}{\partial x_2}\big(\varphi(p)\big) = 0\,.$$

Since $\frac{\partial \tilde{\varphi}^{-1}}{\partial x_1}$ and $\frac{\partial \tilde{\varphi}^{-1}}{\partial x_2}$ are linearly independent vectors in $T_p\Sigma \subset \mathbb{R}^3$, we conclude that, for all $p \in U$,

$$\frac{\partial \lambda}{\partial x_1}(p) = \frac{\partial \lambda}{\partial x_2}(p) = 0\,. \tag{5.39}$$

Since U is connected, it follows that $\lambda \equiv \lambda_0$ a constant in U, as we claimed.

If we assume that $\lambda_0 = 0$, Equation (5.38) shows that

$$\frac{\partial \tilde{\boldsymbol{n}}}{\partial x_1}\big(\varphi(p)\big) = \frac{\partial \tilde{\boldsymbol{n}}}{\partial x_2}\big(\varphi(p)\big) = 0\,,$$

and therefore $\tilde{\boldsymbol{n}}\big(\varphi(p)\big) = \boldsymbol{n}_0 \in \mathbb{R}^3$ in U; on the other hand, for $i = 1, 2$, we have

$$\frac{\partial}{\partial x_i}\big\langle \tilde{\varphi}^{-1}\big(\varphi(p)\big), \boldsymbol{n}_0\big\rangle = \left\langle \frac{\partial \tilde{\varphi}^{-1}}{\partial x_i}\big(\varphi(p)\big), \boldsymbol{n}_0 \right\rangle = 0\,.$$

Thus, we deduce

$$\big\langle \tilde{\varphi}^{-1}\big(\varphi(p)\big), \boldsymbol{n}_0\big\rangle = \text{const}\,, \tag{5.40}$$

showing that $\tilde{\varphi}^{-1}\big(\varphi(p)\big)$ belongs to a plane for all $p \in U$.

Assuming that $\lambda_0 \neq 0$, we obtain from (5.38) that

$$\frac{\partial}{\partial x_1}\left(\tilde{\varphi}^{-1}\big(\varphi(p)\big) - \frac{1}{\lambda_0}\,\tilde{\boldsymbol{n}}\big(\varphi(p)\big)\right) = 0\,,$$
$$\frac{\partial}{\partial x_2}\left(\tilde{\varphi}^{-1}\big(\varphi(p)\big) - \frac{1}{\lambda_0}\,\tilde{\boldsymbol{n}}\big(\varphi(p)\big)\right) = 0\,.$$

Hence there exists $Z_0 \in \mathbb{R}^3$ such that, for all $p \in U$,

$$\tilde{\varphi}^{-1}\big(\varphi(p)\big) - \frac{1}{\lambda_0}\,\tilde{\boldsymbol{n}}\big(\varphi(p)\big) = Z_0\,,$$

implying also that

$$\|\tilde{\varphi}^{-1}\big(\varphi(p)\big) - Z_0\|^2 = \frac{1}{\lambda_0^2}\,. \tag{5.41}$$

In other words, all points of U are contained in a sphere of center Z_0 and radius $1/\lambda_0^2$.

So far we have shown that for every $p \in \Sigma$ there exists an open neighborhood U of p such that U is either included in a sphere or in a plane. – To complete the proof we observe that, since Σ is connected, given any points $p_1, p_2 \in \Sigma$, there exists a curve $\gamma \in C^0([0,1], \Sigma)$ with $\gamma(0) = p_1$ and $\gamma(1) = p_2$. For each point $p \in \gamma([0,1])$ of this curve the above implies the existence of open neighborhoods $U_p \subset \Sigma$ such that each one is contained in a sphere or a plane. By continuity, $\gamma^{-1}(U_p)$ is open in $\mathbb{R}$ and the union

$$\bigcup_{t \in [0,1]} \gamma^{-1}(U_p)$$

covers the closed and bounded interval $[0,1]$. Due to the Heine-Borel theorem, we can extract a finite covering; hence,

$$\gamma([0,1]) \subset \bigcup_{j=1}^{Q} U_{p_j}\,.$$

If the points of one of these neighborhoods are on a plane (sphere), all the others will be on the *same* plane (sphere). Since p_1 and p_2 have been chosen arbitrarily, all the points of Σ belong to this plane (sphere), completing the proof of the theorem. $\square$

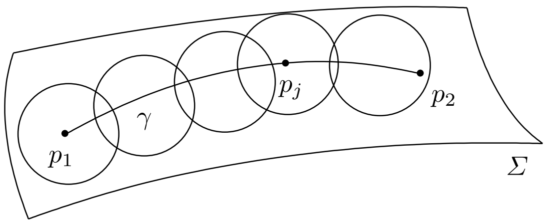

Fig. 5.7. .

Definition 5.31. *Let $p \in \Sigma$ and let $d\boldsymbol{n}_p : T_p\Sigma \longrightarrow T_p\Sigma$ be the differential of the Gauß map. The determinant of the linear map $d\boldsymbol{n}_p$, denoted by $K(p)$, is the **Gaussian curvature** of Σ at p, i.e., in terms of the principal directions*

$$K(p) = det(d\boldsymbol{n}_p) = k_1\, k_2 \,. \tag{5.42}$$

*Moreover, the negative of half of the trace of $d\boldsymbol{n}_p$, denoted by $H(p)$, is the **mean curvature** of Σ at p, i.e., in terms of the principal directions*

$$H(p) = \frac{1}{2}\, Tr(d\boldsymbol{n}_p) = \frac{k_1 + k_2}{2} \,. \tag{5.43}$$

Definition 5.32. *We characterize points $p \in \Sigma$ in the following way:*

1. *A point $p \in \Sigma$ is called **elliptic** if $K(p) > 0$.*
2. *A point $p \in \Sigma$ is called **hyperbolic** if $K(p) < 0$.*
3. *A point $p \in \Sigma$ is called **parabolic** if $K(p) = 0$ and $d\boldsymbol{n}_p \neq 0$.*
4. *A point $p \in \Sigma$ is called **planar** if $d\boldsymbol{n}_p = 0$.*

Let (U, φ) be a local chart about $p \in \Sigma$. As in the case of the first fundamental form I_p, we introduce some useful notations for the coefficients of the second fundamental form II_p in the local chart (U, φ). More precisely, for $X, Y \in T_p\Sigma$ and with the notations

$$\begin{aligned}
e(p) &= -\left\langle d\boldsymbol{n}_p \cdot \frac{\partial}{\partial \varphi_1}(p), \frac{\partial}{\partial \varphi_1}(p) \right\rangle_p \,, \\
f(p) &= -\left\langle d\boldsymbol{n}_p \cdot \frac{\partial}{\partial \varphi_1}(p), \frac{\partial}{\partial \varphi_2}(p) \right\rangle_p = \left\langle d\boldsymbol{n}_p \cdot \frac{\partial}{\partial \varphi_2}(p), \frac{\partial}{\partial \varphi_1}(p) \right\rangle_p \,, \\
g(p) &= -\left\langle d\boldsymbol{n}_p \cdot \frac{\partial}{\partial \varphi_2}(p), \frac{\partial}{\partial \varphi_2}(p) \right\rangle_p \,,
\end{aligned} \tag{5.44}$$

the second fundamental form reads as

$$II_p(X,Y) = -\langle d\boldsymbol{n}_p \cdot X, Y\rangle_p = X_1Y_1\, e(p) + X_1Y_2\, f(p) + X_2Y_1\, f(p) + X_2Y_2\, g(p)\,. \tag{5.45}$$

Next, we introduce some notations for the differential of the Gauß map $d\boldsymbol{n}_p$ in the local chart (U,φ) about p. Since $d\boldsymbol{n}_p \cdot \frac{\partial}{\partial\varphi_i}(p)$ belongs to $T_p\Sigma$ for $i = 1,2$, we may write

$$d\boldsymbol{n}_p \cdot \frac{\partial}{\partial\varphi_i}(p) = \sum_{j=1}^{2} a_{ij}\, \frac{\partial}{\partial\varphi_j}(p)\,. \tag{5.46}$$

Inserting this into (5.44), we get

$$\begin{aligned} e(p) &= -\left\langle a_{11}\, \frac{\partial}{\partial\varphi_1}(p) + a_{12}\, \frac{\partial}{\partial\varphi_2}(p), \frac{\partial}{\partial\varphi_1}(p)\right\rangle_p \\ &= -a_{11}\, E(p) - a_{12}\, F(p)\,, \end{aligned}$$

where we used the notations (5.16) for the first fundamental form. Similarly, we obtain (note that the matrix (a_{ij}) is not necessarily symmetric)

$$\begin{aligned} f(p) &= -a_{11}\, F(p) - a_{12}\, G(p)\,, \\ f(p) &= -a_{21}\, E(p) - a_{22}\, F(p)\,, \\ g(p) &= -a_{21}\, F(p) - a_{22}\, G(p)\,. \end{aligned}$$

In summary, we then arrive at

$$-\begin{pmatrix} e & f \\ f & g \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix}\,, \tag{5.47}$$

and hence

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = -\begin{pmatrix} e & f \\ f & g \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1}\,.$$

It is well-known from linear algebra that the inverse in the last equation reads as

$$\begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1} = a \begin{pmatrix} G & -F \\ -F & E \end{pmatrix}\,,$$

where $a := 1/(E\,G - F^2)$ for the determinant of the matrix formed by the local coefficients of the first fundamental form. By a straightforward computation, we then obtain the *equations of Weingarten*:

$$\begin{aligned} a_{11} &= a(f\,F - e\,G)\,, \qquad & a_{12} &= a(g\,F - f\,G)\,, \\ a_{21} &= a(e\,F - f\,E)\,, \qquad & a_{22} &= a(f\,F - g\,E)\,. \end{aligned} \tag{5.48}$$

For the Gaussian curvature $K(p)$ of a surface at a point p, we then easily get

$$K(p) = \det\bigl(a_{ij}(p)\bigr) = a(p)\bigl(e(p)\,g(p) - f^2(p)\bigr)\,, \tag{5.49}$$

and for the mean curvature $H(p)$, one checks that

$$H(p) = \frac{a(p)}{2}\,\bigl(e(p)\,G(p) - 2f(p)\,F(p) + g(p)\,E(p)\bigr)\,. \tag{5.50}$$

Example 5.33 (Gaussian Curvature of the Torus).

The following proposition gives information about the position of a surface in the neighborhood of an elliptic or an hyperbolic point relative to the tangent space at this point.

Proposition 5.34. *Let $p \in \Sigma$ be an elliptic point of the surface Σ. Then there exists a neighborhood U of p in Σ such that all points in U belong to the same side of the tangent space $T_p\Sigma$. Moreover, let $p \in \Sigma$ be an hyperbolic point of Σ. Then in each neighborhood U of p there are points of U lying in both sides of $T_p\Sigma$.*

Proof. Let (U,φ) be a local chart about $p \in \Sigma$ and the local parameterization $\tilde{\varphi}^{-1} = \iota_\Sigma \circ \varphi^{-1}$ as defined in (5.19). Moreover, for $(0,0) \in \mathbb{R}^2$ we assume that $\tilde{\varphi}^{-1}(0,0) = p$.

The distance d from a point $q = \tilde{\varphi}^{-1}(x_1,x_2) \in U$ to the tangent space $T_p\Sigma$ is given by (see Fig. 5.8)

$$d = \langle \boldsymbol{n}_p, \tilde{\varphi}^{-1}(x_1,x_2) - \tilde{\varphi}^{-1}(0,0)\rangle\,. \tag{5.51}$$

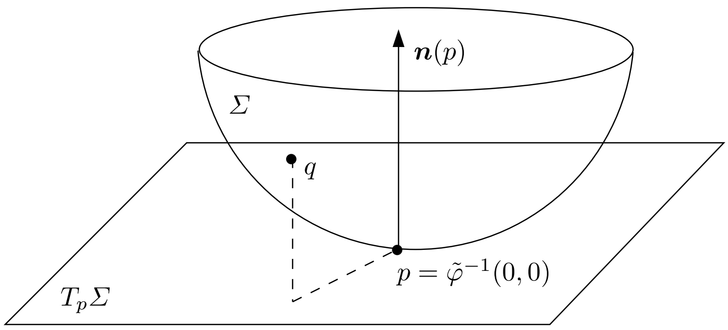

Fig. 5.8. Tangent space of an elliptic point.

Since the local parameterization $\tilde{\varphi}^{-1} : \varphi(\Sigma \cap U) \subset \mathbb{R}^2 \longrightarrow \Sigma \cap U \subset \mathbb{R}^3$ is differentiable, we can apply Taylor's formula to obtain (up to terms of higher order)

$$\tilde{\varphi}^{-1}(x_1,x_2) = \tilde{\varphi}^{-1}(0,0) + x_1 \frac{\partial\tilde{\varphi}^{-1}}{\partial x_1}(0,0) + x_2 \frac{\partial\tilde{\varphi}^{-1}}{\partial x_2}(0,0)$$
$$+\frac{1}{2}\left(x_1^2 \frac{\partial^2\tilde{\varphi}^{-1}}{\partial x_1^2}(0,0) + 2x_1x_2 \frac{\partial^2\tilde{\varphi}^{-1}}{\partial x_1 \partial x_2}(0,0) + x_2^2 \frac{\partial^2\tilde{\varphi}^{-1}}{\partial x_2^2}(0,0)\right).$$

Inserting this into (5.51) implies that

$$d = \frac{1}{2}\left(x_1^2 \left\langle \boldsymbol{n}_p, \frac{\partial^2\tilde{\varphi}^{-1}}{\partial x_1^2}(0,0)\right\rangle + 2x_1x_2 \left\langle \boldsymbol{n}_p, \frac{\partial^2\tilde{\varphi}^{-1}}{\partial x_1 \partial x_2}(0,0)\right\rangle \right.$$
$$\left. +x_2^2 \left\langle \boldsymbol{n}_p, \frac{\partial^2\tilde{\varphi}^{-1}}{\partial x_2^2}(0,0)\right\rangle\right). \tag{5.52}$$

On the other hand, we compute, for the map $\tilde{\boldsymbol{n}} := \boldsymbol{n}\circ\tilde{\varphi}^{-1} : \varphi(\Sigma \cap U) \longrightarrow S^2 \subset \mathbb{R}^3$,

$$0 = \frac{\partial}{\partial x_i}\left\langle \tilde{\boldsymbol{n}}(x_1,x_2), \frac{\partial\tilde{\varphi}^{-1}}{\partial x_j}(x_1,x_2)\right\rangle$$
$$= \left\langle \frac{\partial\tilde{\boldsymbol{n}}}{\partial x_i}(x_1,x_2), \frac{\partial\tilde{\varphi}^{-1}}{\partial x_j}(x_1,x_2)\right\rangle + \left\langle \tilde{\boldsymbol{n}}(x_1,x_2), \frac{\partial^2\tilde{\varphi}^{-1}}{\partial x_i \partial x_j}(x_1,x_2)\right\rangle,$$

where $i,j = 1,2$. Using Definition 5.24 and recalling that $d\boldsymbol{n}\cdot\frac{\partial}{\partial\varphi_i} = \frac{\partial\tilde{\varphi}^{-1}}{\partial x_i}$, for $i = 1,2$, we deduce (note that $\tilde{\boldsymbol{n}}(0,0) = \boldsymbol{n}_p$)

$$\left\langle \boldsymbol{n}_p, \frac{\partial^2\tilde{\varphi}^{-1}}{\partial x_i \partial x_j}(0,0)\right\rangle = -\left\langle d\boldsymbol{n}_p \cdot \frac{\partial}{\partial\varphi_i}(p), \frac{\partial}{\partial\varphi_j}(p)\right\rangle_p = II_p\left(\frac{\partial}{\partial\varphi_i}(p), \frac{\partial}{\partial\varphi_j}(p)\right). \tag{5.53}$$

Inserting (5.53) into (5.52) leads to

$$d = \frac{1}{2}\, II_p\,(X,X)\,,$$

where $X = X_1 \frac{\partial\tilde{\varphi}^{-1}}{\partial x_1}(0,0) + X_2 \frac{\partial\tilde{\varphi}^{-1}}{\partial x_2}(0,0)$.

Since p is elliptic by assumption, the second fundamental form II_p has a fixed sign. Therefore, the same holds for the distance d, showing that all points $q \in U$ belong to the same side of $T_p\Sigma$.

The second assertion of the proposition is left as an exercise. □

Now, we give a geometric interpretation of the Gaussian curvature (see Fig. 5.9).

Proposition 5.35. *Let p be a point of a surface Σ with $K(p) \neq 0$. Denote by $(U_n)_{n\in\mathbb{N}}$ a sequence of neighborhoods of p converging to the point p for n sufficiently large, i.e., $U_n \subset B_{r_n}(p)$ such that the radii r_n of $B_{r_n}(p)$ converge to zero for $n \to \infty$. Moreover, we denote by $\bar{A}_n := A\big(\boldsymbol{n}(U_n)\big)$ the area of the image of U_n by the Gauß map $\boldsymbol{n} : \Sigma \longrightarrow S^2$ and by $A_n := A(U_n)$ the area of $U_n \subset \Sigma$. Then, we have for the Gaussian curvature at p*

$$K(p) = \lim_{n\longrightarrow\infty} \frac{\bar{A}_n}{A_n} = \lim_{n\longrightarrow\infty} \frac{A\big(\boldsymbol{n}(U_n)\big)}{A(U_n)}\,. \tag{5.54}$$

Proof. Let (U,φ) be a local chart for Σ about p such that $U_n \subset U$, for all $n \in \mathbb{N}$. From the discussion in Section 4.2.1, we deduce that

$$A_n = A(U_n) = \int_{\varphi(U_n)} \left\| \frac{\partial\tilde{\varphi}^{-1}}{\partial x_1}(x_1,x_2) \times \frac{\partial\tilde{\varphi}^{-1}}{\partial x_2}(x_1,x_2)\right\| \, dx_1 dx_2\,,$$

where $\tilde{\varphi}^{-1} := \iota_\Sigma \circ \varphi^{-1}$ denotes again the local parameterization of Σ.

On the other hand, since $K(p) = \det(d\boldsymbol{n}_p) \neq 0$ by assumption (implying the invertibility of the linear map $d\boldsymbol{n}_p$), we deduce that the map

$$\tilde{\boldsymbol{n}} := \boldsymbol{n} \circ \tilde{\varphi}^{-1} : \varphi(\Sigma \cap U) \longrightarrow S^2 \subset \mathbb{R}^3$$

gives a local parameterization of $\boldsymbol{n}(U)$. Hence, it follows

$$\bar{A}_n = A\big(\boldsymbol{n}(U_n)\big) = \int_{\varphi(U_n)} \left\| \frac{\partial\tilde{\boldsymbol{n}}}{\partial x_1}(x_1,x_2) \times \frac{\partial\tilde{\boldsymbol{n}}}{\partial x_2}(x_1,x_2)\right\| \, dx_1 dx_2\,.$$

For $i = 1,2$ and $\varphi(p) = (x_1,x_2)$, we now compute (see (5.46))

$$\frac{\partial\tilde{\boldsymbol{n}}}{\partial x_i}(x_1,x_2) = d\boldsymbol{n}_p \cdot \frac{\partial\tilde{\varphi}^{-1}}{\partial x_i}(x_1,x_2) = \sum_{j=1}^{2} a_{ij}\, \frac{\partial\tilde{\varphi}^{-1}}{\partial x_j}(x_1,x_2)\,.$$

We then see that

$$\frac{\partial\tilde{\boldsymbol{n}}}{\partial x_1}(x_1,x_2) \times \frac{\partial\tilde{\boldsymbol{n}}}{\partial x_2}(x_1,x_2) = \det(a_{ij})\, \frac{\partial\tilde{\varphi}^{-1}}{\partial x_1}(x_1,x_2) \times \frac{\partial\tilde{\varphi}^{-1}}{\partial x_2}(x_1,x_2)\,.$$

Since $K(p) = \det(a_{ij})$, the result (5.54) can be deduced. $\square$

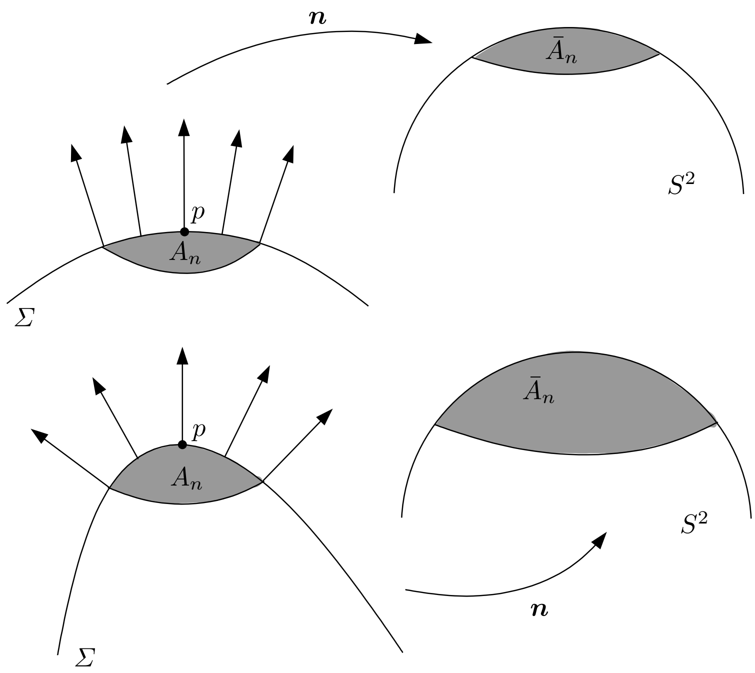

Fig. 5.9. Geometric interpretation of the Gaussian curvature. The first figure illustrates the case of small Gaussian curvature, whereas the second the case of big Gaussian curvature.

6

6.1 Tensor algebra

We consider finite dimensional vector spaces over $\mathcal{F}$. In this section, the ground field $\mathcal{F}$ will be the real number field $\mathbb{R}$ or the complex number field $\mathbb{C}$. Let E and F be two such vector spaces and let $M(E,F)$ be the vector space having the set $E \times F$ as a basis, i.e., the free vector space generated by the pairs (e,f) with $e \in E$ and $f \in F$. Moreover, let $N(E,F)$ be the vector subspace of $M(E,F)$ spanned by the elements of the form

$$\begin{aligned}(e+e',f)-(e,f)-(e',f)\,, &\qquad (e,f+f')-(e,f)-(e,f') \\ (\lambda\, e,f)-\lambda(e,f)\,, &\qquad (e,\lambda\, f)-\lambda(e,f)\,, \end{aligned} \tag{6.1}$$

where $e,e' \in E$, $f,f' \in F$ and $\lambda \in \mathcal{F}$.

Definition 6.1. *Let E and F be two finite dimensional $\mathcal{F}$-vector spaces. Moreover, let $M(E,F)$ and $N(E,F)$ be as above. We define the* ***tensor product*** *of E and F, denoted by $E \otimes F$, to be the quotient vector space $M(E,F)/N(E,F)$.*

The image by the natural projection $M(E,F) \longrightarrow E \otimes F$ of every pair (e,f) – considered as element of $M(E,F)$ – will be denoted by $e \otimes f$. Then, we define the canonical map

$$\begin{aligned} P : E \times F &\longrightarrow E \otimes F \\ (e,f) &\longmapsto e \otimes f\,. \end{aligned} \tag{6.2}$$

It can easily be checked that this map is *bilinear*, since the first relation in (6.1), for example, implies that $(e+e') \otimes f = e \otimes f + e' \otimes f$.

Definition 6.2. *Let G be a vector space and $\psi : E \times F \longrightarrow G$ a bilinear map. We say that the couple (G,ψ) has the* ***universal factorization property*** *for $E \times F$ if for every vector space H and every bilinear map $\varphi : E \times F \longrightarrow H$, there exists a unique linear map $u : G \longrightarrow H$ such that $\varphi = u \circ \psi$.*

There is a first important result.

$$\begin{tikzcd} E \times F \arrow[rr, "\psi"] \arrow[drr, "\varphi"'] & & G \arrow[d, "u"] \\ & & H \end{tikzcd}$$

Fig. 6.1. Universal factorization property.

Proposition 6.3. *The couple* $(E \otimes F, P)$ *has the universal factorization property for* $E \times F$*. Moreover, if another couple* (G, ψ) *has the universal factorization property for* $E \times F$*, then* $(E \otimes F, P)$ *and* (G, ψ) *are isomorphic in the sense that there exists a unique isomorphism* $g : E \otimes F \longrightarrow G$ *such that* $\psi = g \circ P$*.*

Proof. Let H be any vector space and $\varphi : E \times F \longrightarrow H$ any bilinear map. Since $E \times F$ is a basis for $M(E, F)$, we can extend φ to a unique linear map $\tilde{\varphi} : M(E, F) \longrightarrow H$. More precisely, writing $m \in M(E, F)$ as $m = \sum_i \lambda_i\,(e_i, f_i)$, with $(e_i, f_i) \in E \times F$ and $\lambda_i \in \mathcal{F}$, we define

$$\tilde{\varphi}(m) = \tilde{\varphi}\left(\sum_i \lambda_i\,(e_i, f_i)\right) = \sum_i \lambda_i\,\varphi(e_i, f_i)\,.$$

Using the bilinearity of the map φ, we observe that

$$\tilde{\varphi}\big((e + e', f) - (e, f) - (e', f)\big) = \varphi(e + e', f) - \varphi(e, f) - \varphi(e', f) = 0\,.$$

The same holds for the other relations in (6.1). Hence, we deduce that $\tilde{\varphi}$ vanishes on $N(E, F)$. Therefore, the map $\tilde{\varphi}$ induces a linear map[1]

$$u : E \otimes F \longrightarrow H\,.$$

Clearly, we have that $\varphi = u \circ P$. The uniqueness of such a map u follows from the fact that $P(E \times F)$ spans $E \otimes F$ (see (6.2)). – This shows that the couple $(E \otimes F, P)$ has the universal factorization property for $E \times F$.

Let (G, ψ) be another couple having the universal factorization property for $E \times F$. By the universal factorization property of $(E \otimes F, P)$, respectively of (G, ψ), there exists a unique linear map $g : E \otimes F \longrightarrow G$, respectively $\tilde{g} : G \longrightarrow E \otimes F$ such that $\psi = g \circ P$, respectively $P = \tilde{g} \circ \psi$ (see Fig. 6.3). Hence, we obtain that $\psi = g \circ \tilde{g} \circ \psi$ and $P = \tilde{g} \circ g \circ P$. Using the uniqueness of the linear map u in the definition of the universal factorization property, we

[1]

conclude that $\tilde{g} \circ g$ and $g \circ \tilde{g}$ are the identity transformations of $E \otimes F$ and G, respectively. – This shows that $g : E \otimes F \longrightarrow G$ is the desired isomorphism.

$\square$

$$
\begin{tikzcd}
E \times F \arrow[r, "P"] \arrow[dr, "\varphi"'] & E \otimes F \arrow[d, "u"] \\
 & H
\end{tikzcd}
$$

Fig. 6.2. Universal factorization property of $(E \otimes F, P)$ for $E \times F$.

$$
\begin{tikzcd}
E \times F \arrow[r, "\psi"] \arrow[dr, "P"'] & G \arrow[d, "\tilde{g}"] & \\
 & E \otimes F \arrow[d, "g"] & E \times F \arrow[l, "P"'] \arrow[dl, "\psi"] \\
 & G &
\end{tikzcd}
$$

Fig. 6.3. Two couples $(E \otimes F, P)$ and (G, ψ) having the universal factorization property for $E \times F$.

Remark. It is important to note that this proposition will be the main ingredient for the following results concerning the properties of the tensor product. The proofs of Proposition 6.4 to 6.11 will be strongly based on the fact that $(E \otimes F, P)$ has the universal factorization property for $E \times F$.

We have commutativity for the tensor product.

Proposition 6.4. *There exists a unique isomorphism between $E \otimes F$ and $F \otimes E$ which sends $e \otimes f$ onto $f \otimes e$., for all $e \in E$ and $f \in F$.*

Proof. Let φ be the bilinear map defined by

$$\begin{aligned} \varphi : E \times F &\longrightarrow F \otimes E\,, \\ (e, f) &\longmapsto f \otimes e\,. \end{aligned}$$

From Proposition 6.3 we deduce the existence of a unique linear map $u : E \otimes F \longrightarrow F \otimes E$ such that $\varphi = u \circ P$. More precisely, we have that

$$f \otimes e = \varphi(e, f) = u\big(P(e, f)\big) = u(e \otimes f)\,.$$

On the other hand, note that the couple $(F \otimes E, \tilde{P})$, where (see (6.2))

$$\begin{aligned} \tilde{P} : F \times E &\longrightarrow F \otimes E\,, \\ (f, e) &\longmapsto f \otimes e\,, \end{aligned}$$

has the universal factorization property for $F \times E$. So, for the bilinear map

$$\begin{aligned} \tilde{\varphi} : F \times E &\longrightarrow E \otimes F\,, \\ (f, e) &\longmapsto e \otimes f\,, \end{aligned}$$

we obtain the existence of a unique linear map $\tilde{u} : F \otimes E \longrightarrow E \otimes F$ such that $\tilde{u}(f \otimes e) = e \otimes f$. This is again a consequence of the universal factorization property.

Obviously, we see that $\tilde{u} \circ u : E \otimes F \longrightarrow E \otimes F$ and $u \circ \tilde{u} : F \otimes E \longrightarrow F \otimes E$ are the identity transformations. Hence, the linear map u is the desired isomorphism. □

We also have associativity for the tensor product.

Proposition 6.5. *There exists a unique isomorphism between $(E \otimes F) \otimes G$ and $E \otimes (F \otimes G)$ which sends $(e \otimes f) \otimes g$ onto $e \otimes (f \otimes g)$, for all $e \in E$, $f \in F$ and $g \in G$.*

Proof. The proof is similar to the one of Proposition 6.4 and hence left as an exercise. □

Remark. Because of the last proposition, we identify $(E \otimes F) \otimes G$ with $E \otimes (F \otimes G)$ and it makes sense to write $E \otimes F \otimes G$.

Now, let $E_1, \ldots, E_k$ be k vector spaces. Then we can define inductively the tensor product $E_1 \otimes \ldots \otimes E_k$. More precisely, motivated by Proposition 6.3, the tensor product $E_1 \otimes \ldots \otimes E_k$ is characterized by the fact that $(E_1 \otimes \ldots \otimes E_k, P)$, where P now denotes the multilinear map

$$P : E_1 \times \ldots \times E_k \longrightarrow E_1 \otimes \ldots \otimes E_k \,,$$
$$(e_1, \ldots, e_k) \longmapsto e_1 \otimes \ldots \otimes e_k \,,$$

has the universal factorization property for $E_1 \times \ldots \times E_k$. – We mention that Proposition 6.4 can also be generalized to k vector spaces.

Proposition 6.6. *For any permutation $\sigma \in \mathcal{S}_k$ of $\{1, \ldots, k\}$, there exists a unique isomorphism between $E_1 \otimes \ldots \otimes E_k$ and $E_{\sigma(1)} \otimes \ldots \otimes E_{\sigma(k)}$ which sends $e_1 \otimes \ldots \otimes e_k$ onto $e_{\sigma(1)} \otimes \ldots \otimes e_{\sigma(k)}$, for all $e_1 \in E_1$, …, $e_k \in E_k$.*

Next, we study the compatibility of the tensor product with linear maps.

Proposition 6.7. *Let $u_i : E_i \longrightarrow F_i$, $i = 1, 2$, be two linear maps. Then there exists a unique linear map $u : E_1 \otimes E_2 \longrightarrow F_1 \otimes F_2$ such that*

$$u(e_1 \otimes e_2) = u_1(e_1) \otimes u_2(e_2) \in F_1 \otimes F_2 \,,$$

for all $e_1 \in E_1$ and $e_2 \in E_2$. – In the following, the map u will be denoted by $u_1 \otimes u_2$.

Proof. We consider the bilinear map

$$\varphi : E_1 \times E_2 \longrightarrow F_1 \otimes F_2 \,,$$
$$(e_1, e_2) \longmapsto u_1(e_1) \otimes u_2(e_2) \,.$$

From the universal factorization property in Proposition 6.3, we know that there exists a unique linear map $u : E_1 \otimes E_2 \longrightarrow F_1 \otimes F_2$ such that $\varphi = u \circ P$. Hence, we have that

$$u_1(e_1) \otimes u_2(e_2) = \varphi(e_1, e_2) = u \circ P(e_1, e_2) = u(e_1 \otimes e_2) \,.$$

This shows the result. □

Remark. The generalization of this proposition to the case with more than two mappings is obvious.

The next result is a consequence of the last proposition.

Proposition 6.8. *Let $E_1 \oplus E_2$ denote the direct sum between E_1 and E_2. Then there exists a unique isomorphism between $(E_1 \oplus E_2) \otimes F$ and $(E_1 \otimes F) \oplus (E_2 \otimes F)$ which sends $(e_1 + e_2) \otimes f$ onto $e_1 \otimes f + e_2 \otimes f$, for all $e_1 \in E_1$, $e_2 \in E_2$ and $f \in F$. Similarly, there exists a unique isomorphism between $E \otimes (F_1 \oplus F_2)$ and $E \otimes F_1 \oplus E \otimes F_2$.*

Proof. For $i = 1, 2$, let $\iota_i : E_i \hookrightarrow E_1 \oplus E_2$ denote the injections with $\iota_1(e_1) = e_1 + 0$ and $\iota_2(e_2) = 0 + e_2$, respectively. Let also $\pi_i : E_1 \oplus E_2 \longrightarrow E_i$ denote the projections with $\pi_1(e_1 + e_2) = e_1$ and $\pi_2(e_1 + e_2) = e_2$, respectively. (These maps are well-defined since we consider the direct sum of E_1 and E_2.) Then, we see that $\pi_1 \circ \iota_1$ and $\pi_2 \circ \iota_2$ are the identity transformations of E_1

and E_2, respectively. Moreover, both $\pi_2 \circ \iota_1$ and $\pi_1 \circ \iota_2$ are the zero maps on E_1 and E_2, respectively.

Due to Proposition 6.7 the injection ι_1 and the identity transformation id_F on F induce a unique linear map $\iota_1 \otimes id_F : E_1 \otimes F \longrightarrow (E_1 \oplus E_2) \otimes F$ such that $\iota_1 \otimes id_F(e_1 \otimes f) = \iota_1(e_1) \otimes f$. Similarly, a map $\iota_2 \otimes id_F : E_2 \otimes F \longrightarrow (E_1 \oplus E_2) \otimes F$ is obtained. – In order to simplify the notations we write $\tilde{\iota}_i$ instead of $\iota_i \otimes id_F$, for $i = 1, 2$.

In the same manner, applying Proposition 6.7 to the projection π_1 and the identity transformation on F it follows that there exists a unique linear map $\tilde{\pi}_1 : (E_1 \oplus E_2) \otimes F \longrightarrow E_1 \otimes F$ such that $\tilde{\pi}_1\big((e_1 + e_2) \otimes f\big) = \pi_1(e_1 + e_2) \otimes f$. Similarly, the map $\tilde{\pi}_2$ is defined.

As a consequence, $\tilde{\pi}_1 \circ \tilde{\iota}_1$ and $\tilde{\pi}_2 \circ \tilde{\iota}_2$ are the identity transformations on $E_1 \otimes F$ and $E_2 \otimes F$, respectively. For example, we have

$$\tilde{\pi}_1 \circ \tilde{\iota}_1(e_1 \otimes f) = \tilde{\pi}_1\big(\iota_1(e_1) \otimes f\big) = \pi_1\big(\iota_1(e_1)\big) \otimes f = e_1 \otimes f\,.$$

On the other hand, it follows from

$$\tilde{\pi}_2 \circ \tilde{\iota}_1(e_1 \otimes f) = \tilde{\pi}_2\big(\iota_1(e_1) \otimes f\big) = \pi_2\big(\iota_1(e_1)\big) \otimes f = 0 \otimes f$$

that $\tilde{\pi}_2 \circ \tilde{\iota}_1$ is the zero map on $E_1 \otimes F$; and similarly that $\tilde{\pi}_1 \circ \tilde{\iota}_2$ is the zero map on $E_2 \otimes F$.

These results imply that the linear map

$$\begin{aligned} (E_1 \oplus E_2) \otimes F &\longrightarrow (E_1 \otimes F) \oplus (E_2 \otimes F)\,, \\ (e_1 + e_2) \otimes f &\longmapsto \tilde{\pi}_1\big((e_1 + e_2) \otimes f\big) + \tilde{\pi}_2\big((e_1 + e_2) \otimes f\big)\,, \end{aligned}$$

with inverse

$$\begin{aligned} (E_1 \otimes F) \oplus (E_2 \otimes F) &\longrightarrow (E_1 \oplus E_2) \otimes F\,, \\ (e_1 \otimes f) + (e_2 \otimes f) &\longmapsto \tilde{\iota}_1(e_1 \otimes f) + \tilde{\iota}_2(e_2 \otimes f)\,, \end{aligned}$$

gives the first isomorphism. – The proof for the second one is similar. □

Remark. Now, we can write

$$(E_1 \oplus E_2) \otimes F = (E_1 \otimes F) \oplus (E_2 \otimes F)\,.$$

By induction, moreover, we obtain for k vector spaces that

$$(E_1 \oplus \ldots \oplus E_k) \otimes F = E_1 \otimes F \oplus \ldots \oplus E_k \otimes F\,.$$

In a next step, we give a basis for the tensor product of two vector spaces.

Proposition 6.9. *Let $\{e_i\}_{i=1,\ldots,n}$ be a basis for E and $\{f_j\}_{j=1,\ldots,m}$ be a basis for F. Then*

$$\{e_i \otimes f_j\}_{\substack{i=1,\ldots,n \\ j=1,\ldots,m}}$$

is a basis for $E \otimes F$. In particular, we have that $dim\, E \otimes F = dim\, E\; dim\, F$.

Proof. For $i = 1, \ldots, n$, let E_i denote the one-dimensional vector subspace of E spanned by e_i and F_j the one-dimensional vector subspace of F spanned by f_j, $j = 1, \ldots, m$. Thus, we can write

$$E = \bigoplus_{i=1}^{n} E_i \qquad \text{and} \qquad F = \bigoplus_{j=1}^{m} F_j \,.$$

By the remark after Proposition 6.8, it then follows that

$$E \otimes F = \left(\bigoplus_{i=1}^{n} E_i \right) \otimes \left(\bigoplus_{j=1}^{m} F_j \right) = \bigoplus_{i=1}^{n} \bigoplus_{j=1}^{m} E_i \otimes F_j \,.$$

It is left as an exercise to show that each $E_i \otimes F_j$ is a one-dimensional vector space spanned by $e_i \otimes f_j$. Hence, the proposition follows. □

Tensor Product and Dual Vector Space

For a vector space E, we denote by E^* the *dual* vector space of E, i.e., the vector space of linear functionals on E. If $e \in E$ and $e^* \in E^*$, then the value of e^* on e is denoted by $\langle e^*, e \rangle_{E^*,E} \in \mathcal{F}$. – We will simply write $\langle e^*, e \rangle$ when the dual pairing needs no further specifications.

Proposition 6.10. *Let $\mathcal{L}(E^*, F)$ denote the vector space of linear maps from E^* into F. Then there is a unique isomorphism u between $E \otimes F$ and $\mathcal{L}(E^*, F)$ such that*

$$\big(u(e \otimes f)\big)(e^*) = \langle e^*, e \rangle \, f \,, \tag{6.3}$$

for all $e \in E$, $f \in F$ and $e^* \in E^*$.

Proof. We consider the bilinear map

$$\begin{aligned} \varphi : E \times F &\longrightarrow \mathcal{L}(E^*, F) \,, \\ (e, f) &\longmapsto \varphi(e, f) \,, \end{aligned}$$

where $\big(\varphi(e, f)\big)(e^*) = \langle e^*, e \rangle \, f$. By the universal factorization property in Proposition 6.3, there exists a unique $u : E \otimes F \longrightarrow \mathcal{L}(E^*, F)$ such that $\varphi = u \circ P$. More precisely, we have that

$$\langle e^*, e \rangle \, f = \big(\varphi(e, f)\big)(e^*) = \big(u \circ P(e, f)\big)(e^*) = \big(u(e \otimes f)\big)(e^*) \,.$$

Next, we observe that $\dim(E \otimes F) = \dim \mathcal{L}(E^*, F)$ (see Proposition 6.9). Therefore, in order to show that u is an isomorphism, it suffices to establish that u is injective. – Let $\{e_i\}_{i=1,\ldots,n}$, $\{f_j\}_{j=1,\ldots,m}$ and $\{e^i\}_{i=1,\ldots,n}$ denote the basis for E, F and E^*, respectively[2]. Moreover, for some $e \otimes f \in E \otimes F$,

[2] From now on, we will always use this notation: Lower indices for the basis vectors of the vector space itself and upper indices for the dual basis.

assume that $u(e \otimes f) = 0$. From Proposition 6.9, we deduce the existence of coefficients $a_{ij} \in \mathcal{F}$ such that

$$u(e \otimes f) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}\, u(e_i \otimes f_j) = 0\,.$$

In other words, for all $k = 1, \ldots, n$, we have that

$$\sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} \big(u(e_i \otimes f_j)\big)(e^k) \stackrel{(6.3)}{=} \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}\, \langle e^k, e_i \rangle\, f_j = 0\,.$$

Since $\{f_j\}_{j=1,\ldots,m}$ is a basis for F, we obain that, for all $j = 1, \ldots, m$,

$$\sum_{i=1}^{n} \langle e^k, e_i \rangle = a_{kj} = 0\,,$$

implying that $e \otimes f = 0$ and the injectivity of u is shown. □

Proposition 6.11. *Let E and F be two vector spaces. Then there exists a unique isomorphism u between $E^* \otimes F^*$ and $(E \otimes F)^*$ such that*

$$\big(u(e^* \otimes f^*)\big)(e \otimes f) = \langle e^*, e \rangle_{E^*,E} \langle f^*, f \rangle_{F^*,F}\,,$$

for all $e \in E$, $f \in F$, $e^* \in E^*$ and $f^* \in F^*$.

Proof. We apply the universal factorization property of Proposition 6.3 to the bilinear map $\varphi : E^* \times F^* \longrightarrow (E \otimes F)^*$ defined by

$$\big(\varphi(e^* \otimes f^*)\big)(e \otimes f) = \langle e^*, e \rangle_{E^*,E} \langle f^*, f \rangle_{F^*,F}\,.$$

It remains to show that the resulting map $u : E^* \otimes F^* \longrightarrow (E \otimes F)^*$ is an isomorphism. This is done by choosing a basis for the vector spaces E, E^*, F and F^* and proceeding as in the proof of Proposition 6.10. □

Covariant and Contravariant Tensors

We recall that we consider finite dimensional vector spaces over a ground field $\mathcal{F}$ being the real number field $\mathbb{R}$ or the complex number field $\mathbb{C}$. Let E denote such a vector space.

Definition 6.12. *For all $r \in \mathbb{N}$, the space of **contravariant tensors** of degree r on E is defined by*

$$T^r(E) := \underbrace{E \otimes \ldots \otimes E}_{r-\textit{times}}\,.$$

$$\begin{tikzcd}
E^* \times F^* \arrow[r, "P"] \arrow[dr, "\varphi"'] & E^* \otimes F^* \arrow[d, "u"] \\
& (E \otimes F)^*
\end{tikzcd}$$

Fig. 6.4. Universal factorization property for $E^* \times F^*$.

Similarly, for all $s \in \mathbb{N}$, the space of ***covariant tensors*** *of degree s on E is defined by*

$$T_s(E) := \underbrace{E^* \otimes \ldots \otimes E^*}_{s\,-\,times} .$$

Note that for $r = 1$, respectively $s = 1$, we see that $T^1(E) = E$ and $T_1(E) = E^*$, respectively. And, by convention, we set $T^0(E) = T_0(E) = \mathcal{F}$.

Now, we give an explicit expression for these tensors in terms of a basis of E. – Let $\{e_i\}_{i=1,\ldots,n}$ denote a basis for E and $\{e^j\}_{j=1,\ldots,n}$ a basis for E^*, the dual of E. From Proposition 6.9, it follows that

$$\{e_{i_1} \otimes \ldots \otimes e_{i_r}\}_{1 \le i_1,\ldots,i_r \le n}$$

is a basis for $T^r(E)$. Hence, every contravariant tensor K of degree r can be expressed uniquely as linear combination

$$K = \sum_{i_1,\ldots,i_r=1}^{n} K^{i_1\ldots i_r}\, e_{i_1} \otimes \ldots \otimes e_{i_r} , \tag{6.4}$$

where $K^{i_1\ldots i_r} \in \mathcal{F}$ are the components of K with respect to the basis $\{e_i\}_{i=1,\ldots,n}$ of E. Similarly, for every covariant tensor L of degree s, there exist unique components $L_{j_1\ldots j_s} \in \mathcal{F}$ such that

$$L = \sum_{j_1,\ldots,j_s=1}^{n} L_{j_1\ldots j_s}\, e^{j_1} \otimes \ldots \otimes e^{j_s} . \tag{6.5}$$

In a next step, we want to study how the components of tensors transform under a change of basis of E. – Let $\{\hat{e}_i\}_{i=1,\ldots,n}$ denote another basis of E, which is related to the basis $\{e_i\}_{i=1,\ldots,n}$ by the linear transformation

$$e_i = \sum_{k=1}^{n} A_i^k\, \hat{e}_k , \qquad i = 1,\ldots,n . \tag{6.6}$$

We write the corresponding change of the two dual bases in E^* as

$$e^j = \sum_{l=1}^{n} B_l^j \, \hat{e}^l \, , \qquad j = 1, \dots, n \, . \tag{6.7}$$

It is easy to see that, for $i, j = 1, \dots, n$, we have

$$\delta_i^j = \langle e^j, e_i \rangle = \left\langle \sum_{l=1}^{n} B_l^j \, \hat{e}^l, \sum_{k=1}^{n} A_i^k \, \hat{e}_k \right\rangle = \sum_{k=1}^{n} B_k^j A_i^k \, ,$$

implying that $B = (B_i^j)$ is the inverse matrix of $A = (A_i^j)$.

Then, for a contravariant tensor K of degree r, the components $K^{i_1 \dots i_r}$ and $\hat{K}^{i_1 \dots i_r}$ with respect to the basis $\{e_i\}_{i=1,\dots,n}$ and $\{\hat{e}_i\}_{i=1,\dots,n}$, respectively, are related by the following formula:

$$\hat{K}^{i_1 \dots i_r} = \sum_{k_1, \dots, k_r = 1}^{n} A_{k_1}^{i_1} \dots A_{k_r}^{i_r} \, K^{k_1 \dots k_r} \, . \tag{6.8}$$

Similarly, the components of a covariant tensor L of degree s are related by the formula

$$\hat{L}_{j_1 \dots j_s} = \sum_{l_1, \dots, l_s = 1}^{n} B_{j_1}^{l_1} \dots B_{j_s}^{l_s} \, L_{l_1 \dots l_s} \, . \tag{6.9}$$

Remark. These last two formulas follow directly from the representations (6.4) and (6.5) together with Proposition 6.9.

Tensor Algebra

Definition 6.13. *For all $r \in \mathbb{N}$, $s \in \mathbb{N}$, the space of contravariant degree r and covariant degree s **(mixed) tensors** on E, or simply the space of tensors of type (r, s) on E, is defined by the following tensor product:*

$$T_s^r(E) := T^r(E) \otimes T_s(E) = \underbrace{E \otimes \dots \otimes E}_{r-times} \otimes \underbrace{E^* \otimes \dots \otimes E^*}_{s-times} \, .$$

Note that, in particular, we have that $T_0^r(E) = T^r(E)$, $T_s^0(E) = T_s(E)$ and $T_0^0(E) = \mathcal{F}$. – In terms of a basis $\{e_i\}_{i=1,\dots,n}$ of E and its dual basis $\{e^j\}_{j=1,\dots,n}$ of E^*, every tensor K of type (r, s) can be uniquely written as

$$K = \sum_{\substack{i_1, \dots, i_r = 1 \\ j_1, \dots, j_s = 1}}^{n} K_{j_1 \dots j_s}^{i_1 \dots i_r} \, e_{i_1} \otimes \dots \otimes e_{i_r} \otimes e^{j_1} \otimes \dots \otimes e^{j_s} \, , \tag{6.10}$$

where $K_{j_1 \dots j_s}^{i_1 \dots i_r}$ are the components of K with respect to the basis $\{e_i\}_{i=1,\dots,n}$. For a change of basis as in (6.6) and (6.7), we then have the following transformation of components formula for a tensor of type (r, s):

$$\hat{K}^{i_1 \dots i_r}_{j_1 \dots j_s} = \sum_{\substack{k_1, \dots, k_r = 1 \\ l_1, \dots, l_s = 1}}^{n} A^{i_1}_{k_1} \dots A^{i_r}_{k_r} \, B^{l_1}_{j_1} \dots B^{l_s}_{j_s} \, K^{k_1 \dots k_r}_{l_1 \dots l_s} \,. \tag{6.11}$$

Next, we group all the tensor spaces into one. More precisely, we set

$$T(E) := \bigoplus_{r,s=0}^{\infty} T^r_s(E) \,. \tag{6.12}$$

This vector space can be made to an associative algebra over $\mathcal{F}$, called the **tensor algebra** of E, by defining the product of two tensors in $T(E)$ as follows. – As usual, let $\{e_i\}_{i=1,\dots,n}$ denote a basis for E with corresponding dual basis $\{e^j\}_{j=1,\dots,n}$. Then, a tensor $K \in T^r_s(E)$ can be written as in (6.10). Similarly, a tensor $L \in T^p_q(E)$ can be expressed as

$$L = \sum_{\substack{k_1, \dots, k_p = 1 \\ l_1, \dots, l_q = 1}}^{n} L^{k_1 \dots k_p}_{l_1 \dots l_q} \, e_{k_1} \otimes \dots \otimes e_{k_p} \otimes e^{l_1} \otimes \dots \otimes e^{l_q} \,.$$

Then, we define the *product* of K and L, denoted by $K \otimes L$, to be the following element of $T^{r+p}_{s+q}(E)$:

$$K \otimes L = \sum_{\substack{i_1, \dots, i_{r+p} = 1 \\ j_1, \dots, j_{s+q} = 1}}^{n} (K \otimes L)^{i_1 \dots i_{r+p}}_{j_1 \dots j_{s+q}} \, e_{i_1} \otimes \dots \otimes e_{i_{r+p}} \otimes e^{j_1} \otimes \dots \otimes e^{j_{s+q}} \,, \tag{6.13}$$

where the components are precisely

$$(K \otimes L)^{i_1 \dots i_{r+p}}_{j_1 \dots j_{s+q}} := K^{i_1 \dots i_r}_{j_1 \dots j_s} \, L^{i_{r+1} \dots i_{r+p}}_{j_{s+1} \dots j_{s+q}} \,.$$

Remark. The product is a bilinear map from $T^r_s(E) \times T^p_q(E)$ into $T^{r+p}_{s+q}(E)$.

The definition of the product only makes sense if it is independent of the basis. – Let $\{\hat{e}_i\}_{i=1,\dots,n}$ denote another basis of E and let $\hat{K}^{i_1 \dots i_r}_{j_1 \dots j_s}$ and $\hat{L}^{k_1 \dots k_p}_{l_1 \dots l_q}$ be the components of $K \in T^r_s(E)$, respectively $L \in T^p_q(E)$ with respect to this new basis. Moreover, the components of the product of K and L with respect to the new basis are denoted by $(\widehat{K \otimes L})^{i_1 \dots i_{r+p}}_{j_1 \dots j_{s+q}}$. For the definition of the product to be independent of the basis, these components must satisfy the following equation:

$$(\widehat{K \otimes L})^{i_1 \dots i_{r+p}}_{j_1 \dots j_{s+q}} = \hat{K}^{i_1 \dots i_r}_{j_1 \dots j_s} \, \hat{L}^{i_{r+1} \dots i_{r+p}}_{j_{s+1} \dots j_{s+q}} \,. \tag{6.14}$$

We apply three times the change of components formula (6.11) for (r, s)-tensors to obtain

$$(\widehat{K \otimes L})^{i_1 \ldots i_{r+p}}_{j_1 \ldots j_{s+q}} = \sum_{\substack{k_1,\ldots,k_{r+p}=1 \\ l_1,\ldots,l_{s+q}=1}}^{n} A^{i_1}_{k_1} \ldots A^{i_{r+p}}_{k_{r+p}} B^{l_1}_{j_1} \ldots B^{l_{s+q}}_{j_{s+q}} (K \otimes L)^{k_1 \ldots k_{r+p}}_{l_1 \ldots l_{s+q}} , \tag{6.15a}$$

$$\hat{K}^{i_1 \ldots i_r}_{j_1 \ldots j_s} = \sum_{\substack{k_1,\ldots,k_r=1 \\ l_1,\ldots,l_s=1}}^{n} A^{i_1}_{k_1} \ldots A^{i_r}_{k_r} B^{l_1}_{j_1} \ldots B^{l_s}_{j_s} K^{k_1 \ldots k_r}_{l_1 \ldots l_s} , \tag{6.15b}$$

$$\hat{L}^{i_{r+1} \ldots i_{r+p}}_{j_{s+1} \ldots j_{s+q}} = \sum_{\substack{k_{r+1},\ldots,k_{r+p}=1 \\ l_{s+1},\ldots,l_{s+q}=1}}^{n} A^{i_{r+1}}_{k_{r+1}} \ldots A^{i_{r+p}}_{k_{r+p}} B^{l_{s+1}}_{j_{s+1}} \ldots B^{l_{s+q}}_{j_{s+q}} L^{k_{r+1} \ldots k_{r+p}}_{l_{s+1} \ldots l_{s+q}} . \tag{6.15c}$$

Then we observe that (6.15b) multiplied with (6.15c), combined with (6.15a) implies that (6.14) holds, as a consequence of (6.13). Hence, we have shown that the definition of the product is independent of the chosen basis.

Contraction of Tensors

We now define the notion of contraction. – Considering $T^r_s(E)$ together with a basis $\{e_i\}_{i=1,\ldots,n}$ for E and two integers p, q verifying $1 \leq p \leq r$, $1 \leq q \leq s$, we define the (p,q)-*contraction* of $K \in T^r_s(E)$, denoted by CK, to be the following element of $T^{r-1}_{s-1}(E)$:

$$CK = \sum_{\substack{i_1,\ldots,i_{r-1}=1 \\ j_1,\ldots,j_{s-1}=1}}^{n} (CK)^{i_1 \ldots i_{r-1}}_{j_1 \ldots j_{s-1}} e_{i_1} \otimes \ldots \otimes e_{i_{r-1}} \otimes e^{j_1} \otimes \ldots \otimes e^{j_{s-1}} , \tag{6.16}$$

with components

$$(CK)^{i_1 \ldots i_{r-1}}_{j_1 \ldots j_{s-1}} := \sum_{k=1}^{n} K^{i_1 \ldots k \ldots i_{r-1}}_{j_1 \ldots k \ldots j_{p-1}} ,$$

where the subscript k appears at the q-th position and the superscript k at the p-th position of the components of K with respect to the basis $\{e_i\}_{i=1,\ldots,n}$.

Remark. The (p,q)-contraction is a linear map from $T^r_s(E)$ into $T^{r-1}_{s-1}(E)$.

Again, we have to show that the definition is independent of the chosen basis for E. – Let $\{\hat{e}_i\}_{i=1,\ldots,n}$ denote another basis of E and let $\hat{K}^{i_1 \ldots i_r}_{j_1 \ldots j_s}$ be the components of $K \in T^r_s(E)$ with respect to this basis. Moreover, the components of the (p,q)-contraction in the new basis are denoted by $(\widehat{CK})^{i_1 \ldots i_{r-1}}_{j_1 \ldots j_{s-1}}$. From the change of components formula (6.11), we deduce that

$$(\widehat{CK})^{i_1 \ldots i_{r-1}}_{j_1 \ldots j_{s-1}} = \sum_{\substack{k_1,\ldots,k_{r-1}=1 \\ l_1,\ldots,l_{s-1}=1}}^{n} A^{i_1}_{k_1} \ldots A^{i_{r-1}}_{k_{r-1}} B^{l_1}_{j_1} \ldots B^{l_{s-1}}_{j_{s-1}} (CK)^{k_1 \ldots k_{r-1}}_{l_1 \ldots l_{s-1}} .$$

By definition of the (p,q)-contraction, it then follows

$$(\widehat{CK})^{i_1\ldots i_{r-1}}_{j_1\ldots j_{s-1}} = \sum_{k=1}^{n} \sum_{\substack{k_1,\ldots,k_{r-1}=1\\ l_1,\ldots,l_{s-1}=1}}^{n} A^{i_1}_{k_1}\ldots A^{i_{r-1}}_{k_{r-1}}\, B^{l_1}_{j_1}\ldots B^{l_{s-1}}_{j_{s-1}}\, K^{k_1\ldots k\ldots k_{r-1}}_{l_1\ldots k\ldots l_{s-1}}\,, \tag{6.17}$$

where again the subscript k appears at the q-th position and the superscript k at the p-th position.

On the other hand, by the change of components formula (6.11), we have that

$$\hat{K}^{i_1\ldots m\ldots i_{r-1}}_{j_1\ldots m\ldots j_{s-1}} = \sum_{\substack{k_1,\ldots,k,\ldots,k_{r-1}=1\\ l_1,\ldots,l,\ldots,l_{s-1}=1}}^{n} A^{i_1}_{k_1}\ldots A^{m}_{k}\ldots A^{i_{r-1}}_{k_{r-1}}\, B^{l_1}_{j_1}\ldots B^{l}_{m}\ldots B^{l_{s-1}}_{j_{s-1}}\, K^{k_1\ldots k\ldots k_{r-1}}_{l_1\ldots l\ldots l_{s-1}}\,.$$

Since $\sum_{m=1}^{n} A^m_k B^l_m = \delta^l_k$, we obtain

$$\begin{aligned}\sum_{m=1}^{n} \hat{K}^{i_1\ldots m\ldots i_{r-1}}_{j_1\ldots m\ldots j_{s-1}} &= \sum_{\substack{k_1,\ldots,k,\ldots,k_{r-1}=1\\ l_1,\ldots,l,\ldots,l_{s-1}=1}}^{n} \delta^l_k\, A^{i_1}_{k_1}\ldots A^{i_{r-1}}_{k_{r-1}}\, B^{l_1}_{j_1}\ldots B^{l_{s-1}}_{j_{s-1}}\, K^{k_1\ldots k\ldots k_{r-1}}_{l_1\ldots l\ldots l_{s-1}}\\ &= \sum_{k=1}^{n} \sum_{\substack{k_1,\ldots,k_{r-1}=1\\ l_1,\ldots,l_{s-1}=1}}^{n} A^{i_1}_{k_1}\ldots A^{i_{r-1}}_{k_{r-1}}\, B^{l_1}_{j_1}\ldots B^{l_{s-1}}_{j_{s-1}}\, K^{k_1\ldots k\ldots k_{r-1}}_{l_1\ldots k\ldots l_{s-1}}\,.\end{aligned}$$

This agrees with (6.17) showing that

$$(\widehat{CK})^{i_1\ldots i_{r-1}}_{j_1\ldots j_{s-1}} = \sum_{m=1}^{n} \hat{K}^{i_1\ldots m\ldots i_{r-1}}_{j_1\ldots m\ldots j_{s-1}}\,,$$

and hence the definition of the contraction does not depend on the choice of the basis for E.

Example 6.14. If $e \in E$ and $e^* \in E^*$, then $e \otimes e^* \in T^1_1(E)$ is a (mixed) tensor of type $(1,1)$. The contraction, or more precisely $(1,1)$-contraction, maps $e \otimes e^*$ into $\langle e^*, e\rangle \in \mathcal{F} = T^0_0(E)$ (see also Example 6.19).

Tensors and Multilinear Maps

We give an important interpretation of covariant and contravariant tensors as multilinear maps.

Proposition 6.15. *The vector space $T_s(E)$ of covariant tensors of degree s is isomorphic to the vector space of s-multilinear maps from $E \times \ldots \times E$ into $\mathcal{F}$. More precisely, there is an isomorphism, denoted by Σ_s, between $T_s(E)$ and s-multilinear maps from $E \times \ldots \times E$ into $\mathcal{F}$ given by*

$$\Sigma_s\big(e_1^* \otimes \ldots \otimes e_s^*\big)(e_1, \ldots, e_s) = \prod_{i=1}^{s} \langle e_i^*, e_i \rangle \,, \tag{6.18}$$

where $e_1^* \otimes \ldots \otimes e_s^* \in T_s(E)$ *and* $e_1, \ldots, e_s \in E$.

Proof. By generalizing Proposition 6.11, we deduce that $T_s(E) = E^* \otimes \ldots \otimes E^*$ can be seen as the dual space of $T^s(E) = E \otimes \ldots \otimes E$. On the other hand, it follows from the universal factorization property of the tensor product (see Proposition 6.3) that the space of linear maps from $T^s(E) = E \otimes \ldots \otimes E$ into $\mathcal{F}$ is isomorphic to the space of s-linear maps from $E \times \ldots \times E$ into $\mathcal{F}$. – For more details in the case of covariant tensors of degree 2 we refer to the example below. □

Remark. Following the last proposition, we consider $e_1^* \otimes \ldots \otimes e_s^* \in T_s(E)$ as s-multilinear map $E \times \ldots \times E \longrightarrow \mathcal{F}$ and use the notation $e_1^* \otimes \ldots \otimes e_s^*(e_1, \ldots, e_s) \in \mathcal{F}$, for $e_1, \ldots, e_s \in E$.

Example 6.16 (Covariant Tensors of degree 2). The universal factorization property in Proposition 6.3 implies that for every bilinear map $\varphi : E \times E \longrightarrow \mathcal{F}$, there exists a unique linear map $u : E \otimes E = T^2(E) \longrightarrow \mathcal{F}$ such that $\varphi = u \circ P$ (see Fig. 6.5). In other words, there is a one-to-one correspondence between bilinear maps from $E \times E$ into $\mathcal{F}$ and $(E \otimes E)^*$, i.e., linear maps from $E \otimes E$ into $\mathcal{F}$. On the other hand, from Proposition 6.11, we know that $(E \otimes E)^* = E^* \otimes E^*$. Hence, there is an isomorphism, denoted by Σ_2, between $E^* \otimes E^* = T_2(E)$ and the vector space of bilinear maps from $E \times E$ into $\mathcal{F}$ given by

$$\Sigma_2\big(e^* \otimes \tilde{e}^*\big)(e, \tilde{e}) = \langle e^*, e \rangle \, \langle \tilde{e}^*, \tilde{e} \rangle \,,$$

for all $e, \tilde{e} \in E$ and $e^*, \tilde{e}^*$.

$$\begin{tikzcd}
E \times E \arrow[r, "P"] \arrow[dr, "\varphi"'] & E \otimes E \arrow[d, "u"] \\
& \mathcal{F}
\end{tikzcd}$$

Fig. 6.5. Universal factorization property for $E \times E$.

Similarly, for contravariant tensors there is the following

Proposition 6.17. *The vector space $T^r(E)$ of contravariant tensors of degree r is isomorphic to the vector space of r-multilinear maps from $E^* \times \ldots \times E^*$ into $\mathcal{F}$.*

Proposition 6.18. *Let V be a vector space. Then the vector space $T_s(E) \otimes V$ is isomorphic to the vector space of s-multilinear maps from $E \times \ldots \times E$ into V. More precisely, there is an isomorphism, denoted by Σ_s^V, between $T_s(E) \otimes V$ and s-multilinear maps from $E \times \ldots \times E$ into the vector space V given by*

$$\Sigma_s^V \big(e_1^* \otimes \ldots \otimes e_s^* \otimes v\big)(e_1, \ldots, e_s) = \prod_{i=1}^{n} \langle e_i^*, e_i \rangle \, v \,, \tag{6.19}$$

where $e_1^* \otimes \ldots \otimes e_s^* \in T_s(E)$, $e_1, \ldots, e_s \in E$ and $v \in V$.

Proof. By Proposition 6.10, we know that $T_s(E) \otimes V$ is isomorphic to $\mathcal{L}\big(T_s(E)^*, V\big)$. Since $T^s(E)$ is the dual space of $T_s(E)$ (see Proposition 6.11), we obtain the existence of an isomorphism between $T_s(E) \otimes V$ and $\mathcal{L}\big(T^s(E), V\big)$. Again, by the universal factorization property of the tensor product, we get that $\mathcal{L}\big(T^s(E), V\big)$ can be identified with s-multilinear maps from $E \times \ldots \times E$ into the vector space V. □

Example 6.19. Consider $K \in T_1^1(E)$ having components $K_j^i \in \mathcal{F}$ with respect to a basis $\{e_i\}_{i=1,\ldots,n}$ of E. The endomorphism Σ_1^E corresponding to K sends $e_j \in E$ to $\sum_{i=1}^n K_j^i \, e_i$. This is a direct consequence of (6.19). Moreover, the contraction CK of K equals $\sum_{i=1}^n K_i^i$.

Skew-Symmetric Forms as Covariant Tensors

We want to show that due to Proposition 6.15 we have $\bigwedge^s(E) \subset T^s(E)$[3]. – Let $\alpha \in \bigwedge^s(E)$ and let $\{e^j\}_{j=1,\ldots,n}$ denote a basis of E^*. Then, we can write

$$\alpha = \sum_J \alpha_J \, e^{j_1} \wedge \ldots \wedge e^{j_s} \,,$$

where $J = \{(j_1, \ldots, j_s) \ : \ 1 \leq j_1 < \ldots < j_s \leq n\}$. Using formula (3.6) for the wedge product, it follows, for $e_1, \ldots, e_s \in E$,

$$\begin{aligned} e^{j_1} \wedge \ldots \wedge e^{j_s}(e_1, \ldots, e_s) &= \sum_{\sigma \in \mathcal{S}_s} (-1)^{|\sigma|} \left\langle e^{j_{\sigma(1)}}, e_1 \right\rangle \ldots \left\langle e^{j_{\sigma(s)}}, e_s \right\rangle \\ &= \sum_{\sigma \in \mathcal{S}_s} (-1)^{|\sigma|} \prod_{k=1}^{s} \left\langle e^{j_{\sigma(k)}}, e_k \right\rangle . \end{aligned}$$

[3] At this stage, the reader should be familiar with Section 3.1.1. In particular, it is not difficult to generalize Definition 3.1 to alternating p-multilinear maps from $E \times \ldots \times E$ into $\mathcal{F}$.

On the other hand, Proposition 6.15 implies that

$$\prod_{k=1}^{s} \langle e^{j_{\sigma(k)}}, e_k \rangle = e^{j_{\sigma(1)}} \otimes \ldots \otimes e^{j_{\sigma(s)}}(e_1, \ldots, e_s)\,.$$

Thus we arrive at the following representation for $\alpha \in \bigwedge^s(E)$:

$$\begin{aligned}
\alpha &= \sum_J \sum_{\sigma \in \mathcal{S}_s} (-1)^{|\sigma|} \alpha_J \, e^{j_{\sigma(1)}} \otimes \ldots \otimes e^{j_{\sigma(s)}} \\
&= \sum_{j_1,\ldots,j_s=1}^{n} \sum_{\sigma \in \mathcal{S}_s} \frac{1}{s!} (-1)^{|\sigma|} \alpha_J \, e^{j_{\sigma(1)}} \otimes \ldots \otimes e^{j_{\sigma(s)}}\,.
\end{aligned} \tag{6.20}$$

This shows that indeed $\bigwedge^s(E) \subset T^s(E)$.

Example 6.20. Let $E = \mathbb{R}^3$ with canonical basis $\{e_1, e_2, e_3\}$ and let $\alpha \in \bigwedge^2(\mathbb{R}^3)$ be given by

$$\alpha = \alpha_{12}\, e^1 \wedge e^2 + \alpha_{13}\, e^1 \wedge e^3 + \alpha_{23}\, e^2 \wedge e^3\,.$$

For $v, w \in \mathbb{R}^3$, it then follows

$$\alpha(v,w) = \alpha_{12}(v^1 w^2 - w^1 v^2) + \alpha_{13}(v^1 w^3 - w^1 v^3) + \alpha_{23}(v^2 w^3 - w^2 v^3)\,.$$

On the other hand, using (6.20), we obtain

$$\begin{aligned}
\alpha = &\frac{\alpha_{12}}{2}(e^1 \otimes e^2 - e^2 \otimes e^1) + \frac{-\alpha_{12}}{2}(e^2 \otimes e^1 - e^1 \otimes e^2) \\
&+ \frac{\alpha_{13}}{2}(e^1 \otimes e^3 - e^3 \otimes e^1) + \frac{-\alpha_{13}}{2}(e^3 \otimes e^1 - e^1 \otimes e^3) \\
&+ \frac{\alpha_{23}}{2}(e^2 \otimes e^3 - e^3 \otimes e^2) + \frac{-\alpha_{23}}{2}(e^3 \otimes e^2 - e^2 \otimes e^3)\,.
\end{aligned}$$

Applying this to $v, w \in \mathbb{R}^3$ we find indeed the above result for $\alpha(v,w)$.

Transporting Tensors

Let E and F be two finite dimensional $\mathcal{F}$-vector spaces. Moreover, let $A : E \longrightarrow F$ be an isomorphism . Then we define the map

$$\begin{aligned}
A^* : F^* &\longrightarrow E^*\,, \\
f^* &\longmapsto A^* f^*\,,
\end{aligned}$$

where $A^* f^* \in E^*$ satisfies, for all $e \in E$,

$$\langle A^* f^*, e \rangle_{E^*,E} = \langle f^*, A\, e \rangle_{F^*,F}\,. \tag{6.21}$$

The map A^* is clearly an isomorphism between F^* and E^* with inverse $(A^{-1})^* : E^* \longrightarrow F^*$ defined in the same way.

Using Proposition 6.7, we then deduce the existence of a unique linear map

$$\begin{aligned} A \otimes (A^{-1})^* : E \otimes E^* &\longrightarrow F \otimes F^* \\ e \otimes e^* &\longmapsto A\, e \otimes (A^{-1})^* e^* , \end{aligned} \tag{6.22}$$

where $e \in E$ and $e^* \in E^*$. It is straightforward to check that the map $A \otimes (A^{-1})^*$ is an isomorphism.

Next, we show that the map $A \otimes (A^{-1})^*$ commutes with contractions on $T_1^1(E) = E \otimes E^*$. – Recall that $(1,1)$-contractions on $T_1^1(E)$ are given by (see Example 6.14)

$$C(e \otimes e^*) = \langle e^*, e\rangle_{E^*,E} .$$

Then, we compute

$$\begin{aligned} C\big(A \otimes (A^{-1})^*(e \otimes e^*)\big) &= C\big(A\, e \otimes (A^{-1})^* e^*\big) \\ &= \big\langle (A^{-1})^* e^*, A\, e\big\rangle_{F^*,F} \\ &\overset{(6.21)}{=} \big\langle e^*, A^{-1} A\, e\big\rangle_{E^*,E} \\ &= \langle e^*, e\rangle_{E^*,E} = C(e \otimes e^*) , \end{aligned}$$

showing that $A \otimes (A^{-1})^*$ indeed respects the contraction operation.

For (mixed) tensors of arbitrary type, we have the following generalization of the previous results:

Proposition 6.21. *Let E, F be two finite dimensional $\mathcal{F}$-vector spaces and $A : E \longrightarrow F$ an isomorphism. Then there exists a unique isomorphism from $T_s^r(E)$ to $T_s^r(F)$, denoted by $A^{\otimes r} \otimes \big((A^{-1})^*\big)^{\otimes s}$, such that*

$$\begin{aligned} &A^{\otimes r} \otimes \big((A^{-1})^*\big)^{\otimes s}(e_1 \otimes \ldots \otimes e_r \otimes e_1^* \otimes \ldots \otimes e_s^*) \\ &\qquad = A\, e_1 \otimes \ldots \otimes A\, e_r \otimes (A^{-1})^* e_1^* \otimes \ldots \otimes (A^{-1})^* e_s^* , \end{aligned}$$

for all $e_1, \ldots, e_r \in E$ and all $e_1^*, \ldots, e_s^* \in E^*$. Moreover, the tensor algebra isomorphism

$$\tilde{A} := \bigoplus_{r,s=0}^{\infty} A^{\otimes r} \otimes \big((A^{-1})^*\big)^{\otimes s} : T(E) \longrightarrow T(F) ,$$

*called the **extension** of A, preserves the type of the tensors and commutes with all contractions.*

Proof. □

Now, we want to check the consistency of the extension $\tilde{A}$ with previously defined operations, namely with the pull-back of skew-symmetric s-forms on E.

For $f_1, \ldots, f_s \in F$, we deduce from Definition 3.8 that the pull-back $(A^{-1})^*\alpha \in \bigwedge^s(F)$ of α by A^{-1} is given by

$$\begin{aligned}\big((A^{-1})^*\alpha\big)(f_1, \ldots, f_s) &= \alpha(A^{-1}f_1, \ldots, A^{-1}f_s) \\ &= \sum_J \alpha_J \, e^{j_1} \wedge \ldots \wedge e^{j_s}(A^{-1}f_1, \ldots, A^{-1}f_s)\end{aligned}$$

Then, using formula (3.6) for the wedge product, we obtain

$$\begin{aligned}\big((A^{-1})^*\alpha\big)(f_1, \ldots, f_s) &= \sum_J \sum_{\sigma \in \mathcal{S}_s} (-1)^{|\sigma|} \alpha_J \, e^{j_{\sigma(1)}}(A^{-1}f_1) \ldots e^{j_{\sigma(s)}}(A^{-1}f_s) \\ &= \sum_J \sum_{\sigma \in \mathcal{S}_s} (-1)^{|\sigma|} \alpha_J \prod_{k=1}^{s} \big\langle e^{j_{\sigma(k)}}, A^{-1}f_k \big\rangle_{E^*,E} . \qquad (6.23)\end{aligned}$$

On the other hand, using (6.20), we obtain

$$\tilde{A}\alpha = \sum_J \sum_{\sigma \in \mathcal{S}_s} (-1)^{|\sigma|} \alpha_J \; \tilde{A}(e^{j_{\sigma(1)}} \otimes \ldots \otimes e^{j_{\sigma(s)}}) .$$

It follows, using Proposition 6.21 for the extension $\tilde{A}$,

$$\tilde{A}(e^{j_{\sigma(1)}} \otimes \ldots \otimes e^{j_{\sigma(s)}}) = (A^{-1})^* e^{j_{\sigma(1)}} \otimes \ldots \otimes (A^{-1})^* e^{j_{\sigma(s)}} .$$

For $f_1, \ldots, f_s \in F$, we then deduce that

$$\begin{aligned}(A^{-1})^* e^{j_{\sigma(1)}} \otimes \ldots \otimes (A^{-1})^* e^{j_{\sigma(s)}}(f_1, \ldots, f_s) \quad &= \quad \prod_{k=1}^{s} \big\langle (A^{-1})^* e^{j_{\sigma(k)}}, f_k \big\rangle_{F^*,F} \\ &\overset{(6.21)}{=} \prod_{k=1}^{s} \big\langle e^{j_{\sigma(k)}}, A^{-1}f_k \big\rangle_{E^*,E} .\end{aligned}$$

This agrees with the expression (6.23) for the pull-back.

6.2 Tensor Fields on Manifolds

We can now do with tensors what we did with alternating forms when defining differential forms on a manifold: To each point of a manifold a tensor on the tangent space of this point is assigned in a smooth way. The utilized concepts will be exactly the same as for the tangent bundle and the cotangent bundle.

6.2.1 The Tensor Bundle

Let M be a n-dimensional C^k-differentiable manifold and $p \in M$. We already know that the tangent space T_pM at p is a n-dimensional vector space over the ground field $\mathcal{F} = \mathbb{R}$ (see Definition 2.35). We denote by

$$T_s^r(T_pM)$$

the tensors of type (r,s) on T_pM. From a "set" point of view the **tensor bundle** of type (r,s) on M is then defined by

$$T_s^r(M) = \bigcup_{p\in M} T_s^r(T_pM)\,.$$

For $K \in T_s^r(T_pM)$, the projection map to the base point reads as

$$\begin{aligned}\pi : T_s^r(M) &\longrightarrow M\,,\\ K &\longmapsto p\,.\end{aligned} \tag{6.24}$$

Let (U,x) be a local chart on M. A basis for T_pM , $p \in U$, is then given by $\{\frac{\partial}{\partial x_i}(p)\}_{i=1,\ldots,n}$ (see (2.15)) and the corresponding dual basis for T_p^*M is $\left\{dx^j(p)\right\}_{j=1,\ldots,n}$ (see (3.26)). Using Proposition 6.9, we deduce that

$$\left\{\frac{\partial}{\partial x_{i_1}}(p)\otimes\ldots\otimes\frac{\partial}{\partial x_{i_r}}(p)\otimes dx^{j_1}(p)\otimes\ldots\otimes dx^{j_s}(p)\right\}\,, \tag{6.25}$$

with all indices running from 1 to n, is a basis of $T_s^r(T_pM)$. – At this stage, it is important to note that we have slightly changed the notation for the dual basis $\left\{dx^j(p)\right\}_{j=1,\ldots,n}$ of T_p^*M. More precisely, in order to be consistent with the previous section we write the elements of the basis of T_p^*M with an *upper* index, whereas in Chapter 3.2 a *lower* index is used, in order to emphasize that the elements are one-forms on U corresponding to the coordinate functions $x_i : U \longrightarrow \mathbb{R}$, $i = 1,\ldots,n$.

Definition 6.22. *Let M be a C^k-differentiable manifold of dimension n. A C^{k-1}-**tensor field** of type (r,s) on M is a map*

$$T : M \longrightarrow T_s^r(M)$$

such that $\pi\circ T = id_M$, and for all local charts (U,x) on M^n there exist functions $T^{i_1\ldots i_r}_{j_1\ldots j_s} \in C^{k-1}(U,\mathbb{R})$ satisfying

$$T(p) = \sum_{\substack{i_1,\ldots,i_r=1\\ j_1,\ldots,j_s=1}}^{n} T^{i_1\ldots i_r}_{j_1\ldots j_s}(p)\,\frac{\partial}{\partial x_{i_1}}(p)\otimes\ldots\otimes\frac{\partial}{\partial x_{i_r}}(p)\otimes dx^{j_1}(p)\otimes\ldots\otimes dx^{j_s}(p)\,, \tag{6.26}$$

for all $p \in U$. – We denote the vector space of C^∞-tensor fields of type (r,s) on M by $\mathcal{T}_s^r(M)$.

Next, we want to construct a differentiable structure on the tensor bundle $T_s^r(M)$ of type (r,s) in such a way that C^k-tensor fields on M are just C^k-*sections* of $T_s^r(M)$, i.e., $T \in C^k(M, T_s^r(M))$ and $\pi\circ T = id_M$. The method will be very similar to the one used for the construction of the differentiable structure on the tangent bundle TM in Section 2.5.3 and on the cotangent bundle $\bigwedge^p T^*M$ of degree p in Section 3.2.1. Recall that in the case of $r=1$, $s=0$ we are dealing with the tangent bundle and in the case of $r=0$, $s=1$ with the cotangent bundle. – For the sake of simplicity, we assume $r=s=1$ for the construction of the differential structure on the tensor bundle.

Let $\{(U_i,\varphi_i)\}_{i\in I}$ be an atlas for M. We set

$$T^r_s(U_i) = \bigcup_{p\in U_i} T^r_s(T_pM) = \pi^{-1}(U_i)\,,$$

and define, combining (2.19) with (3.29),

$$\begin{aligned} \Phi_i^{1,1} : T^r_s(U_i) &\longrightarrow \mathbb{R}^n \times \big(\mathbb{R}^n \otimes (\mathbb{R}^n)^*\big)\,, \\ T &\longmapsto \big(\varphi_i \circ \pi(T), \varphi_i^{1,1}(T)\big)\,, \end{aligned} \tag{6.27}$$

with

$$\varphi_i^{1,1}(T) := (d\varphi_i)_p \otimes \big((d\varphi_i^{-1})_{\varphi_i(p)}\big)^*(T)\,, \tag{6.28}$$

where $\pi(T)=p\in U_i$. More precisely, writing $T=[\gamma]\otimes w_p \in T_pM\otimes T_p^*M$, it follows from Proposition 6.21 that

$$\varphi_i^{1,1}([\gamma]\otimes w_p) = (d\varphi_i)_p\cdot[\gamma]\otimes\big((d\varphi_i^{-1})_{\varphi_i(p)}\big)^* w_p \in \mathbb{R}^n\otimes(\mathbb{R}^n)^*\,.$$

It is clear that $\Phi_i^{1,1}$ is a bijection from $T^r_s(U_i)$ into $\varphi_i(U_i)\times\big(\mathbb{R}^n\otimes(\mathbb{R}^n)^*\big)$. We then say that $\Omega\subset T^r_s(M)$ is *open* if and only if for all $i\in I$ the set $\Phi_i^{1,1}(\Omega\cap T^r_s(U_i)$ is open in $\mathbb{R}^n\times\big(\mathbb{R}^n\otimes(\mathbb{R}^n)^*\big)$.

Proposition 6.23. *These open sets define a (separated) topology on $T^r_s(M)$ which depends only on the differentiable structure of M, and not on the atlas $\{(U_i,\varphi_i)\}_{i\in I}$ representing the fixed differentiable structure on M. Moreover, the maps $\Phi_i^{1,1}$, $i\in I$, defined in (6.27) are homeomorphisms for this topology and the projection π is continuous.*

Proposition 6.24. *Let M be a n-dimensional C^k-manifold with $k\geq 2$ and let $\{(U_i,\varphi_i)\}_{i\in I}$ be an atlas for M^n. Then $\{(T^1_1(U_i),\Phi_i^{1,1})\}_{i\in I}$ defines a C^{k-1}-differentiable structure on the tensor bundle $T^1_1(M)$ of type $(1,1)$, depending only on the differential structure on M^n. And $T^1_1(M)$ is a C^{k-1}-differentiable manifold of dimension $n+n^2$ for this differentiable structure. Moreover, the projection π is a C^{k-1}-submersion for this differentiable structure.*

The proofs of the last two propositions are essentially based on the explicit expression for the transition functions (this was already the case for the tangent bundle and the cotangent bundle of degree p, see Section 2.5.3 and 3.2.1). Therefore, we first state the following

Lemma 6.25. *Let $\{(U_i, \varphi_i)\}_{i\in I}$ and $\{(V_j, \psi_j)\}_{j\in J}$ be two equivalent systems of charts for M. Moreover, let $i \in I$ and $j \in J$ such that $U_i \cap V_j \neq \emptyset$. Then on $\Phi_i^{1,1}(T_1^1(U_i \cap V_j)) \subset \mathbb{R}^n \times \big(\mathbb{R}^n \otimes (\mathbb{R}^n)^*\big)$ the following formula for the transition functions on $T_1^1(M)$ holds:*

$$\Psi_j^{1,1} \circ \big(\Phi_i^{1,1}\big)^{-1}(x, \xi \otimes \beta) = \\ \Big(\psi_j \circ \varphi_i^{-1}(x) \;\; , \;\; d(\psi_j \circ \varphi_i^{-1})_x \cdot \xi \otimes \big(d(\varphi_i \circ \psi_j^{-1})_{\psi_j \circ \varphi_i^{-1}(x)}\big)^* \beta\Big)\,, \tag{6.29}$$

where $\varphi_i^{-1}(x) = p \in U_i \cap V_j$, $\xi \in \mathbb{R}^n$ and $\beta \in (\mathbb{R}^n)^*$.

Proof. The result is a straightforward consequence of (2.21) and (3.31). □

As mentioned before, the previous results, generalized to arbitrary $r \in \mathbb{N}$ and $s \in \mathbb{N}$, allow us to interpret tensor fields as sections of the tensor bundle.

Proposition 6.26. *Let M be a $C^{k'}$-differentiable manifold of dimension n, with $k' \geq k + 1$. Then a C^k-tensor field of type (r, s) on M is a C^k-section of $T_r^s(M)$ for the above defined differential structure.*

Proof. It is not necessary to give a proof because similar results have already been shown in great detail at several places in previous chapters (see especially Proposition 3.22). □

Remark. The last proposition can be interpreted as alternative definition for tensor fields on manifolds.

Examples of Tensor Fields

Example 6.27 (Vector Fields and One-Forms). Let X be a C^k-vector field on M^n. By Definition 2.44, we then have that $X(p) \in T_pM = T_0^1(T_pM)$, for all $p \in M^n$. Moreover, in a local chart (U, x) the C^k-vector field can be represented as

$$X(p) = \sum_{i=1}^{n} X^i(p)\, \frac{\partial}{\partial x_i}(p)\,,$$

where $X^i \in C^k(U, \mathbb{R})$ and $p \in U$. Note again that, compared to (2.27), the components of the vector field have now an *upper* index in order to be consistent with (6.26). Hence, a C^k-vector field can be seen as C^k-tensor field of type $(1, 0)$.

Let $\omega \in \Omega_k^1(M)$ be a C^k-differential one-form on M. We then have that $\omega(p) \in \bigwedge^1(T_pM) = T_p^*M = T_1^0(T_p^*M)$, for all $p \in M$. Moreover, in a local chart (U, x) the C^k-differential one-form can be represented as

$$\omega(p) = \sum_{j=1}^{n} \omega_j(p)\, dx^j(p)\,,$$

where $\omega_j \in C^k(U,\mathbb{R})$ and $p \in U$. Hence, a C^k-differential one-form can be seen as C^k-tensor field of type $(0,1)$.

Example 6.28 (Riemannian Metric). There is a very important example for a tensor field given by the following

Definition 6.29. *Let M^n be a C^∞-manifold of dimension n. A* ***Riemannian metric*** *g on M is a tensor field of type $(0,2)$ (or a covariant tensor field of degree 2) such that, for all $p \in M$,*

(i) it is non-negative, i.e., $g_p(X,X) \geq 0$, for all $X \in T_pM$;
(ii) it is non-degenerate, i.e., $g_p(X,X) = 0$ if and only if $X = 0$, for $X \in T_pM$;
(iii) it is symmetric, i.e., $g_p(X,Y) = g_p(Y,X)$, for all $X,Y \in T_pM$.

In terms of a local chart (U,x), the Riemannian metric reads as

$$g = \sum_{i,j=1}^{n} g_{ij}\, dx^i \otimes dx^j \,, \tag{6.30}$$

where obviously for the components

$$g_{ij}(p) = g_p\left(\frac{\partial}{\partial x_i}(p), \frac{\partial}{\partial x_j}(p)\right)\,, \qquad p \in U\,.$$

Note that by Proposition 6.15, for all $p \in U$, the covariant tensor g_p of type $(0,2)$ is now interpreted as bilinear map $T_pM \times T_pM \longrightarrow \mathbb{R}$. Moreover, $\big(g_{ij}(p)\big)$ is a symmetric, positive definite $n \times n$-matrix.

* * *

Next, we give an important characterization of covariant tensor fields. – We denote C^∞-functions on M by $\mathcal{F}(M)$. Note that for $f \in \mathcal{F}(M)$ and $X \in \mathcal{X}(M)$ the assignment $p \longmapsto f(p)X(p)$ defines a new C^∞-vector field on M denoted by $f\,X$.

Proposition 6.30. *Let $T \in \mathcal{T}^0_s(M)$ be a covariant tensor field of degree s on M. Then, it can be considered as s-multilinear map, also denoted by T, from $\mathcal{X}(M) \times \ldots \times \mathcal{X}(M)$ into $\mathcal{F}(M)$ such that*

$$T(f_1\,X_1,\ldots,f_s\,X_s) = f_1\cdots f_s\,T(X_1,\ldots,X_s)\,, \tag{6.31}$$

for all $f_1,\ldots,f_s \in \mathcal{F}(M)$ and $X_1,\ldots,X_s \in \mathcal{X}(M)$. Conversely, any s-multilinear map satisfying (6.31) can be considered as covariant tensor field of degree s on M.

Proof. Let $T \in \mathcal{T}^0_s(M)$. Then Proposition 6.15 implies that, for all $p \in M$, the covariant tensor $T_p \in T^0_s(T_pM) = T^*_pM \otimes \ldots \otimes T^*_pM$ identifies to an s-multilinear map, also denoted by T_p, from $T_pM \times \ldots \times T_pM$ into $\mathbb{R}$. More precisely, for $X_1, \ldots, X_s \in \mathcal{X}(M)$ and all $p \in M$, the map

$$\begin{aligned} T_p : T_pM \times \ldots \times T_pM &\longrightarrow \mathbb{R}\,, \\ \big(X_1(p), \ldots, X_s(p)\big) &\longmapsto T_p\big(X_1(p), \ldots, X_s(p)\big) \end{aligned}$$

is s-multilinear.

Next, we consider the s-multilinear map defined by

$$\begin{aligned} T : \mathcal{X}(M) \times \ldots \times \mathcal{X}(M) &\longrightarrow \mathcal{F}(M)\,, \\ (X_1, \ldots, X_s) &\longmapsto T(X_1, \ldots, X_s)\,, \end{aligned}$$

where

$$\begin{aligned} T(X_1, \ldots, X_s) : M &\longrightarrow \mathbb{R}\,, \\ p &\longmapsto T\big(X_1, \ldots, X_s\big)(p) := T_p\big(X_1(p), \ldots, X_s(p)\big)\,. \end{aligned}$$

Obviously, we see that the map T satisfies (6.31). However, it remains to show that indeed $T(X_1, \ldots, X_s) \in \mathcal{F}(M)$. – For this purpose, let (U, x) be a local chart for M implying that $T \in \mathcal{T}^0_s(M)$ can be written as

$$T = \sum_{j_1, \ldots, j_s = 1}^{n} T_{j_1 \ldots j_s}\, dx^{j_1} \otimes \ldots \otimes dx^{j_s}\,.$$

Denoting by $\langle \cdot, \cdot \rangle_{T^*_pM, T_pM}$ the dual pairing of T^*_pM and T_pM, we deduce from Proposition 6.15 that

$$T_p\big(X_1(p), \ldots, X_s(p)\big) = \sum_{j_1, \ldots, j_s = 1}^{n} T_{j_1 \ldots j_s}(p) \prod_{k=1}^{n} \big\langle dx^{j_k}(p), X_k(p) \big\rangle_{T^*_pM, T_pM} \in \mathbb{R}\,,$$

and moreover that

$$T(X_1, \ldots, X_s) = \sum_{j_1, \ldots, j_s = 1}^{n} T_{j_1 \ldots j_s} \prod_{k=1}^{n} \big\langle dx^{j_k}, X_k \big\rangle_{T^*M, TM}\,,$$

implying that $T(X_1, \ldots, X_s)$ is indeed a real-valued C^∞-function on M.

Conversely, let $T : \mathcal{X}(M) \times \ldots \times \mathcal{X}(M) \longrightarrow \mathcal{F}(M)$ be a s-multilinear map satisfying (6.31). – The essential point of the proof is to show that the value of the C^∞-function $T(X_1, \ldots, X_s)$ at a point $p \in M$ depends only on the values of the vector fields $X_1, \ldots, X_s \in \mathcal{X}(M)$ at the *same* point p. So this will be done first.

By the s-multilinearity of the map T, this translates to the assertion that if $X_1(p) = 0$, for some $p \in M$, then for all $X_2, \ldots, X_s \in \mathcal{X}(M)$ the following holds:

$$T\big(X_1, \ldots, X_s\big)(p) = 0\,.$$

In order to show this, we choose a local chart (U, x) about p so that we can write

$$X_1 = \sum_{i=1}^{n} f^i \, \frac{\partial}{\partial x_i} \qquad \text{on } \; U\,.$$

Next, we extend $f^i \in C^\infty(U, \mathbb{R})$ to $g^i \in \mathcal{F}(M)$ such that $f^i = g^i$ on some neighborhood $U' \subset U$ of p. Similarly, we take vector fields $Z_i \in \mathcal{X}(M)$ such that $\frac{\partial}{\partial x_i} = Z_i$ on U'. Hence, we have that

$$X_1 = \sum_{i=1}^{n} g^i \, Z_i \qquad \text{on } \; U'\,.$$

By the localization property of the map T in the Lemma 6.31 below, we then deduce that

$$T(X_1, X_2, \ldots, X_s) = T\left(\sum_{i=1}^{n} g^i \, Z_i, X_2, \ldots, X_s\right) \overset{(6.31)}{=} \sum_{i=1}^{n} g^i \; T(Z_i, X_2, \ldots, X_s)\,.$$

Since $f^i(p) = g^i(p) = 0$, for $i = 1, \ldots, n$, it follows that the map $T(X_1, \ldots, X_s)$ vanishes at $p \in M$.

Now, we are ready to establish the fact that T is a covariant C^∞-tensor field of degree s on M. – Since, as shown before, the value of $T(X_1, \ldots, X_s)$ at $p \in M$ depends only on the values of $X_1, \ldots, X_s \in \mathcal{X}(M)$ at p, we get that T induces for each p an assignment of a s-multilinear map of $T_pM \times \ldots \times T_pM$ into $\mathbb{R}$. This is by Proposition 6.15 equivalent to an assignment for each $p \in M$ of an element in $T_p^*M \otimes \ldots \otimes T_p^*M$. In terms of a local chart (U, x), there exist thus components $T_{j_1 \ldots j_s}(p) \in \mathbb{R}$ such that, for all $p \in U$,

$$T(p) = \sum_{j_1, \ldots, j_s = 1}^{n} T_{j_1 \ldots j_s}(p) \, dx^{j_1}(p) \otimes \ldots \otimes dx^{j_s}(p)\,.$$

In order to show that $T \in \mathcal{T}_s^0(M)$, we still have to check that the functions $p \longmapsto T_{j_1 \ldots j_s}(p)$ are C^∞-functions on U. – For this purpose, we construct as before extensions $Z_1, \ldots, Z_n$ in $\mathcal{X}(M)$ of $\frac{\partial}{\partial x_i}$, $i = 1, \ldots, n$, such that $\frac{\partial}{\partial x_i} = Z_i$ on U', where $U' \subset U$. Since by assumption $T(Z_{j_1}, \ldots, Z_{j_s}) \in \mathcal{F}(M)$, we obtain that

$$T(Z_{j_1}, \ldots, Z_{j_s}) = T_{j_1 \ldots j_s} \in C^\infty(U', \mathbb{R})\,.$$

Moreover, assuming that we start with an atlas $\{(U_i, x_i)\}_{i \in I}$ on M such that $\{(U_i', x_i)\}_{i \in I}$ is still an atlas on M, it follows[4] that $T \in \mathcal{T}_s^0(M)$. – This completes the proof of the proposition. $\square$

[4] If T has C^∞-components with respect to one particular atlas, it also has C^∞-components with respect to every other atlas representing the differentiable structure on M.

Lemma 6.31 (Localization Property). *Let $T : \mathcal{X}(M) \times \ldots \times \mathcal{X}(M) \longrightarrow \mathcal{F}(M)$ be a s-multilinear map satisfying (6.31). Moreover, let U be an open subset of M, $X_1, \ldots, X_s \in \mathcal{X}(M)$ and $Y_1, \ldots, Y_s \in \mathcal{X}(M)$. Then, if there exists $i = 1, \ldots, s$ such that $X_i = Y_i$ on U, we have that*

$$T(X_1, \ldots, X_s) = T(Y_1, \ldots, Y_s) \qquad \text{on } U\,.$$

Proof. Because of the s-multilinearity of the map T, it suffices to show that if $X_1 = 0$ on U, then for all $X_2, \ldots, X_s \in \mathcal{X}(M)$, we have that

$$T(X_1, \ldots, X_s) = 0 \qquad \text{on } U\,.$$

For any $y \in U$, let $f_1 \in \mathcal{F}(M)$ such that $f_1(y) = 0$ and $f_1 \equiv 1$ outside of U. It is clear that $f_1\, X_1 = X_1$. Moreover, using (6.31), it follows that

$$T(X_1, \ldots, X_s) = T(f_1\, X_1, \ldots, X_s) = f_1\, T(X_1, \ldots, X_s)\,.$$

This implies that $T(X_1, \ldots, X_s)$ vanishes at y, showing the lemma. □

Differential Forms as Covariant Tensor Fields

We want to establish that C^∞-differential s-forms on M identify to covariant C^∞-tensor fields of degree s on M which are *skew-symmetric*. Note that the case $s = 1$ is clear, as described in Example 6.27.

Recall that in terms of a local chart (U, x) a C^∞-differential s-form ω on M can be written as (note again the change in the indexation compared to (3.28))

$$\omega = \sum_J \omega_J\, dx^{j_1} \wedge \ldots \wedge dx^{j_s}\,,$$

where $J = \{(j_1, \ldots, j_s) \ : \ 1 \le j_1 < \ldots < j_s \le n\}$. For $X_1, \ldots, X_s \in T_pM$, we deduce, using formula (3.6) for the wedge product,

$$dx^{j_1}(p) \wedge \ldots \wedge dx^{j_s}(p)\big(X_1, \ldots, X_s\big) = \sum_{\sigma \in \mathcal{S}_s} (-1)^{|\sigma|} \prod_{k=1}^{s} \big\langle dx^{j_{\sigma(k)}}(p), X_k \big\rangle_{T_p^*M, T_pM}\,.$$

On the other hand, Proposition 6.15 implies that

$$\prod_{k=1}^{s} \big\langle dx^{j_{\sigma(k)}}(p), X_k \big\rangle_{T_p^*M, T_pM} = dx^{j_{\sigma(1)}}(p) \otimes \ldots \otimes dx^{j_{\sigma(s)}}(p)\big(X_1, \ldots, X_s\big)\,.$$

Thus we arrive at the following representation for $\omega \in \Omega^s(M)$:

$$\begin{aligned}
\omega &= \sum_J \sum_{\sigma \in \mathcal{S}_s} (-1)^{|\sigma|} \omega_J\, dx^{j_{\sigma(1)}} \otimes \ldots \otimes dx^{j_{\sigma(s)}} \\
&= \sum_{j_1, \ldots, j_s = 1}^{n} \sum_{\sigma \in \mathcal{S}_s} \frac{1}{s!} (-1)^{|\sigma|} \omega_J\, dx^{j_{\sigma(1)}} \otimes \ldots \otimes dx^{j_{\sigma(s)}}\,, \qquad (6.32)
\end{aligned}$$

showing that $\omega \in \mathcal{T}_s^0(M)$.

Starting with a covariant tensor field one can also ask if it is nothing else than a differential form. This motivates the following

Definition 6.32. *Let $T \in \mathcal{T}_s^0(M)$ be a covariant tensor field of type $(0, s)$. We define the **alternation** AT of T by*

$$AT_p(X_1, \dots, X_s) = \frac{1}{s!} \sum_{\sigma \in \mathcal{S}_s} (-1)^{|\sigma|}\, T_p(X_{\sigma(1)}, \dots, X_{\sigma(s)})\,,$$

*where $X_1, \dots, X_s \in T_pM$. Moreover, we define the **symmetrization** ST of T by*

$$ST_p(X_1, \dots, X_s) = \frac{1}{s!} \sum_{\sigma \in \mathcal{S}_s} T_p(X_{\sigma(1)}, \dots, X_{\sigma(s)})\,.$$

The following proposition, which can be easily verified, gives an interpretation of skew-symmetric covariant tensor fields as differential forms.

Proposition 6.33. *Let $T \in \mathcal{T}_s^0(M)$ be a covariant tensor field of type $(0, s)$. Then, we have that $AT \in \Omega^s(M)$ and $T \in \Omega^s(M)$ if and only if $AT = T$. Moreover, for $\omega \in \Omega^s(M)$ and $\tilde{\omega} \in \Omega^{\tilde{s}}(M)$, we have that $\omega \otimes \tilde{\omega} \in \mathcal{T}_{s+\tilde{s}}^0(M)$ and $A(\omega \otimes \tilde{\omega}) = \omega \wedge \tilde{\omega}$.*

Example 6.34. Consider $T \in \mathcal{T}_2^0(M)$ given by

$$T = \sum_{i,j=1}^{n} T_{ij}\, dx^i \otimes dx^j\,.$$

We want to determine the alternation $A(dx^i \otimes dx^j)$. – Let $X_1, X_2 \in T_pM$. From Definition 6.32, we then deduce

$$\begin{aligned} A\big(dx^i(p) \otimes dx^j(p)\big)(X_1, X_2) &= \frac{1}{2} \sum_{\sigma \in \mathcal{S}_2} (-1)^{|\sigma|} dx^i(p) \otimes dx^j(p)\big(X_{\sigma(1)}, X_{\sigma(2)}\big) \\ &= \frac{1}{2}\Big(dx^i(p) \otimes dx^j(p)\big(X_1, X_2\big) - dx^i(p) \otimes dx^j(p)\big(X_2, X_1\big)\Big)\,. \end{aligned}$$

In terms of the wedge product the right-hand side becomes

$$\frac{1}{2} dx^i(p) \wedge dx^j(p)\big(X_1, X_2\big)\,.$$

Thus, we end up with

$$AT = \sum_{i,j=1}^{n} \frac{1}{2} T_{ij}\, dx^i \wedge dx^j = \sum_{i<j=1}^{n} T_{ij}\, dx^i \wedge dx^j\,,$$

showing that indeed $AT \in \Omega^2(M)$.

Vector-Valued Differential Forms

At this stage, we want to introduce differential forms on a manifold which are *vector-valued.* More precisely, let V be a vector space. Then we consider s-linear and alternating maps $E \times \ldots \times E \longrightarrow V$. Generalizing the results of Chapter 3 to these maps, we can define V-valued C^k-differential s-forms on a manifold, denoted by $\Omega_k^s(M, V)$.

Recall that by Proposition 6.18, a s-linear map $E \times \ldots \times E \longrightarrow V$ can be seen as an element of $T_s(E) \otimes V$. Because of the interpretation of (real-valued) differential forms as skew-symmetric covariant tensor fields, we then deduce the identification of $\Omega_k^s(M, V)$ with C^k-sections of $\bigwedge^s T^*M \otimes V$.

6.2.2 Transporting Tensor Fields

The results of Section 6.1 concerning the transport of tensors are now generalized to tensor fields. – For this purpose, we first set

$$\mathcal{T}(M) = \bigoplus_{r,s=0}^{n} \mathcal{T}_s^r(M)\,.$$

Note that $\mathcal{T}(M)$ is an associative algebra over $\mathbb{R}$, the product $\otimes$ being defined pointwise, i.e., if $T_1, T_2 \in \mathcal{T}(M)$, then $(T_1 \otimes T_2)_p = T_1(p) \otimes T_2(p)$, for all $p \in M$ (see (6.13) for the definition of the product). We call $\mathcal{T}(M)$ the **tensor field algebra**.

Now, let $\varphi : M \longrightarrow N$ be a C^∞-diffeomorphism between two C^∞-differentiable manifolds M and N. Then, for all $p \in M$, the tangent map $d\varphi_p$ of φ at p is an isomorphism between the tangent spaces T_pM and $T_{\varphi(p)}N$. By Proposition 6.21 this isomorphism can be extended to an isomorphism $\tilde{\varphi}_p$ between the tensor algebras $T(T_pM)$ and $T(T_{\varphi(p)}N)$. Moreover, given a C^∞-tensor field $T \in \mathcal{T}_s^r(M)$, we define a tensor field $\tilde{\varphi}\,T$ on N by

$$(\tilde{\varphi}\,T)_q = \tilde{\varphi}_{\varphi^{-1}(q)}(T_{\varphi^{-1}(q)})\,, \qquad q \in N\,. \tag{6.33}$$

More precisely, we have with Proposition 6.21 that

$$(\tilde{\varphi}\,T)_q = \left(d\varphi_{\varphi^{-1}(q)}\right)^{\otimes r} \otimes \left((d\varphi_q^{-1})^*\right)^{\otimes s} (T_{\varphi^{-1}(q)})\,. \tag{6.34}$$

In this way, every diffeomorphism $\varphi : M \longrightarrow N$ induces an algebra isomorphism between $\mathcal{T}(M)$ and $\mathcal{T}(N)$ which preserves the type and commutes with all contractions.

It is clear that $(\tilde{\varphi}\,T)_q \in T_s^r(T_qN)$, for all $q \in N$. Next, we show that $\tilde{\varphi}T \in \mathcal{T}_s^r(N)$. – Let (V, y) be a chart about q. Moreover, let $p = \varphi^{-1}(q) \in M$ and set $U = \varphi^{-1}(V)$, $x = y \circ \varphi$. Then, we see that (U, x) is a chart about p. In terms of this chart, the tensor field $T \in \mathcal{T}_s^r(M)$ can be written as (see Definition 6.22)

$$T = \sum_{\substack{i_1,\dots,i_r=1 \\ j_1,\dots,j_s=1}}^{n} T^{i_1\dots i_r}_{j_1\dots j_s} \frac{\partial}{\partial x_{i_1}} \otimes \dots \otimes \frac{\partial}{\partial x_{i_r}} \otimes dx^{j_1} \otimes \dots \otimes dx^{j_s} .$$

From (6.34), we deduce

$$\tilde{\varphi}\left(\frac{\partial}{\partial x_{i_1}} \otimes \dots \otimes \frac{\partial}{\partial x_{i_r}} \otimes dx^{j_1} \otimes \dots \otimes dx^{j_s} \right)$$
$$= d\varphi \cdot \frac{\partial}{\partial x_{i_1}} \otimes \dots \otimes d\varphi \cdot \frac{\partial}{\partial x_{i_r}} \otimes (\varphi^{-1})^* dx^{j_1} \otimes \dots \otimes (\varphi^{-1})^* dx^{j_s} .$$

Well-known results for the tangent map and the pull-back then give the existence of components $\tilde{T}^{i_1\dots i_r}_{j_1\dots j_s} \in C^\infty(V, \mathbb{R})$ such that

$$\tilde{\varphi} T = \sum_{\substack{i_1,\dots,i_r=1 \\ j_1,\dots,j_s=1}}^{n} \tilde{T}^{i_1\dots i_r}_{j_1\dots j_s} \frac{\partial}{\partial y_{i_1}} \otimes \dots \otimes \frac{\partial}{\partial y_{i_r}} \otimes dy^{j_1} \otimes \dots \otimes dy^{j_s} ,$$

showing that indeed $\tilde{\varphi} T \in \mathcal{T}^r_s(N)$.

6.3 Lie Derivative of Tensor Fields

Let M be a differentiable manifold and let $Y \in \mathcal{X}(M)$. From Section 2.7, we know that for all $p \in M$ the local flow $\Gamma^Y_t : U \longrightarrow M$ of Y, where U is an open neighborhood of p and $t \in I \subset \mathbb{R}$ a time for which the local flow exists, defines a C^∞-diffeomorphism between U and $\Gamma^Y_t(U)$. For the sake of simplicity, we assume that Y is *complete*, i.e., we can take $U = M$ and $I = \mathbb{R}$; in other words, the flow Γ^Y_t exists globally. Applying the results of the previous Section 6.2.2 to the diffeomorphism $\Gamma^Y_t : M \longrightarrow M$, we can construct for each $t \in \mathbb{R}$ an algebra automorphism $\tilde{\Gamma}^Y_t$ of $\mathcal{T}(M)$, called the *extension* of Γ^Y_t.

Definition 6.35. *Let $T \in \mathcal{T}^r_s(M)$ and $Y \in \mathcal{X}(M)$. We define the **Lie derivative** of T with respect to Y to be the following element of $\mathcal{T}^r_s(M)$:*

$$(L_Y T)_p = \lim_{t \longrightarrow 0} \frac{1}{t} \big(T(p) - \tilde{\Gamma}^Y_t T(p) \big) , \qquad p \in M . \tag{6.35}$$

Remark. a) In the particular case of a vector field $X \in \mathcal{X}(M) = \mathcal{T}^1_0(M)$, we deduce from (6.34) that the defining equation (6.35) reduces to

$$(L_Y X)_p = \lim_{t \longrightarrow 0} \frac{1}{t} \Big(X(p) - \big(d\Gamma^Y_t\big)_{\Gamma^Y_{-t}(p)} \cdot X\big(\Gamma^Y_{-t}(p)\big) \Big) .$$

This agrees up to a sign with (2.32), being the defining equation for the bracket $[X, Y]$ of the vector fields X and Y (see Section 2.7.1). Hence, for the Lie derivative of vector fields, we have

$$L_Y X = [Y, X] . \tag{6.36}$$

b) In the case of a function $f \in \mathcal{F}(M) = \mathcal{T}_0^0(M)$, we observe that (recall the defining equation for the flow of a vector field)

$$(L_Y f)_p \overset{(6.35)}{=} \lim_{t \longrightarrow 0} \frac{1}{t}\Big(f(p) - f\big(\Gamma_{-t}^Y(p)\big)\Big)$$
$$= -\lim_{t \longrightarrow 0} \frac{1}{t}\Big(f\big(\Gamma_{-t}^Y(p)\big) - f(p)\Big) = -df \cdot (-Y) = df \cdot Y\,.$$

Hence, for the Lie derivative of a function, we have

$$L_Y f = df \cdot Y\,. \tag{6.37}$$

Proposition 6.36. *Let L_Y be the Lie derivative with respect to $Y \in \mathcal{X}(M)$. Then*

(i) it is linear, i.e., for all T_1, $T_2 \in \mathcal{T}(M)$, we have

$$L_Y(T_1 + T_2) = L_Y(T_1) + L_Y(T_2)\,;$$

(ii) it satisfies a Leibniz rule, i.e., for all T_1, $T_2 \in \mathcal{T}(M)$, we have

$$L_Y(T_1 \otimes T_2) = L_Y T_1 \otimes T_2 + T_1 \otimes L_Y T_2\,;$$

(iii) it preserves the type, i.e., $L_Y(\mathcal{T}_s^r(M)) \subset \mathcal{T}_s^r(M)$;
(iv) and it commutes with every contraction of a tensor field.

Proof. □

By a *derivation* of $\mathcal{T}(M)$, we mean a map of $\mathcal{T}(M)$ into itself satisfying the conditions of the previous proposition. Hence, we have shown that the Lie derivative L_Y with respect to the vector field $Y \in \mathcal{X}(M)$ is a derivation of $\mathcal{T}(M)$ which satisfies also $L_Y f = df \cdot Y$ and $L_Y X = [Y, X]$, for all $f \in \mathcal{F}(M)$ and $X \in \mathcal{X}(M)$.

Next, we note that the set of all derivations of $\mathcal{T}(M)$ forms a Lie algebra over $\mathbb{R}$ with respect to the canonical addition and multiplication and the bracket operation for two derivations D_1, D_2 defined by

$$[D_1, D_2]T = D_1(D_2 T) - D_2(D_1 T)\,, \qquad T \in \mathcal{T}(M)\,.$$

The following proposition then shows that the map associating to a vector field its corresponding Lie derivative is a homomorphism of Lie algebras (see Example 9.92 below).

Proposition 6.37. *Let X and $Y \in \mathcal{X}(M)$. For their corresponding Lie derivatives, we then have*

$$L_{[X,Y]} = [L_X, L_Y]\,.$$

Proof. By virtue of the Lemma 6.38 below, it is sufficient to show that the derivations $L_{[X,Y]}$ and $[L_X, L_Y]$ agree on $\mathcal{F}(M)$ and $\mathcal{X}(M)$. – For $f \in \mathcal{F}(M)$, we observe that (see the proof of Proposition 2.53)

$$L_{[X,Y]}f \overset{(6.37)}{=} df \cdot [X,Y] = d(df \cdot Y) \cdot X - d(df \cdot X) \cdot X \,.$$

It is easy to check that the right-hand side equals $[L_X, L_Y]f$. Moreover, for $Z \in \mathcal{X}(M)$, we have

$$\begin{aligned} [L_X, L_Y]Z &= L_X(L_Y Z) - L_Y(L_X Z) \\ &\overset{(6.36)}{=} \big[X,[Y,Z]\big] - \big[Y,[X,Z]\big] = \big[[X,Y],Z\big] \end{aligned}$$

where we used the Jacobi identity in Proposition 2.56 for the last equality. Again by (6.36), we have that $\big[[X,Y],Z\big] = L_{[X,Y]}Z$. This completes the proof of the proposition. □

Lemma 6.38. *Two derivations of the algebra of tensor fields $\mathcal{T}(M)$ are equal if they agree on $\mathcal{F}(M)$ and $\mathcal{X}(M)$.*
Proof. □

Proposition 6.39. *Let $T \in \mathcal{T}^0_s(M)$ be a covariant tensor field of degree s and $X_1, \ldots, X_s \in \mathcal{X}(M)$. Then, for the Lie derivative $L_Y T \in \mathcal{T}^0_s(M)$ with respect to any $Y \in \mathcal{X}(M)$, we have*

$$\begin{aligned} L_Y T(\, X_1 \,, \ldots, X_s) & \\ &= d\big(T(X_1, \ldots, X_s)\big) \cdot Y - \sum_{i=1}^{s} T\big(X_1, \ldots, L_Y X_i, \ldots, X_s\big) \\ &, \end{aligned} \tag{6.38}$$

where $L_Y X_i = [Y, X_i]$ is the i-th argument of the covariant tensor field T in the sum of the right-hand side.

Proof. As a consequence of Proposition 6.30, we observe that

$$T(X_1, \ldots, X_s) = C_1 \ldots C_s(T \otimes X_1 \otimes \ldots \otimes X_s)\,,$$

where $C_1, \ldots, C_s$ are obvious contractions (see Section 6.1). Thus, it follows

$$L_Y\big(T(X_1, \ldots, X_s)\big) = L_Y\big(C_1 \ldots C_s(T \otimes X_1 \otimes \ldots \otimes X_s)\big)\,. \tag{6.39}$$

On the other hand, it follows from the Leibniz rule (ii) in Proposition 6.36 that

$$\begin{aligned} L_Y(T \otimes X_1 \otimes \ldots \otimes X_s) &= L_Y T \otimes X_1 \otimes \ldots \otimes X_s \\ &+ T \otimes L_Y X_1 \otimes \ldots \otimes X_s + \ldots + T \otimes X_1 \ldots \otimes L_Y X_s \,. \end{aligned}$$

Since by (iv) in the same Proposition 6.36 the Lie derivative L_Y commutes with contractions, we then obtain, using (6.39),

$$\begin{aligned} L_Y\big(T(X_1,\dots,X_s)\big) &= C_1\dots C_s(L_Y T\otimes X_1\otimes\dots\otimes X_s) \\ &\quad+\sum_{i=1}^{s} C_1\dots C_s(T\otimes X_1\otimes\dots\otimes L_Y X_i\otimes\dots X_s)\,. \end{aligned}$$

Using (6.37) for the left-hand side, this becomes

$$\begin{aligned} d\big(T(X_1,\dots,X_s)\big)\cdot Y &= L_Y T(X_1,\dots,X_s) \\ &\quad+\sum_{i=1}^{s} T\big(X_1,\dots,L_Y X_i,\dots,X_s\big)\,, \end{aligned}$$

showing the result. □

Example 6.40. Let $\omega \in \mathcal{T}_1^0(M) = \Omega^1(M)$ and $X \in \mathcal{X}(M)$. Then (6.38) reduces to

$$L_Y\omega(X) = d\big(\omega(X)\big)\cdot Y - \omega([Y,X])\,. \tag{6.40}$$

In particular, for $\omega = df$ with $f \in \mathcal{F}(M)$, we obtain

$$\begin{aligned} L_Y df(X) &= d\big(df(X)\big)\cdot Y - df([Y,X]) \\ &= d\big(df(X)\big)\cdot Y - d\big(df(X)\big)\cdot Y + d\big(df(Y)\big)\cdot X \\ &= d\big(df(Y)\big)\cdot X = d(L_Y f)\cdot X\,. \end{aligned}$$

Thus, on $\mathcal{F}(M)$ it follows

$$L_Y\circ d = d\circ L_Y\,. \tag{6.41}$$

6.4 Relation Among the Operations d, L_X and int_X

In this section, we consider $\omega \in \Omega^s(M)$. The exterior derivative $d\omega \in \Omega^{s+1}(M)$ of ω as well as the interior product $\mathrm{int}_X\omega \in \Omega^{s-1}(M)$ of ω and $X \in \mathcal{X}(M)$ are defined in Section 3.2.2. Now we want to establish a relation between these two operations and the previously defined Lie derivative[5] $L_X\omega \in \Omega^s(M)$ of ω with respect to X.

Proposition 6.41 (Cartan's Formula). *Let $\omega \in \Omega^s(M)$ and $X \in \mathcal{X}(M)$. Then, we have*

$$L_X\omega = d(\mathit{int}_X\omega) + \mathit{int}_X(d\omega)\,. \tag{6.42}$$

[5] Recall that $L_X\omega$ is well-defined because $\Omega^s(M) \subset \mathcal{T}_s^0(M)$.

Proof. Let $q \in M$ and assume first that the vector field $X \in \mathcal{X}(M)$ does not vanish at q, i.e., $X(q) \neq 0$. From the Straightening Theorem 1.25 generalized to vector fields on manifolds, we then deduce that there exists a local chart (U, x) about q such that

$$X = \frac{\partial}{\partial x_1} \qquad \text{on } U .$$

a) In a first step, we have to compute $L_X\omega$, for $\omega \in \Omega^s(M)$. – By Definition 6.35, we have

$$(L_X\omega)_p = \lim_{t\longrightarrow 0} \frac{1}{t}\big(\omega(p) - \tilde{\Gamma}_t^X \omega(p)\big) , \qquad p \in U . \tag{6.43}$$

Recall that $\tilde{\Gamma}_t^X$ denotes the extension of the flow Γ_t^X of X, where the time $t \in \mathbb{R}$ is sufficiently small. Moreover, from (6.34) we deduce (see also the Definition 3.23 of the pull-back)

$$\begin{aligned}\tilde{\Gamma}_t^X \omega(p) &= \big(d(\Gamma_t^X)_p^{-1}\big)^* \omega\big((\Gamma_t^X)^{-1}(p)\big) \\ &= \big(d(\Gamma_{-t}^X)_p\big)^* \omega\big(\Gamma_{-t}^X(p)\big) = \big((\Gamma_{-t}^X)^*\omega\big)(p) .\end{aligned}$$

From the particular choice of the chart, we deduce that the coordinate expression for the inverse flow of X is given by

$$\Gamma_{-t}^X(x_1, \ldots, x_n) = (x_1 - t, x_2, \ldots, x_n) , \tag{6.44}$$

if the manifold M is assumed to be n-dimensional. Assuming that ω has the local form

$$\omega = \sum_I \omega_I \, dx^{i_1} \wedge \ldots \wedge dx^{i_s} \qquad \text{on } U ,$$

we conclude, using well-known properties of the pull-back,

$$\begin{aligned}\big((\Gamma_{-t}^X)^*\omega\big)(p) \quad &= \quad \sum_I \omega_I\big(\Gamma_{-t}^X(p)\big)\, d\big(x_{i_1} \circ \Gamma_{-t}^X\big) \wedge \ldots \wedge d\big(x_{i_s} \circ \Gamma_{-t}^X\big) \\ &\overset{(6.44)}{=} \sum_I \omega_I(x_1 - t, x_2, \ldots, x_n)\, dx^{i_1} \wedge \ldots \wedge dx^{i_s} .\end{aligned}$$

Inserting this result into (6.43), we obtain that the Lie derivative $(L_X\omega)_p$ at $p \in U$ equals

$$\lim_{t\longrightarrow 0} \sum_I \frac{\omega_I(x_1, x_2, \ldots, x_n) - \omega_I(x_1 - t, x_2, \ldots, x_n)}{t}\, dx^{i_1} \wedge \ldots \wedge dx^{i_s} ,$$

and thus

$$(L_X\omega)_p = \sum_I \frac{\partial \omega_I}{\partial x_1}(p)\, dx^{i_1} \wedge \ldots \wedge dx^{i_s} . \tag{6.45}$$

b) In a second step, we compute $d(\mathrm{int}_X\omega)$. – For this purpose, we first note that in the local chart (U,x) the interior product is given by (see Definition 3.28)

$$\begin{aligned}\mathrm{int}_X dx^{i_1}\wedge\ldots\wedge dx^{i_s}(X_2,\ldots,X_s) &= \mathrm{int}_{\frac{\partial}{\partial x_1}} dx^{i_1}\wedge\ldots\wedge dx^{i_s}(X_2,\ldots,X_s)\\ &= dx^{i_1}\wedge\ldots\wedge dx^{i_s}\left(\frac{\partial}{\partial x_1},X_2,\ldots,X_s\right).\end{aligned}$$

where $X_2,\ldots,X_s\in\mathcal{X}(M)$. By a well-known formula for the wedge product, the right-hand side of the last equation equals

$$\sum_{\sigma\in\mathcal{S}_s}(-1)^{|\sigma|}\left\langle dx^{i_{\sigma(1)}},\frac{\partial}{\partial x_1}\right\rangle\prod_{j=2}^{s}\left\langle dx^{i_{\sigma(j)}},X_{\sigma(j)}\right\rangle.$$

Clearly, the expression $\left\langle dx^{i_{\sigma(1)}},\frac{\partial}{\partial x_1}\right\rangle$ vanishes if $i_{\sigma(1)}\neq 1$ and otherwise equals 1. Thus, it follows

$$\begin{aligned}\mathrm{int}_{\frac{\partial}{\partial x_1}} dx^1\wedge dx^{i_2}\wedge\ldots\wedge dx^{i_p}(X_2,\ldots,X_s) &= \sum_{\sigma\in\mathcal{S}_{s-1}}(-1)^{|\sigma|}\prod_{j=2}^{s}\left\langle dx^{i_{\sigma(j)}},X_{\sigma(j)}\right\rangle\\ &= dx^{i_2}\wedge\ldots\wedge dx^{i_s}(X_2,\ldots,X_s).\end{aligned}\tag{6.46}$$

This result[6] together with the local expression of ω in the chart (U,x), then implies

$$\mathrm{int}_{\frac{\partial}{\partial x_1}}\omega=\sum_{\tilde I}\omega_{\tilde I}\,dx^{i_2}\wedge\ldots\wedge dx^{i_s},$$

where $\tilde I=\{(i_1,\ldots,i_s)\ :\ 1=i_1<i_2<\ldots<i_s\leq n\}\subset I$. Taking the exterior derivative of the last equation, we get

$$\begin{aligned}d(\mathrm{int}_{\frac{\partial}{\partial x_1}}\omega) &= \sum_{\tilde I}\frac{\partial\omega_{\tilde I}}{\partial x_1}\,dx^1\wedge dx^{i_2}\wedge\ldots\wedge dx^{i_s}\\ &\quad+\sum_{k=2}^{n}\sum_{\tilde I}\frac{\partial\omega_{\tilde I}}{\partial x_k}\,dx^k\wedge dx^{i_2}\wedge\ldots\wedge dx^{i_s}.\end{aligned}\tag{6.47}$$

c) In a third step, we compute $\mathrm{int}_X(d\omega)$. – We write the exterior derivative of $\omega\in\Omega^s(M)$ in the form

$$\begin{aligned}d\omega &= \sum_{\tilde I^c}\frac{\partial\omega_{\tilde I^c}}{\partial x_1}\,dx^1\wedge dx^{i_1}\wedge\ldots\wedge dx^{i_s}\\ &\quad+\sum_{k=2}^{n}\sum_{I}\frac{\partial\omega_I}{\partial x_k}\,dx^k\wedge dx^{i_1}\wedge\ldots\wedge dx^{i_s},\end{aligned}$$

[6] Compare with 3.49.

where $\tilde{I}^c = \{(i_1, \ldots, i_s) \ : \ 1 < i_1 < i_2 < \ldots < i_s \leq n\} \subset I$. Using (6.46) for both terms, we obtain[7]

$$
\begin{aligned}
\operatorname{int}_{\frac{\partial}{\partial x_1}}(d\omega) &= \sum_{\tilde{I}^c} \frac{\partial \omega_{\tilde{I}^c}}{\partial x_1} \, dx^{i_1} \wedge \ldots \wedge dx^{i_p} \\
&\quad - \sum_{k=2}^{n} \sum_{\tilde{I}} \frac{\partial \omega_{\tilde{I}}}{\partial x_k} \, dx^k \wedge dx^{i_2} \wedge \ldots \wedge dx^{i_p} \, . \qquad (6.48)
\end{aligned}
$$

d) Now, we put the results togther. –Adding (6.47) and (6.48), we obtain

$$
d(\operatorname{int}_X \omega) + \operatorname{int}_X(d\omega) = \sum_{\tilde{I}} \frac{\partial \omega_{\tilde{I}}}{\partial x_1} \, dx^1 \wedge dx^{i_2} \wedge \ldots \wedge dx^{i_p} + \sum_{\tilde{I}^c} \frac{\partial \omega_{\tilde{I}^c}}{\partial x_1} \, dx^{i_1} \wedge \ldots \wedge dx^{i_p} \, ,
$$

and conclude that this agrees with (6.45). Hence, the formula (6.42) holds locally under the assumption that $X(p) \neq 0$. □

As an application of Cartan's formula (6.42), we prove the following coordinate-free characterization for the exterior derivative.

Proposition 6.42. *Let $\omega \in \Omega^s(M)$ and let $X_0 \in \mathcal{X}(M)$. For the exterior derivative of ω, we then have*

$$
\begin{aligned}
d\omega(X_0, X_1, \ldots, X_s) &= \sum_{i=0}^{s} (-1)^i \, d\big(\omega(X_0, \ldots, \hat{X}_i, \ldots, X_s)\big) \cdot X_i \\
&\quad + \sum_{0 \leq i < j \leq s} (-1)^{i+j} \, \omega\big([X_i, X_j], X_0, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_s\big) \, ,
\end{aligned}
\qquad (6.49)
$$

where the "hat" over a vector field means that the latter is omitted and $X_1, \ldots, X_s \in \mathcal{X}(M)$. In particular, for $\omega \in \Omega^1(M)$, we have

$$
d\omega(X_0, X_1) = d\big(\omega(X_1)\big) \cdot X_0 - d\big(\omega(X_0)\big) \cdot X_1 - \omega\big([X_0, X_1]\big) \, . \qquad (6.50)
$$

Proof. The proof is by induction on s. – First note that if ω is an element of $\mathcal{F}(M)$, i.e., if $s = 0$, then (6.49) states that $d\omega(X_0) = d\omega \cdot X_0$, which is obviously correct.

For $s = 1$, we have $\omega \in \Omega^1(M)$ and from Cartan's formula (6.42) it follows

$$
\begin{aligned}
L_{X_0}\omega(X_1) &= d(\operatorname{int}_{X_0}\omega) \cdot X_1 + \big(\operatorname{int}_{X_0}(d\omega)\big)(X_1) \\
&= d\big(\omega(X_0)\big) \cdot X_1 + d\omega(X_0, X_1) \, .
\end{aligned}
$$

Using Proposition 6.39 for $L_{X_0}\omega$, we conclude

[7] The minus sign in the last line comes from the permutation of dx^k and dx^1 before applying formula (6.46) for the interior product.

$$\begin{aligned} d\omega(X_0, X_1) &= L_{X_0}\omega(X_1) - d\big(\omega(X_0)\big) \cdot X_1 \\ &= d\big(\omega(X_1)\big) \cdot X_0 - \omega(L_{X_0}X_1) - d\big(\omega(X_0)\big) \cdot X_1 \\ &= d\big(\omega(X_1)\big) \cdot X_0 - d\big(\omega(X_0)\big) \cdot X_1 - \omega\big([X_0, X_1]\big) , \end{aligned}$$

Hence, formula (6.50) is established.

We now complete the proof by induction. – Assume that the formula (6.49) is correct for all differential forms of degree $s-1$. If $\omega \in \Omega^s(M)$ we start as before with Cartan's formula (6.42). Then we use the induction assumption for the term $\mathrm{int}_{X_0}\omega \in \Omega^{s-1}(M)$, the Definition 3.28 of the interior product for the term $\mathrm{int}_{X_0}(d\omega)$ and for $L_{X_0}\omega$ the explicit expression (6.38). (This is the same strategy we used before for one-forms.) After a short calculation, we observe that (6.49) remains true for differential forms of degree s. □

* * *

Exercises.

Exercise 6.43. Show that the bilinear map T on $\mathcal{X}(M) \times \mathcal{X}(M)$, given by

$$T(X, Y) = d\big(\omega(Y)\big) \cdot X - d\big(\omega(X)\big) \cdot Y - \omega\big([X, Y]\big) ,$$

defines a covariant tensor field of type $(0, 2)$ on M.

Hint. In a first step, show that the map

$$\begin{aligned} T(X, Y) : M &\longrightarrow \mathbb{R} . \\ p &\longmapsto T\big(X, Y\big)(p) := \big(T(X, Y)\big)_p \end{aligned}$$

is an element of $\mathcal{F}(M)$. Hence, it follows

$$T : \mathcal{X}(M) \times \mathcal{X}(M) \longrightarrow \mathcal{F}(M) .$$

In a second step, show that the value of $T(X, Y) \in \mathcal{F}(M)$ at some point $p \in M$ depends only on the values of the vector fields X, Y at the *same* point p. Then, as in the proof of Proposition 6.30, we obtain the result.

7 An Introduction to Lie Groups

In this chapter, we give the basic definitions and properties of Lie groups. We develop here enough tools concerning Lie groups in order to be able to study fiber bundle in Chapter 9.

7.1 Lie Groups

Definition 7.1. *A **Lie group** G is a differentiable manifold which is endowed with a group structure such that the group operations*

$$G \times G \longrightarrow G\,, \qquad (g_1, g_2) \longmapsto g_1 g_2\,,$$

and

$$G \longrightarrow G\,, \qquad g \longmapsto g^{-1}\,,$$

are differentiable maps between $G \times G$ and G, respectively, G and G.

Remark. In other words, for a Lie group the group structure is consistent with the differentiable structure.

Definition 7.2. *We define the **left translation** by $g_0 \in G$ to be the following map:*

$$\begin{aligned} L_{g_0} : G &\longrightarrow G\,, \\ g &\longmapsto g_0 g\,. \end{aligned} \tag{7.1}$$

We show that the left translation is a diffeomorphism between G and G. – By Definition 7.1, we see that $L_{g_0} \in C^\infty(G, G)$. Moreover, we observe that

$$L_{g_0^{-1}} \circ L_{g_0}(g) = L_{g_0^{-1}}(g_0 g) = g_0^{-1}(g_0 g) = g\,, \qquad g \in G\,.$$

Hence, the map L_{g_0} is invertible with inverse $L_{g_0^{-1}}$. The inverse being also differentiable by definition, we conclude that the left translation is a diffeomorphism.

Remark. a) The differentiability of the inversion map follows from the differentiability of the multiplication map. Hence, it would be sufficient to require the differentiability of the multiplication in the Definition 7.1 of Lie groups. – Indeed, denoting the neutral element by $e \in G$, we want to find a solution $h(g) \in G$ of $gh(g) = e$, for $g \in G$. Since the partial derivative of the multiplication map with respect to the second argument is just the differential of the left translation L_g being an isomorphism, we deduce by the Implicit Function Theorem 1.14 that the solution $h(g) = g^{-1}$ is a smooth function on G.

b) Similarly, it is possible to define the *right* translation $R_{g_0} : G \longrightarrow G$, $g \longmapsto gg_0$. But we prefer the left translation, since $L_{g_1} \circ L_{g_0} = L_{g_1 g_0}$ in opposite to the more "complicated" formula $R_{g_1} \circ R_{g_0} = R_{g_0 g_1}$.

The next definition combines the concepts of subgroup and submanifold.

Definition 7.3. *A **(regular) Lie subgroup** H of a Lie group G is a subgroup of G that is also a submanifold of G. A Lie subgroup H of G is itself a Lie group.*

Theorem 7.4. *Let H be a closed subgroup of a Lie group G. Then H is a Lie subgroup.*

Remark. The previous theorem gives a powerful tool in order to find examples for Lie subgroups (see Example 7.7 below).

Examples of Lie Groups

Example 7.5. The simplest example for a Lie group is $\mathbb{R}^n$ with the usual vector addition. Left translation by $v \in \mathbb{R}^n$ is just $L_v w = v + w$, for all $w \in \mathbb{R}^n$.

Example 7.6 (The General Linear Group $GL_n(\mathbb{R})$). The **general linear group** is defined by

$$GL_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) \ : \ \det A \neq 0\}\,, \tag{7.2}$$

where $M_n(\mathbb{R})$ denotes the space of real square matrices of order n which can be identified with $\mathbb{R}^{n^2}$. The general linear group is a differentiable manifold of dimension n^2, since it is an open subset of $\mathbb{R}^{n^2}$. Indeed, consider the continuous map

$$\begin{aligned} \det : \mathbb{R}^{n^2} &\longrightarrow \mathbb{R}\,, \\ A &\longmapsto \det A\,. \end{aligned}$$

Then, we observe that $GL_n(\mathbb{R})$ is the inverse image of the open set $\mathbb{R} \setminus \{0\}$ under this map. Moreover, for the usual matrix multiplication and matrix inversion, with identity element $\mathbf{1} \in GL_n(\mathbb{R})$ given by the identity $n \times n$ matrix, the general linear group can be made to a Lie group. Note that the smoothness of the group operations is obvious, since the formulas for the product and inverse of matrices are smooth functions of the matrix components. – Note also that $GL_n(\mathbb{R})$ is obviously non-compact, since it is open.

Example 7.7 (The Lie Subgroups of $GL_n(\mathbb{R})$).

a) As a direct consequence of Example 2.4, we get that the orthogonal group $O(n)$ is a Lie subgroup of $GL_n(\mathbb{R})$ with dimension $n(n-1)/2$. Note that, it is also compact. Indeed, from the determinant map we see that $O(n)$ is closed and the boundedness follows directly from the definition of the orthogonal group.
b) In an analogous manner, we deduce from Example 2.5 that the special orthogonal group $SO(n)$ is also a compact Lie subgroup of $GL_n(\mathbb{R})$.

Remark. Considering the case of *complex* matrices, we recover the same type of matrix Lie groups as for real matrices. – More precisely, the (**complex**) **general linear group** $GL_n(\mathbb{C})$ defined by

$$GL_n(\mathbb{C}) = \{A \in M_n(\mathbb{C}) \ : \ \det A \neq 0\} \tag{7.3}$$

is a Lie group of (real) dimension $2n^2$ with the **unitary group**

$$U(n) = \{A \in GL_n(\mathbb{C}) \ : \ A^\dagger A = A\, A^\dagger = \mathbf{1}\}\,, \tag{7.4}$$

where $A^\dagger$ denotes the hermitian conjugate $\bar{A}^T$ of $A \in GL_n(\mathbb{C})$, and the **special unitary group**

$$SU(n) = \{A \in U(n) \ : \ \det A = 1\} = U(n) \cap SL(n, \mathbb{C}) \tag{7.5}$$

as Lie subgroups. Note also that $U(n)$ and $SU(n)$ are both compact. We will come back later to these matrix Lie groups. – For more details, we refer to [].

In the following, the tangent space T_eG of G at the neutral element e will play an important role for the study of geometrical aspects of Lie groups. This becomes already clear in the next definition.

Definition 7.8. *Let G be a Lie group with neutral element e. We then define $\theta \in \Omega^1(G, T_eG)$ by*

$$\theta_g(X) = (dL_{g^{-1}})_g \cdot X\,, \tag{7.6}$$

*where $X \in T_gG$. The T_eG-valued differential one-form θ is called the **canonical one-form** or **Maurer-Cartan form**.*

Remark. Note that the Maurer-Cartan form acting on vectors in T_eG coincides with the identity of T_eG. More precisely, for $A \in T_eG$, we have[1]

$$\theta_e(A) = (dL_e)_e \cdot A = A\,. \tag{7.7}$$

[1] The particular choice of notation for elements in T_eG will become clear later.

Next, we want to show that the Maurer-Cartan form is *left invariant*, i.e., it satisfies

$$(L_g)^*\theta = \theta\,, \tag{7.8}$$

for all $g \in G$. – Let $X \in T_gG$ and $g_0 \in G$. By definition of the pull-back, we have

$$\big((L_{g_0})^*\theta\big)_g(X) = \theta_{L_{g_0(g)}}\big((dL_{g_0})_g \cdot X\big) = \theta_{g_0 g}\big((dL_{g_0})_g \cdot X\big)\,.$$

Using 7.6, it follows

$$\big((L_{g_0})^*\theta\big)_g(X) = (dL_{(g_0 g)^{-1}})_{g_0 g} \cdot \big((dL_{g_0})_g \cdot X\big)\,.$$

From the chain rule, we then conclude

$$\begin{aligned}\big((L_{g_0})^*\theta\big)_g(X) &= d(L_{(g_0 g)^{-1}} \circ L_{g_0})_g \cdot X = (dL_{g^{-1}})_g \cdot X \\ &= \theta_g(X)\,,\end{aligned}$$

showing that θ is indeed left invariant.

Proposition 7.9. *Let G be a Lie group and V a vector space. Then $\omega \in \Omega^1(G,V)$ is left invariant if and only if there exists a linear map $f \in C^\infty(T_eG,V)$ such that*

$$\omega = f \circ \theta\,.$$

Proof. □

In opposite to usual manifolds, there exists a special class of vector fields on Lie groups characterized by an invariance under left translations.

Definition 7.10. *Let $X \in \mathcal{X}(G)$ be a vector field on a Lie group G. Then X is said to be* ***left invariant*** *if $(L_g)_*X = X$, that is, if*

$$(dL_{g_0})_g \cdot X(g) = X(g_0 g)\,, \tag{7.9}$$

for all left translations L_{g_0}, $g_0 \in G$ and all $g \in G$. We denote the vector space of left invariant vector fields by $\mathcal{X}_L(G)$.

Proposition 7.11. *Let G be a Lie group and $A \in T_eG$. Then we have that $X \in \mathcal{X}_L(G)$ if and only if $\theta_g\big(X(g)\big) = A$, for all $g \in G$.*

Proof. Let $X \in \mathcal{X}_L(M)$ be a left invariant vector field on G. Using the defining equation 7.6 for the Maurer-Cartan form, we observe that

$$\theta_g\big(X(g)\big) = (dL_{g^{-1}})_g \cdot X(g) = X(g^{-1}g) = X(e)\,.$$

Setting $X(e) = A$, this shows that for all $g \in G$, it follows $\theta_g\big(X(g)\big) = A$.

Conversely, let $X \in \mathcal{X}(M)$ satisfying $\theta_g\big(X(g)\big) = A$, for all $g \in G$, with $A \in T_eG$. First note that because of (7.7), we must have $A = X(e)$. By definition of the pull-back, we get

$$\big((L_{g^{-1}})^*\theta\big)_g\big(X(g)\big) = \theta_e\big((dL_{g^{-1}})_g \cdot X(g)\big) \overset{(7.7)}{=} (dL_{g^{-1}})_g \cdot X(g)\,.$$

Since the Maurer-Cartan form θ is left invariant, we then deduce

$$\theta_g\big(X(g)\big) = (dL_{g^{-1}})_g \cdot X(g)\,.$$

Using the assumption on $X \in \mathcal{X}(M)$, it follows $(dL_{g^{-1}})_g \cdot X(g) = X(e)$, for all $g \in G$, being equivalent to

$$X(g) = (dL_g)_e \cdot X(e)\,.$$

It is easy to check that this implies $(dL_{g_0})_g \cdot X(g) = X(g_0 g)$, showing that X is a left invariant vector field. □

The next proposition shows that every left invariant vector field corresponds to an unique element of T_eG.

Proposition 7.12. *Let G be a Lie group and let $\mathcal{X}_L(M)$ be the vector space of left invariant vector fields on G. Then the map*

$$\mathcal{X}_L(M) \longrightarrow T_eG\,, \qquad X \longmapsto X(e)$$

is an isomorphism of vector spaces. In particular, $\dim G = \dim \mathcal{X}_L(M)$.

Proof. Let $X \in \mathcal{X}_L(M)$ such that $X(e) = 0$. Since X is left invariant, we have $X(g) = dL_g \cdot X(e)$, for all $g \in G$. This implies that $X(g) = 0$, for all $g \in G$, and hence the map in the proposition is injective.

For the surjectivity, let $A \in T_eG$ and define, for all $g \in G$,

$$X_A(g) := (dL_g)_e \cdot A\,. \tag{7.10}$$

Clearly, $X_A(e) = A$. It remains to show that $X_A \in \mathcal{X}_L(M)$. For this purpose, we compute

$$\begin{aligned}\theta_g\big(X_A(g)\big) &= \theta_g\big((dL_g)_e \cdot A\big)\\ &= \big((L_g)^*\theta\big)_e(A) = \theta_e(A) = A\,,\end{aligned}$$

where we used (7.7) and (7.8) for the Maurer-Cartan form. From Proposition 7.11, we then deduce that $X_A \in \mathcal{X}_L(M)$. □

The flow of a left invariant vector field has some interesting properties. More precisely, we have

Proposition 7.13. *Let G be a Lie group and let $X \in \mathcal{X}_L(M)$ be a left invariant vector field. Then the flow Γ_t^X of X is defined on $\mathbb{R} \times G$. Moreover, the flow of X commutes with left translations, i.e., for all $g \in G$ and all $t \in \mathbb{R}$, we have*

$$L_g \circ \Gamma_t^X = \Gamma_t^X \circ L_g\,. \tag{7.11}$$

In particular, we have

$$g\Gamma_t^X(e) = \Gamma_t^X(g)\,. \tag{7.12}$$

Proof. We first establish (7.11) locally. – Let $\tilde{g} \in G$. From Section 2.7 it is well-known that there exist an open neighborhood U of $\tilde{g}$ and $0 < T \in \mathbb{R}$ such that the existence of the flow Γ_t^X on $[-T, T] \times U$ can be deduced. Then for $g_0 \in G$ consider the map

$$L_{g_0}(U) \subset G \longrightarrow G\,, \qquad g \longmapsto L_{g_0}\big(\Gamma_t^X(g_0^{-1}g)\big)\,.$$

Note that the map is well-defined since $g_0^{-1}g \in U$. We compute, using the chain rule and the defining equation for the flow,

$$\begin{aligned}\frac{\partial(L_{g_0} \circ \Gamma_t^X)}{\partial t}(g_0^{-1}g) &= dL_{g_0} \cdot \left(\frac{\partial \Gamma_t^X}{\partial t}(g_0^{-1}g)\right) \\ &= dL_{g_0} \cdot X\big(\Gamma_t^X(g_0^{-1}g)\big)\,.\end{aligned}$$

Since $X \in \mathcal{X}_L(M)$, we deduce

$$\frac{\partial(L_{g_0} \circ \Gamma_t^X)}{\partial t}(g_0^{-1}g) = X\big(L_{g_0} \circ \Gamma_t^X(g_0^{-1}g)\big)\,.$$

Moreover, it is easy to see that

$$L_{g_0} \circ \Gamma_0^X(g_0^{-1}g) = g_0\Gamma_0^X(g_0^{-1}g) = g_0(g_0^{-1}g) = g\,,$$

showing that the flow Γ_t^X of X exists also on $[-T, T] \times L_{g_0}(U)$. Note that $L_{g_0}(U)$ is open.

As a consequence, we get on $[-T, T] \times L_{g_0}(U)$ that

$$\Gamma_t^X = L_{g_0} \circ \Gamma_t^X \circ L_{g_0^{-1}}\,,$$

implying that (7.11) holds locally on U, for all $t \in [-T, T]$.

Next, we show that Γ_t^X is defined on $\mathbb{R} \times G$. – From the first part of the proof, we know that if the flow Γ_t^X of X exists on $[-T, T] \times U$, where $U \subset G$ open neighborhood of $\tilde{g}$, then the flow also exists on $[-T, T] \times U'$, where $U' = L_{g_0}(U)$ open neighborhood of g' with $g_0 = g'\tilde{g}^{-1}$. This holds for all $g' \in G$. Hence, the flow Γ_t^X is defined on $[-T, T] \times G$. This implies that it is also defined on $\mathbb{R} \times G$. Indeed, for $\tilde{t} \in [-T, T]$ such that $2|\tilde{t}| > |T|$, we deduce from $\Gamma_{\tilde{t}}^X \circ \Gamma_{\tilde{t}}^X = \Gamma_{2\tilde{t}}^X$ that $\Gamma_{2\tilde{t}}^X$ is also defined. – This proves the proposition. $\square$

Definition 7.14. *We define the* ***exponential map*** *by*

$$\begin{aligned}\exp : T_eG &\longrightarrow G\,,\\ A &\longmapsto \Gamma_1^{X_A}(e)\,,\end{aligned}$$

where X_A, defined in (7.10), is the unique left invariant vector field generated by $A \in T_eG$ and $\Gamma_t^{X_A}$ its flow.

Proposition 7.15. *Let $A \in T_eG$ and $\tilde{t}$, $t \in \mathbb{R}$. Then the exponential map satisfies*

$$\exp(tA) = \Gamma_t^{X_A}(e)\,, \tag{7.13}$$

and

$$\exp\big((t+\tilde{t})A\big) = \exp(tA)\,\exp(\tilde{t}A)\,. \tag{7.14}$$

Remark. Equation (7.14) can be seen as a reason for calling the map of Definition 7.14 "exponential" map. Moreover, we will show in Example 7.20 below that the exponential map for matrix Lie groups is simply given by the exponential of matrices.

Proof. First, we fix t and introduce a new "time" variable $\tau \in \mathbb{R}$. Then, we have

$$\frac{\partial \Gamma_{t\tau}^{X_A}}{\partial \tau}(g) = t\,X_A\big(\Gamma_{t\tau}^{X_A}(g)\big)\,,$$

and clearly $\Gamma_0^{X_A}(g) = g$. We deduce that $\Gamma_{t\tau}^{X_A}$ is the flow of the left invariant vector field tX_A satisfying $tX_A(e) = tA$. On the other hand, let $\Gamma_\tau^{X_{tA}}$ be the flow of the left invariant vector field $X_{tA} \in \mathcal{X}_L(M)$ generated by $tA \in T_eG$. From the uniqueness of the flow, we deduce that $\Gamma_\tau^{X_{tA}}(g) = \Gamma_{t\tau}^{X_A}(g)$, for all $g \in G$. In particular, for $\tau = 1$ we get, using Definition 7.14,

$$\Gamma_t^{X_A}(e) = \Gamma_1^{X_{tA}}(e) = \exp(tA)\,.$$

This shows (7.13).

By definition of the exponential map, Equation (7.14) translates to

$$\Gamma_1^{X_{(t+\tilde{t})A}}(e) = \Gamma_1^{X_{tA}}(e)\,\Gamma_1^{X_{\tilde{t}A}}(e)\,. \tag{7.15}$$

In order to show this equality, note that we have, using (7.13),

$$\Gamma_1^{X_{tA}}(e)\,\Gamma_1^{X_{\tilde{t}A}}(e) = \Gamma_t^{X_A}(e)\Gamma_{\tilde{t}}^{X_A}(e)\,.$$

Since by Proposition 7.13 the flow commutes with the left translation by $\Gamma_1^{X_{tA}}(e)$, we obtain

$$\Gamma_t^{X_A}(e)\,\Gamma_{\tilde{t}}^{X_A}(e) \overset{(7.12)}{=} \Gamma_{\tilde{t}}^{X_A}\big(\Gamma_t^{X_A}(e)\big) = \Gamma_{t+\tilde{t}}^{X_A}(e)\,,$$

where we also used a standard formula for the composition of flows. Using again (7.13) for the right-hand side, Equation (7.15) follows. □

7.2 Lie Algebras

Definition 7.16. *Let E be a $\mathbb{R}$-vector space. We call E a* ***Lie algebra*** *if there exists a bilinear map $[\cdot,\cdot] : E \times E \longrightarrow E$ which is anti-symmetric, i.e., $[e,f] = -[f,e]$, and which satisfies*

$$\big[[e,f],g\big] + \big[[g,e],f\big] + \big[[f,g],e\big] = 0\,. \tag{7.16}$$

for all $e, f, g \in E$. The last condition is often called Jacobi identity.

Example 7.17. Let M be a differentiable manifold. Then the vector space $\mathcal{X}(M)$ of C^∞-vector fields on M is a Lie algebra for the usual bracket operation on vector fields defined in Section 2.7.1.

Coming back to Lie groups we note the following important fact. – For the left translation $L_g : G \longrightarrow G$ being a diffeomorphism, we know from (2.40) that

$$(L_g)_*[X,Y] = \big[(L_g)_*X, (L_g)_*Y\big]\,,$$

for all $X, Y \in \mathcal{X}(G)$. In particular, for left invariant vector fields $X, Y \in \mathcal{X}_L(G)$, we have

$$(L_g)_*[X,Y] = \big[(L_g)_*X, (L_g)_*Y\big] = [X,Y]\,,$$

showing that the vector field $[X,Y]$ remains invariant under left translations, i.e., $[X,Y] \in \mathcal{X}_L(G)$. – As a consequence, the following definition is well-posed.

Definition 7.18. *Let G be a Lie group. The vector space $\mathcal{X}_L(M)$ of left invariant vector fields on G together with the bracket operation on vector fields is called the **Lie algebra** of the Lie group G. We denote this Lie algebra by $\mathfrak{g}$.*

Remark. Recalling Proposition 7.12 we see that the Lie algebra $\mathfrak{g}$ of a Lie group G can be identified with the tangent space T_eG of G at the identity. Thus, the previous definition can be reformulated in the following way: Let $A, B \in T_eG$ and $X_A, X_B \in \mathcal{X}_L(M)$ be the unique left invariant vector fields generated by A, respectively B. Then the vector space T_eG can be made into a Lie algebra for the bracket operation defined by

$$[A,B] = [X_A, X_B](e)\,. \tag{7.17}$$

Examples of Lie Algebras

Example 7.19 (Lie algebra of $\mathbb{R}^n$). Obviously, the Lie algebra of $\mathbb{R}^n$ can be identified with $\mathbb{R}^n$ itself. The constant vector field $X_A(v) = A$, for all $v \in \mathbb{R}^n$, generated by $A \in \mathbb{R}^n$ is left invariant. Indeed, we have

$$X_A(L_{v_0}v) = X_A(v_0 + v) = A = (dL_{v_0})_v \cdot X_A(v)\,.$$

Therefore, the Lie algebra of $\mathbb{R}^n$ is $\mathbb{R}^n$ itself together with the trivial bracket $[v,w] = 0$, for all $v, w \in \mathbb{R}^n$. Moreover, the exponential map $\exp : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is the identity map.

Example 7.20 (Lie algebra of $GL_n(\mathbb{R})$). We show that the Lie algebra of $GL_n(\mathbb{R})$, denoted by $\mathfrak{gl}_n(\mathbb{R})$, is $M_n(\mathbb{R})$. Moreover, for all $A, B \in M_n(\mathbb{R})$, we have

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

To see this, we recall that $GL_n(\mathbb{R})$ is open in $M_n(\mathbb{R})$. Thus, its Lie algebra $\mathfrak{gl}_n(\mathbb{R})$ can be identified with $M_n(\mathbb{R})$. To compute the bracket, note that for every $A \in M_n(\mathbb{R})$ the vector field $X_A(P) := PA$ on $GL_n(\mathbb{R})$ is left invariant. Indeed, for $P_0 \in GL_n(\mathbb{R})$, we have

$$X_A(L_{P_0}P) = X_A(P_0P) = P_0PA = (dL_{P_0})_P \cdot X_A(P)\,.$$

Note also that the Maurer-Cartan form reads as

$$\theta_P\big(X_A(P)\big) = (dL_{P^{-1}})_P \cdot X_A(P) = P^{-1}X_A(P) = P^{-1}PA\,.$$

Using the local formula (2.33) – written in terms of Jacobian matrices – for the bracket of vector fields, we then obtain

$$[A,B] \overset{(7.17)}{=} [X_A, X_B](\mathbb{1}) = (JX_B)_{\mathbb{1}} \cdot X_A(\mathbb{1}) - (JX_A)_{\mathbb{1}} \cdot X_B(\mathbb{1})\,.$$

By linearity of the map X_A, it follows that $(JX_A)_{\mathbb{1}} \cdot P = PA$. Hence, we have $(JX_A)_{\mathbb{1}} \cdot X_B(\mathbb{1}) = BA$. For the bracket, we then conclude

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

Next, we want to give an explicit expression for the exponential map $\exp : \mathfrak{gl}_n(\mathbb{R}) \longrightarrow GL_n(\mathbb{R})$. – Let $X_A(P) = PA$, for $P \in GL_n(\mathbb{R})$, denote the left invariant vector field generated by $A \in M_n(\mathbb{R})$. Its corresponding flow $\varGamma_t^{X_A}$ must satisfy, for all $P \in GL_n(\mathbb{R})$ and $t \in \mathbb{R}$,

$$\frac{\partial \varGamma_t^{X_A}}{\partial t}(P) = X_A\big(\varGamma_t^{X_A}(P)\big) = \varGamma_t^{X_A}(P)A\,, \qquad \varGamma_0^{X_A}(P) = P\,. \tag{7.18}$$

We fix $P_0 \in GL_n(\mathbb{R})$ and claim that

$$\varGamma_t^{X_A}(P_0) = P_0 \exp tA = P_0 \sum_{k=1}^{\infty} \frac{(tA)^k}{k!} \tag{7.19}$$

solves (7.18). Indeed, we have $P_0 \exp 0 = P_0$ and

$$\frac{d}{dt}\left(P_0 \sum_{k=1}^{\infty} \frac{(tA)^k}{k!}\right) = P_0 \sum_{k=1}^{\infty} \frac{t^{k-1}}{(k-1)!}\, A^k = P_0 \exp(tA)\, A\,.$$

Moreover, for the exponential map, we conclude

$$\exp(A) = \varGamma_1^{X_A}(\mathbb{1}) = \sum_{k=1}^{\infty} \frac{A^k}{k!}\,, \tag{7.20}$$

showing that for matrix Lie groups the exponential map equals the usual exponential of matrices.

Now, we compute again the bracket of $A, B \in \mathfrak{gl}_n(\mathbb{R})$ using the explicit expression for its flows. – From the expression (2.32) for the bracket of vector fields, we get

$$[A, B] = [X_A, X_B](\mathbb{1}) = \lim_{t \longrightarrow 0} \frac{1}{t}\Big((d\Gamma_t^{X_B})_{\Gamma_{-t}^{X_B}(\mathbb{1})} \cdot X_A\big(\Gamma_{-t}^{X_B}(\mathbb{1})\big) - X_A(\mathbb{1}) \Big) .$$

The linearity of the flow $\Gamma_t^{X_B}$ (see (7.19)) implies

$$[A, B] = \lim_{t \longrightarrow 0} \frac{1}{t}\Big(X_A\big(\Gamma_{-t}^{X_B}(\mathbb{1})\big) \exp tB - X_A(\mathbb{1}) \Big) .$$

By definition of the left invariant vector field X_A, it follows

$$[A, B] = \lim_{t \longrightarrow 0} \frac{1}{t}\Big(\Gamma_{-t}^{X_B}(\mathbb{1}) A \, \exp tB - A \Big) .$$

Using the explicit expression for the flow, we arrive at the same result as before:

$$\begin{aligned} [A, B] &= \lim_{t \longrightarrow 0} \frac{1}{t}\left(\sum_{k=1}^{\infty} \frac{(-tB)^k}{k!} A \sum_{l=1}^{\infty} \frac{(tB)^l}{l!} - A \right) \\ &= \lim_{t \longrightarrow 0} \frac{1}{t}\Big(\big(A + (-tB\,A) + A\,tB + \ldots\big) - A \Big) = AB - BA . \end{aligned}$$

Example 7.21 (Lie algebra of $O(n)$). Recall that $O(n) = g^{-1}(0)$, where the map g is explicitly given in Example 2.4. Using Proposition 2.33, it follows

$$T_{\mathbb{1}} O(n) = \ker dg_{\mathbb{1}} = \big\{ A \in M_n(\mathbb{R}) \;:\; A^T + A = 0 \big\} .$$

As a direct consequence, the Lie algebra $\mathfrak{o}(n)$ of $O(n)$ is the space of skew-symmetric $n \times n$ matrices together with the bracket as in Example 7.20.

Example 7.22 (Lie algebra of $SL(n, \mathbb{R})$). Recall from Exercise 2.5 that $SL(n, \mathbb{R}) = g^{-1}(0)$, where $g : GL_n(\mathbb{R}) \longrightarrow \mathbb{R}$ is given by $A \longmapsto \det A - 1$. For the tangent space of $SL(n, \mathbb{R})$ at $\mathbb{1}$, we have, using Proposition 2.33,

$$T_{\mathbb{1}} SL(n, \mathbb{R}) = \ker dg_{\mathbb{1}} .$$

Exercise 2.7 then implies that $T_{\mathbb{1}} SL(n, \mathbb{R})$ consists of all $n \times n$ matrices with trace zero. As a direct consequence, the Lie algebra $\mathfrak{sl}(n, \mathbb{R})$ is given by the space of all $n \times n$ matrices with trace zero together with the bracket as in Example 7.20. – Note that we also have $\mathfrak{o}(n) = \mathfrak{so}(n)$, where $\mathfrak{so}(n)$ denotes the Lie algebra of $SO(n)$.

* * *

Next, we introduce some definitions choosing a more general setting which will be also useful in the next chapters. – Let M be a differentiable manifold and let $\omega \in \Omega^1(M, \mathfrak{g})$ be a Lie algebra-valued one-form on M. Then, in a basis $E_1, \ldots, E_n$ of $\mathfrak{g}$, we write ω as

$$\omega = \sum_{i=1}^{n} \omega^i \, E_i \,,$$

where $\omega^i \in \Omega^1(M)$, for $i = 1, \ldots, n$, are real-valued one-forms on M such that $\{\omega^i(p)\}_{i=1,\ldots,n}$ form a basis of $T_p^* M$, for each $p \in M$. We call $\{\omega^i\}_{i=1,\ldots,n}$ a *coframe* of M. Moreover, we define

$$\omega \wedge \omega = \sum_{i,j=1}^{n} \omega^i \wedge \omega^j \, [E_i, E_j] \,, \tag{7.21}$$

being an element of $\Omega^2(M, \mathfrak{g})$. – It is easy to check that $\omega \wedge \omega$ is independent of the choice of the basis for $\mathfrak{g}$.

* * *

Coming back to Lie groups, there is the following important

Definition 7.23. *Let G be a Lie group with finite dimensional Lie algebra $\mathfrak{g}$. Moreover, let $E_1, \ldots, E_n$ denote a basis of $\mathfrak{g}$. We define the **structure coefficients** $c_{ij}^k \in \mathbb{R}$ of $\mathfrak{g}$ with respect to the basis $E_1, \ldots, E_n$ by*

$$[E_i, E_j] = \sum_{k=1}^{n} c_{ij}^k \, E_k \,. \tag{7.22}$$

Proposition 7.24. *Let G be a Lie group and $\mathfrak{g}$ its Lie algebra. Then for the Maurer-Cartan form $\theta \in \Omega^1(G, \mathfrak{g})$ the following equation holds on G:*

$$d\theta + \frac{1}{2} \theta \wedge \theta = 0 \,. \tag{7.23}$$

*This equation is called **Maurer-Cartan's structure equation**.*

Proof. Recall that the Maurer-Cartan form θ is left invariant, i.e., $L_{g_0}^* \theta = \theta$, for all $g_0 \in G$. Since the pull-back and the exterior differentiation commute, it follows that $L_{g_0}^*(d\theta) = d(L_{g_0}^* \theta) = d\theta$, showing that $d\theta$ is also left invariant. Obviously, the same holds for $\theta \wedge \theta$. Thus, it suffices to show the Maurer-Cartan equation at $e \in G$, i.e.,

$$d\theta(e) + \frac{1}{2} \theta \wedge \theta(e) = 0 \,.$$

Let A and $B \in \mathfrak{g}$ with corresponding left invariant vector fields X_A and $X_B \in \mathcal{X}_L(M)$ (see Proposition 7.12). We deduce from (6.50) that

$$d\theta_e(A, B) = d\big(\theta(X_B)\big)_e \cdot A - d\big(\theta(X_A)\big)_e \cdot B - \theta_e\big([X_A, X_B](e)\big) \,.$$

Since by Proposition 7.11 the functions $\theta(X_A)$ and $\theta(X_B)$ are constant on G, we obtain

$$d\theta_e(A,B) = -\theta_e\big([X_A, X_B](e)\big)\,. \tag{7.24}$$

Now, let $E_1,\ldots,E_n$ be a basis of the Lie algebra $\mathfrak{g}$. We then write the constants $\theta(X_A)$ and $\theta(X_B)$ as $\theta(X_A) = \sum_{i=1}^n A^i\, E_i$, respectively as $\theta(X_B) = \sum_{i=1}^n B^i\, E_i$, where $A^i, B^i \in \mathbb{R}$. On the other hand, it is clear that $\theta(X_A) = X_A(e) = A$ and $\theta(X_B) = X_B(e) = B$. For the right-hand side of (7.24), we get by linearity

$$\theta_e\big([X_A, X_B](e)\big) = [A,B] = \sum_{i<j=1}^{n} (A^i B^j - A^j B^i)\,[E_i, E_j]\,. \tag{7.25}$$

Next, writing $\theta = \sum_{i=1}^n \theta^i\, E_i$, where $\{\theta^i\}_{i=1,\ldots,n}$ is the so-called *Maurer-Cartan coframe*[2] on G, we observe that $\theta^i(X_A) = A^i$, since

$$\theta(X_A) = \sum_{i=1}^{n} \theta^i(X_A)\, E_i = \sum_{i=1}^{n} A^i\, E_i\,.$$

Inserting this in (7.25), we deduce

$$\sum_{i<j=1}^{n} (A^i B^j - A^j B^i)\,[E_i, E_j] = \sum_{i<j=1}^{n} \theta^i \wedge \theta^j (X_A, X_B)\,[E_i, E_j]\,.$$

With (7.24) it then follows

$$d\theta = -\sum_{i<j=1}^{n} \theta^i \wedge \theta^j\,[E_i, E_j]\,.$$

From (7.21), we then deduce Maurer-Cartan's structure equation. □

Remark. In a basis $\{E_1,\ldots,E_n\}$ of $\mathfrak{g}$ Maurer-Cartan's structure equation becomes

$$d\theta^k = -\sum_{i<j=1}^{n} c_{ij}^k\, \theta^i \wedge \theta^j\,. \tag{7.26}$$

Using (7.22) and writing $d\theta = \sum_{k=1}^n d\theta^k\, E_k$, this follows directly from (7.23).

* * *

Exercises.

Exercise 7.25. Show that the structure coefficients are skew-symmetric, i.e., $c_{ij}^k = -c_{ij}^k$, and satisfy the Jacobi identity, i.e., for all $i,j,l,m \in \{1,\ldots,n\}$,

$$\sum_{k=1}^{n} (c_{ij}^k\, c_{kl}^m + c_{li}^k\, c_{kj}^m + c_{jl}^k\, c_{ki}^m) = 0\,.$$

[2] Obviously, $\theta^i \in \Omega^1(G)$ are left invariant, for $i = 1,\ldots,n$.

Exercise 7.26 (Lie algebra of $SU(n)$). With the help of Example 7.21 and 7.22 show that the Lie algebra $\mathfrak{su}(n)$ of the special unitary group $SU(2)$ defined in (7.5) is given by

$$\mathfrak{su}(n) = \big\{A \in M_n(\mathbb{C}) \ : \ A^\dagger + A = 0 \ \text{ and } \ \mathrm{Tr}(A) = 0\big\}\,,$$

with bracket $[A,B] = AB - BA$, for $A, B \in \mathfrak{su}(n)$.

7.3 Lie Groups Acting on Manifolds

Definition 7.27. *Let G be a Lie group and M a differentiable manifold. We say that G is a **Lie transformation group** on M or that G is a **(right) action** on M if there exists a differentiable map*

$$\mathcal{A} : G \times M \longrightarrow M$$

*such that $\mathcal{A}(g,\cdot) : M \longrightarrow M$ is a diffeomorphism, for all $g \in G$, and such that $\mathcal{A}(g\tilde{g},p) = \mathcal{A}\big(\tilde{g},\mathcal{A}(g,p)\big)$, for all $g, \tilde{g} \in G$, $p \in M$. Moreover, a **(left) action** must satisfy $\mathcal{A}(\tilde{g}g,p) = \mathcal{A}\big(\tilde{g},\mathcal{A}(g,p)\big)$.*

Remark. For a right action, we will often write pg instead of $\mathcal{A}(g,p)$ and $R_g : M \longrightarrow M$ for the map $\mathcal{A}(g,\cdot)$. Note also that R_e is the identity transformation on M, i.e., $R_e = id_M$. In fact, since by definition we have $R_e p = pe = p(ee) = (pe)e = R_e(R_e p)$, it follows that

$$p = (R_e)^{-1} R_e(R_e p) = R_e p\,.$$

Definition 7.28. *Let G be a right action on a manifold M. We say that G acts **freely** on M if $R_g p = p$, for some $p \in M$, implies that $g = e$. In other words, the action G is free if $e \in G$ is the only element in G having a fixed point in M. Moreover, we say that G acts **effectively** on M if $R_g p = p$, for all $p \in M$, implies that $g = e$.*

Next, we introduce a particular vector field on M which will be important for Section 9.92. – Let G act on M on the right. Then we assign to each $A \in \mathfrak{g}$ the vector field $A^* \in \mathcal{X}(M)$ defined by

$$A^*(p) = \left.\frac{d}{dt}\right|_{t=0} R_{\exp tA}\, p = \left.\frac{d}{dt}\right|_{t=0} p \exp tA\,, \tag{7.27}$$

for all $p \in M$ and $t \in \mathbb{R}$. – Recall Definition 7.14 for the exponential map.

The vector field $A^* \in \mathcal{X}(M)$ can also be defined in the following manner: For every $p \in M$ consider the map

$$\begin{aligned} \sigma_p : G &\longrightarrow M\,, \\ g &\longmapsto pg\,. \end{aligned} \tag{7.28}$$

Obviously, we have $\sigma_p = \mathcal{A}(\cdot, p)$. The differential of σ_p at e – which we will simply denote by $d\sigma_p$ – applied to the left invariant vector field $X_A \in \mathcal{X}_L(G)$ generated by $A \in \mathfrak{g}$ reads as

$$\begin{aligned} d\sigma_p \cdot X_A(e) &= d\sigma_p \cdot \frac{d}{dt}\bigg|_{t=0} \Gamma_t^{X_A}(e) = \frac{d}{dt}\bigg|_{t=0} p\,\Gamma_t^{X_A}(e) \\ &= \frac{d}{dt}\bigg|_{t=0} p\exp(tA) = \frac{d}{dt}\bigg|_{t=0} R_{\exp tA}\, p = A^*(p)\,. \end{aligned} \tag{7.29}$$

Next, we define a map from $\mathfrak{g}$ into $\mathcal{X}(M)$, also denoted by σ, as follows:

$$\begin{aligned} \sigma : \mathfrak{g} &\longrightarrow \mathcal{X}(M)\,, \\ A &\longmapsto A^*\,. \end{aligned} \tag{7.30}$$

Proposition 7.29. *Let a Lie group G act on M on the right. Then, the map $\sigma : \mathfrak{g} \longrightarrow \mathcal{X}(M)$ is a Lie algebra homomorphism. Moreover, if G acts freely on M, then $\sigma(A)$ never vanishes on M, whenever $A \in \mathfrak{g}$ is non-zero.*

Proof. In this proof, we will use the notation $a_t = \exp tA \in G$ and note that the linearity is clear since $\sigma(A)_p = A^*(p) = d\sigma_p \cdot A$ by (7.29). – First, we determine the flow corresponding to $A^* \in \mathcal{X}(M)$. We observe that

$$\frac{d}{dt}\bigg|_{t=t_0} R_{a_t} p = \frac{d}{dt}\bigg|_{t=0} R_{a_{t+t_0}} p\,.$$

As a consequence of Definition 7.27 and (7.14), we also have that

$$R_{a_{t+t_0}} p = p\exp\big((t+t_0)A\big) = p\big(\exp(t_0 A)\,\exp(tA)\big) = R_{a_t}(p\,a_{t_0})\,.$$

This gives, using (7.27),

$$\frac{d}{dt}\bigg|_{t=t_0} R_{a_t} p = \frac{d}{dt}\bigg|_{t=0} R_{a_t}(p\,a_{t_0}) = A^*(p\,a_{t_0}) = A^*(R_{a_{t_0}} p)\,.$$

It follows that $R_{a_t} : M \longrightarrow M$ is the flow of A^* (note also that $R_{a_0} p = p$).

Now, we show that the map σ commutes with the brackets of the Lie algebras $\mathfrak{g}$ and $\mathcal{X}(M)$ (see (7.17) and Example 7.17). – Let $A, B \in \mathfrak{g}$ and $A^* = \sigma(A)$, $B^* = \sigma(B)$. Using formula (2.32) for the bracket of vector fields and the fact that R_{b_t} is the flow of B^*, we obtain that

$$[A^*, B^*](p) = \lim_{t \longrightarrow 0} \frac{1}{t}\Big(\big(dR_{b_t}\big)_{R_{b_{-t}} p} \cdot A^*\big(R_{b_{-t}} p\big) - A^*(p)\Big)\,. \tag{7.31}$$

Using (7.29), we observe that

$$\begin{aligned} dR_{b_t} \cdot A^*(R_{b_{-t}} p) &= dR_{b_t} \cdot A^*(p b_{-t}) \\ &= dR_{b_t} \cdot d\sigma_{pb_{-t}} \cdot X_A(e) = d(R_{b_t} \circ \sigma_{pb_{-t}})_e \cdot X_A(e)\,. \end{aligned}$$

Note that in terms of conjugation $c_{b_{-t}} : G \longrightarrow G$ defined in Example 7.30 below, we can write

$$R_{b_t} \circ \sigma_{pb_{-t}}(g) = p\, b_{-t} g b_t = \sigma_p \circ c_{b_{-t}}(g)\,.$$

With the adjoint action defined in (7.37) below, we then arrive at

$$\begin{aligned} dR_{b_t} \cdot A^*(R_{b_{-t}} p) &= d(\sigma_p \circ c_{b_{-t}})_e \cdot X_A(e) \\ &= d\sigma_p \cdot (dc_{b_{-t}})_e \cdot X_A(e) = d\sigma_p \cdot Ad_{b_{-t}}\big(X_A(e)\big)\,. \end{aligned} \tag{7.32}$$

Inserting this into (7.31), we deduce

$$[A^*, B^*](p) = \lim_{t \longrightarrow 0} \frac{1}{t}\Big(d\sigma_p \cdot \Big(Ad_{b_{-t}}\big(X_A(e)\big) - X_A(e)\Big)\Big)\,.$$

Next, we compute the following limit:

$$\begin{aligned} \lim_{t \longrightarrow 0} \frac{1}{t}\Big(Ad_{b_{-t}}\big(X_A(e)\big) - X_A(e)\Big) &= \lim_{t \longrightarrow 0} \frac{1}{t}\Big(dR_{b_t} \cdot dL_{b_{-t}} \cdot X_A(e) - X_A(e)\Big) \\ &= \lim_{t \longrightarrow 0} \frac{1}{t}\Big(dR_{b_t} \cdot X_A(b_{-t}) - X_A(e)\Big) \\ &= [X_A, X_B](e)\,, \end{aligned}$$

where we used that X_A is left invariant and that X_B has the flow R_{b_t}. (The last statement is a direct consequence of (7.34) below.) From this we finally deduce that

$$\begin{aligned} \big[\sigma(A), \sigma(B)\big](p) &= [A^*, B^*](p) \\ &= d\sigma_p \cdot \Big(\lim_{t \longrightarrow 0} \frac{1}{t}\Big(Ad_{b_{-t}}\big(X_A(e)\big) - X_A(e)\Big)\Big) \\ &= d\sigma_p \cdot [X_A, X_B](e) \overset{(7.29)}{=} \sigma\big([A,B]\big)(p)\,. \end{aligned}$$

Hence, we have shown that $\sigma : \mathfrak{g} \longrightarrow \mathcal{X}(M)$ is a Lie algebra homomorphism.

To prove the last assertion of the proposition, we assume that $\sigma(A)$ vanishes at some point $p_0 \in M$ and claim that this implies

$$R_{a_t} p_0 = p_0\,,$$

for all $t \in \mathbb{R}$. – Since $R_{a_t} : M \longrightarrow M$ is the flow of the vector field $A^* \in \mathcal{X}(M)$, we have

$$\left.\frac{d}{dt}\right|_{t=0} R_{a_t} p_0 = A^*(R_{a_0} p_0) = A^*(p_0) = 0\,.$$

Solving the resulting ordinary differential equation with this initial data in a chart the claim follows directly. Thus, if G acts freely on M, we conclude that $a_t = e$, for all $t \in \mathbb{R}$, and moreover A^* vanishes on M as direct consequence of (7.27). □

* * *

We consider the particular case of a Lie group acting on itself. So, let G act on G on the right. Then for each $A \in \mathfrak{g}$ the vector field $A^* \in \mathcal{X}(G)$ is given by

$$A^*(g) = \left.\frac{d}{dt}\right|_{t=0} R_{\exp tA}\, g = \left.\frac{d}{dt}\right|_{t=0} g \exp tA\,. \tag{7.33}$$

Now, we claim that

$$X_A(g) = A^*(g)\,, \tag{7.34}$$

where $X_A \in \mathcal{X}_L(G)$ defined in (7.10) is the unique left invariant vector field generated by A. In order to prove the claim, we compute

$$\begin{aligned} X_A(g) &= (dL_g)_e \cdot A = (dL_g)_e \cdot X_A(e) = (dL_g)_e \cdot \left.\frac{d}{dt}\right|_{t=0} \varGamma_t^{X_A}(e) \\ &= \left.\frac{d}{dt}\right|_{t=0} g\, \varGamma_t^{X_A}(e) = \left.\frac{d}{dt}\right|_{t=0} g \exp(tA) \overset{(7.33)}{=} A^*(g)\,. \end{aligned}$$

Remark. Note that this calculation is completely similar to the one in (7.29), which comes from the fact that the map defined in (7.28) translates to left translation on G in the particular case of G acting on itself. We also emphasize that we need a *right* action of G on itself in order to obtain a *left* invariant vector field.

Example 7.30. There is a particular (left) action of a Lie group G on itself.
– Let $g_0 \in G$ and define the *conjugation* c_{g_0} by g_0 to be the map

$$\begin{aligned} c_{g_0} : G &\longrightarrow G\,, \\ g &\longmapsto L_{g_0} \circ R_{g_0^{-1}}(g) = g_0 g g_0^{-1}\,. \end{aligned} \tag{7.35}$$

Clearly, the conjugation is a homomorphism. Next, we consider the map

$$\begin{aligned} c : G \times G &\longrightarrow G\,, \\ (g_0, g) &\longmapsto c_{g_0}(g) \end{aligned} \tag{7.36}$$

being a left action of G on itself in the sense of Definition 7.27. Noting that $c_{g_0}(e) = e$ and restricting the tangent map dc_{g_0} to T_eG – which we identify with the Lie algebra $\mathfrak{g}$ – we obtain the following (left) action of G on the Lie algebra $\mathfrak{g}$:

$$\begin{aligned} Ad : G \times \mathfrak{g} &\longrightarrow \mathfrak{g}\,, \\ (g_0, A) &\longmapsto (dc_{g_0})_e \cdot A\,. \end{aligned} \tag{7.37}$$

This is the so-called **adjoint action** and we will also write $Ad_{g_0} : \mathfrak{g} \longrightarrow \mathfrak{g}$ for the diffeomorphism $(dc_{g_0})_e$.

* * *

Exercises.

Exercise 7.31. Show that in the case of a matrix Lie group the diffeomorphism $Ad_{P_0} : \mathfrak{g} \longrightarrow \mathfrak{g}$ is simply given by

$$Ad_{P_0}(A) = P_0 A P_0^{-1} \,. \tag{7.38}$$

8 Distributions and Frobenius' Theorem

Consider a vector field X on a manifold M and assume that X vanishes nowhere. The map

$$p \longmapsto \operatorname{span}\{X(p)\}$$

assigns to each point $p \in M$ a one-dimensional subspace of T_pM. Then the integral curves of X – which always exist (locally) – are one-dimensional submanifolds of M being everywhere tangent to these given subspaces of T_pM.

Instead of an one-dimensional subspace we now assign a k-dimensional subspace of T_pM to each point $p \in M$. The following question arises: Is it possible to find k-dimensional submanifolds of M being everywhere tangent to these subspaces? Note that, in general, such submanifolds do not exist, even locally. In Section 8.2 we will give necessary and sufficient criteria guaranteeing the existence of such submanifolds. – We begin with precisions in the following section.

8.1 Distributions on Manifolds

Let M^n be an n-dimensional differentiable manifold. From a "set" point of view, we define

$$G_k(TM) = \bigcup_{p \in M} G_k(T_pM)\,, \tag{8.1}$$

where $G_k(T_pM)$ denotes the set of k-dimensional subspaces of T_pM. In Appendix 9.92 we show that $G_k(TM)$ can be made to a differentiable manifold, called the **Grassmannian bundle**. For $\mathcal{D}(p) \in G_k(T_pM)$, the projection map to the base point reads as

$$\begin{aligned} \pi : G_k(M) &\longrightarrow M\,, \\ \mathcal{D}(p) &\longmapsto p\,. \end{aligned} \tag{8.2}$$

Note that the following concept appears at various places in previous chapters: Starting with some bundle, we construct a natural object corresponding to the bundle – called section – which is equivalently characterized by some particular local representation. For example, take the cotangent bundle and the differential one-forms. This concept also applies to the Grassmannian bundle. More precisely, we have

Definition 8.1. *Let M^n be a differentiable manifold of dimension n and $k \leq n$. A k-dimensional* ***distribution*** *on M is a map*

$$\mathcal{D} : M \longrightarrow G_k(M)$$

such that $\pi \circ \mathcal{D} = id_M$, and for all local charts (U, x) on M^n there exists a function $\tilde{\mathcal{D}} \in C^\infty(U, G_k(\mathbb{R}^n))$ satisfying

$$\Phi_p^k\big(\mathcal{D}(p)\big) = \tilde{\mathcal{D}}(p) \tag{8.3}$$

for all $p \in U$, where $\Phi_p^k : G_k(T_pM) \longrightarrow G_k(\mathbb{R}^n)$ is defined by

$$\begin{aligned}&\Phi_p^k\left(span\left\{\sum_{i=1}^n a_{i1}\,\frac{\partial}{\partial x_i}(p), \dots, \sum_{i=1}^n a_{ik}\,\frac{\partial}{\partial x_i}(p)\right\}\right)\\ &= span\left\{\sum_{i=1}^n a_{i1}\,e_i, \dots, \sum_{i=1}^n a_{ik}\,e_i\right\}\end{aligned}$$

with $\{e_i\}_{i=1,\dots,n}$ the canonical basis of $\mathbb{R}^n$ and $a_{ij} \in \mathbb{R}$, for $j = 1 \dots, k$.

Lemma 8.2. *Let $\mathcal{D}$ be a k-dimensional distribution on M. Then, for all $p \in M$, there exists $U \subset M$ open containing p, and*

(i) smooth, linearly independent one-forms $\omega^{k+1}, \dots, \omega^n$ on U such that

$$\mathcal{D}(q) = \bigcap_{j=k+1}^n ker\,\omega^j(q)\,, \tag{8.4}$$

for $q \in U$.

(ii) smooth, linearly independent vector fields $X_1, \dots, X_k$ on U such that

$$span\big\{X_1(q), \dots, X_k(q)\big\} = \mathcal{D}(q)\,, \tag{8.5}$$

for $q \in U$.

Proof. Let $p \in M^n$. We choose a chart (U, x) about p such that the k-dimensional distribution at p can be written as

$$\mathcal{D}(p) = \text{span}\left\{\frac{\partial}{\partial x_1}(p), \dots, \frac{\partial}{\partial x_k}(p)\right\} \subset T_pM\,,$$

and thus by definition

$$\tilde{\mathcal{D}}(p) = \text{span}\{e_1, \dots, e_k\} \subset \mathbb{R}^n\,.$$

Let $\iota_{\tilde{\mathcal{D}}(p)} : \tilde{\mathcal{D}}(p) \hookrightarrow \mathbb{R}^n$ denote the canonical inclusion and observe that $\iota^*_{\tilde{\mathcal{D}}(p)} e^1 \wedge \ldots \wedge e^k \neq 0$. By continuity of $\tilde{\mathcal{D}} : U \longrightarrow G_k(\mathbb{R}^n)$, we deduce that there exists a neighborhood $V \subset U$ of p such that $\iota^*_{\tilde{\mathcal{D}}(q)} e^1, \ldots, \iota^*_{\tilde{\mathcal{D}}(q)} e^k$ are linearly independent in V, i.e.,

$$\iota^*_{\tilde{\mathcal{D}}(q)} e^1 \wedge \ldots \wedge e^k = \iota^*_{\tilde{\mathcal{D}}(q)} e^1 \wedge \ldots \wedge \iota^*_{\tilde{\mathcal{D}}(q)} e^k \neq 0\,,$$

for all $q \in V$. Since $\dim \tilde{\mathcal{D}}(q) = k$, it follows that $\{\iota^*_{\tilde{\mathcal{D}}(q)} e^i\}_{i=1,\ldots,k}$ forms a basis for the dual $\big(\tilde{\mathcal{D}}(q)\big)^*$ of $\tilde{\mathcal{D}}(q) \subset \mathbb{R}^n$. Hence, we can write, for $j = k+1, \ldots, n$,

$$\iota^*_{\tilde{\mathcal{D}}(q)} e^j = \sum_{i=1}^{k} \lambda_i^j(q)\, \iota^*_{\tilde{\mathcal{D}}(q)} e^i\,. \tag{8.6}$$

Note that since $\tilde{\mathcal{D}} \in C^\infty(V, G_k(\mathbb{R}^n))$, we clearly have that $\lambda_i^j \in C^\infty(V, \mathbb{R})$. By linearity of the pull-back, (8.6) can be written as

$$\iota^*_{\tilde{\mathcal{D}}(q)} \left(e^j - \sum_{i=1}^{k} \lambda_i^j(q)\, e^i \right) = 0\,.$$

Recalling that the last expression is an element of $\big(\tilde{\mathcal{D}}(q)\big)^*$, we obtain that

$$\bigcap_{j=k+1}^{n} \ker \left(e^j - \sum_{i=1}^{k} \lambda_i^j(q)\, e^i \right) \subset \tilde{\mathcal{D}}(q)\,. \tag{8.7}$$

Moreover, it is not difficult to check that

$$\dim \bigcap_{j=k+1}^{n} \ker \left(e^j - \sum_{i=1}^{k} \lambda_i^j(q)\, e^i \right) = n - (n-k) = k\,.$$

Hence, we deduce from (8.7) that

$$\tilde{\mathcal{D}}(q) = \bigcap_{j=k+1}^{n} \ker \left(e^j - \sum_{i=1}^{k} \lambda_i^j(q)\, e^i \right)\,,$$

for all $q \in V$. As a direct consequence, it then follows that

$$\mathcal{D}(q) = \bigcap_{j=k+1}^{n} \ker \left(dx^j(q) - \sum_{i=1}^{k} \lambda_i^j(q)\, dx^i(q) \right)\,, \tag{8.8}$$

for all $q \in V$. Defining smooth, linearly independent one-forms ω^j, for $j = k+1, \ldots, n$, on V by

$$\omega^j = dx^j - \sum_{i=1}^{k} \lambda_i^j \, dx^i \,, \tag{8.9}$$

we see that the first part of the proposition is proved.

For the second part, we define smooth vector fields X_i, $i = 1, \ldots, k$, on V by

$$X_i = \frac{\partial}{\partial x_i} + \sum_{j=k+1}^{n} \lambda_i^j \, \frac{\partial}{\partial x_j} \,. \tag{8.10}$$

Note that since $\lambda_i^j \in C^\infty(V, \mathbb{R})$ (see (8.6)) these vector fields are linearly independent. Moreover, we have, using (8.9), for $j = k+1, \ldots, n$, that

$$\begin{aligned}
\omega^j(X_l) &= dx^j - \sum_{i=1}^{k} \lambda_i^j \, dx^i \left(\frac{\partial}{\partial x_l} + \sum_{m=k+1}^{n} \lambda_l^m \, \frac{\partial}{\partial x_m} \right) \\
&= \sum_{m=k+1}^{n} \lambda_l^m \, dx^j \left(\frac{\partial}{\partial x_m} \right) - \sum_{i=1}^{k} \lambda_i^j \, dx^i \left(\frac{\partial}{\partial x_l} \right) \\
&= \lambda_l^j - \lambda_l^j = 0 \,.
\end{aligned}$$

This implies that, for $q \in V$ and $i = 1, \ldots, k$,

$$X_i(q) \in \bigcap_{j=k+1}^{n} \ker \omega^j(q) = \mathcal{D}(q) \,,$$

showing the second part of the proposition. $\square$

Remark. Since $\mathcal{D}$ must be a solution of the system of partial differential equations defined in (8.4) by the vanishing of a collection of differential forms, it is also often called **differential system**.

In the next section, it will be useful to work not only with the one-forms $\omega^{k+1}, \ldots, \omega^n$ of the previous lemma but with *all* differential forms having a vanishing restriction to $\mathcal{D}$. More precisely, we associate to each k-dimensional distribution $\mathcal{D}$ the set $\mathcal{I}(\mathcal{D}) \subset \Omega(M)$ defined by

$$\mathcal{I}(\mathcal{D}) = \left\{ \omega \in \Omega(M) \;:\; \iota^*_{\mathcal{D}(p)} \omega(p) = 0 \,, \forall p \in M \right\} , \tag{8.11}$$

i.e., all differential forms in the exterior algebra $\Omega(M)$ whose restriction to $\mathcal{D}$ vanishes for all $p \in M$. The set $\mathcal{I}(\mathcal{D})$ is called the *annihilator*[1] of the distribution $\mathcal{D}$. It is easy to verify that $\mathcal{I}(\mathcal{D})$ is *locally generated* by $\omega^{k+1}, \ldots, \omega^n$. This means that every $\omega \in \mathcal{I}(\mathcal{D})$ can be expressed on U as

[1] Note that $\mathcal{I}(\mathcal{D})$ is an ideal of the exterior algebra $\Omega(M)$, i.e., if $\omega_1, \omega_2 \in \mathcal{I}(\mathcal{D})$, then $\omega_1 + \omega_2 \in \mathcal{I}(\mathcal{D})$; and if $\omega \in \mathcal{I}(\mathcal{D})$, $\tilde{\omega} \in \Omega(M)$, then $\omega \wedge \tilde{\omega} \in \mathcal{I}(\mathcal{D})$.

$$\omega = \sum_{i=k+1}^{n} \theta^i \wedge \omega^i \,, \tag{8.12}$$

for some smooth differential forms θ^i on U. Moreover, we say that the annihilator is a *differential ideal* if whenever $\omega \in \mathcal{I}(\mathcal{D})$, then also $d\omega \in \mathcal{I}(\mathcal{D})$. This will be often written as $d\mathcal{I}(\mathcal{D}) \subset \mathcal{I}(\mathcal{D}) \subset$.

8.2 Frobenius' Theorem

In this section, we will give an answer to the question at the beginning of this chapter. This will be first done in the language of vector fields.

Vector Field Version of Frobenius' Theorem

Definition 8.3. *Let $\mathcal{D}$ be a k-dimensional distribution on a n-dimensional manifold M^n. A k-dimensional submanifold N^k of M^n is called an **integral manifold** of $\mathcal{D}$ if for every $p \in N^k$, we have*

$$(d\iota_N)_p(T_pN^k) = \mathcal{D}(p) \,, \tag{8.13}$$

*where $\iota_N : N^k \hookrightarrow M^n$ is the inclusion map. Moreover, we say that the distribution $\mathcal{D}$ on M is **integrable** if an integral manifold passes through each point $p \in M$. These integral manifolds will be denoted by $N(p)$, $p \in M$.*

We want to find out when a distribution is integrable. – For this purpose, we consider vector fields which belong to a distribution $\mathcal{D}$. More precisely, we say that a vector field $X \in \mathcal{X}(M)$ *belongs to* $\mathcal{D}$ if $X(p) \in \mathcal{D}(p)$, for all $p \in M$. We then write $X \in \mathcal{D}$.

In a first step, we suppose that the distribution $\mathcal{D}$ is integrable with integral manifolds $N(p)$, $p \in M$. Let $X, Y \in \mathcal{D}$ be two vector fields belonging to $\mathcal{D}$. Since $(d\iota_{N(p)})_q : T_qN(p) \longrightarrow \mathcal{D}(q)$ is an isomorphism for every $q \in N(p)$ by assumption, there exist smooth unique vector fields $\bar{X}$ and $\bar{Y}$ on $N(p)$ such that $X(q) = (d\iota_{N(p)})_q \cdot \bar{X}(q)$ and $Y(q) = (d\iota_{N(p)})_q \cdot \bar{Y}(q)$, for all $q \in N(p)$. (Note that the smoothness of $\bar{X}$ and $\bar{Y}$ is not difficult to show.) From Proposition 2.55, we deduce that

$$(d\iota_{N(p)})_q \cdot [\bar{X}, \bar{Y}](q) = [X, Y](q)$$

Since $[\bar{X}, \bar{Y}](q) \in T_qN(p)$, we conclude that $[X, Y](q) \in \mathcal{D}(q)$, for all $q \in N(p)$. – This suggests to introduce the following concept:

Definition 8.4. *Let $X, Y \in \mathcal{D}$. The k-dimensional distribution $\mathcal{D}$ on M is then said to be **involutive** if the bracket of X and Y also belongs to $\mathcal{D}$, i.e., if $[X, Y] \in \mathcal{D}$.*

Using this terminology, we have previously shown that if a distribution $\mathcal{D}$ is integrable then it is involutive. – It turns out that the converse is also true.

Theorem 8.5 (Frobenius' Theorem (Vector Field Version)). *Let $\mathcal{D}$ be a k-dimensional distribution on a manifold M. Assume that $\mathcal{D}$ is involutive. Then the distribution $\mathcal{D}$ is integrable.*

More precisely, the following holds: For every point $p \in M$, there exists a chart (U, x) about p, an interval $I \subset \mathbb{R}$ around 0, such that $x(p) = 0$, $x(U) = I^n$, and for any $a_{k+1}, \dots, a_n \in I$, the slice

$$N_{a_{k+1},\dots,a_n}(p) := \{q \in U \ : \ x_{k+1}(q) = a_{k+1}, \dots, x_n(q) = a_n\} \tag{8.14}$$

is an integral manifold of $\mathcal{D}$. Moreover, any connected integral manifold of $\mathcal{D}$ contained in U is of this form.

Proof. The statement being a local one, we can work in $\mathbb{R}^n$ and choose $p = 0 \in \mathbb{R}^n$. Moreover, we can assume that $\mathcal{D}(0) \subset T_0\mathbb{R}^n$ is spanned by

$$\left\{\frac{\partial}{\partial x_1}(0), \dots, \frac{\partial}{\partial x_k}(0)\right\} ,$$

where $(x_1, \dots, x_n)$ denote the standard coordinate functions on $\mathbb{R}^n$.

Let $\pi : \mathbb{R}^n \longrightarrow \mathbb{R}^k$ be the canonical projection. Then, we see that

$$d\pi|_{\mathcal{D}(0)} : \mathcal{D}(0) \longrightarrow \mathbb{R}^k$$

is an isomorphism. By continuity, the tangent map $d\pi$ remains an isomorphism on $\mathcal{D}(q)$, for all $q \in \mathbb{R}^n$ in some neighborhood $\tilde{U}$ of the point $0 \in \mathbb{R}^n$. It follows that there are unique vector fields $X_1, \dots, X_k$ on $\tilde{U}$ belonging to $\mathcal{D}$ such that, for $i = 1, \dots, k$,

$$(d\pi)_q \cdot X_i(q) = \frac{\partial}{\partial x_i}\big(\pi(q)\big) .$$

Using Proposition 2.55, we then obtain

$$(d\pi)_q \cdot [X_i, X_j](q) = \left[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right]\big(\pi(q)\big) . \tag{8.15}$$

The right-hand side vanishes as a consequence of the "Double" Straightening Theorem 2.58 applied to $\pi(\tilde{U}) \subset \mathbb{R}^k$. Since $\mathcal{D}$ was assumed to be involutive, we have $[X_i, X_j] \in \mathcal{D}$, thus $[X_i, X_j](q) = 0$, for all $q \in \tilde{U}$, using (8.15). According to a slight generalization of the "double" straightening theorem , there exists a chart (U, y) about $0 \in \mathbb{R}^n$ with $y(U) = I^n$ and

$$X_i(q) = \frac{\partial}{\partial y_i}(q) , \qquad q \in U , \tag{8.16}$$

for some interval $I \subset \mathbb{R}$ around 0 and $i = 1, \ldots, k$. – Now, it is not difficult to see that the slices

$$\{q \in U \ : \ y_{k+1}(q) = a_{k+1}, \ldots, y_n(q) = a_n\}$$

are integral manifolds of $\mathcal{D}$, since their tangent spaces are spanned by (8.16).

Next, we show that any *connected* integral manifold contained in U is of the form (8.14). For this purpose, suppose $N(p)$ is an integral manifold of $\mathcal{D}$ contained in U. For $X(q) \in T_q N(p)$ and $i = k+1, \ldots, n$, we then have

$$d(x_i \circ \iota_{N(p)})_q \cdot X(q) = dx_i\big((d\iota_{N(p)})_q \cdot X(q)\big) = 0\,,$$

using the fact that by assumption $(d\iota_{N(p)})_q \cdot X(q)$ is an element of $\mathcal{D}(q)$, which is spanned by the vectors in (8.16). Thus $d(x_i \circ \iota_{N(p)})_q = 0$ for every $q \in N(p)$, implying that $x_i \circ \iota_N$ is constant on the connected manifold $N(p)$. In other words, we have $N(p) = N_{a_{k+1},\ldots,a_n}(p)$, for some $a_{k+1}, \ldots, a_n \in I$. – This completes the proof of Frobenius' theorem in the vector field version.

□

Differential Form Version of Frobenius' Theorem

Next, we reformulate Frobenius' theorem in terms of differential forms. – For this purpose, we first translate Definition 8.3 in the language of differential forms.

Proposition 8.6. *Let $\mathcal{D}$ be a k-dimensional distribution on a manifold M. Then $\mathcal{D}$ is involutive if and only if the annihilator $\mathcal{I}(\mathcal{D})$ is a differential ideal.*

Proof. From the proof of Lemma 8.2, we know that the smooth, linearly independent one-forms $\omega^{k+1}, \ldots, \omega^n$ on U defined in (8.9) belong to $\mathcal{I}(\mathcal{D})$. We complete $\omega^{k+1}, \ldots, \omega^n$ by adding $\omega^1, \ldots, \omega^k$ in such a way that $\{\omega^j\}_{j=1,\ldots,n}$ form a coframe on U, i.e., $\{\omega^1(p), \ldots, \omega^n(p)\}$ is a basis for T_p^*M, for all $p \in U$. Moreover, we denote by $\{X_1(p), \ldots, X_n(p)\}$ the dual of $\{\omega^1(p), \ldots, \omega^n(p)\}$. It is easy to check that $X_1, \ldots, X_n$ are smooth, linearly independent vector fields on U. Note that $\omega^j(p)\big(X_i(p)\big) = \delta_i^j$, for all $p \in U$. In particular, this implies that

$$\omega^j(p)\big(X_i(p)\big) = 0\,, \tag{8.17}$$

for all $j = k+1, \ldots, n$ and $i = 1, \ldots, k$. From the second part of Lemma 8.2, we then deduce that

$$\operatorname{span}\big\{X_1(p), \ldots, X_k(p)\big\} = \mathcal{D}(p)\,, \tag{8.18}$$

for all $p \in U$.

Next, we observe that because of (8.12) it is sufficient to check the condition $d\mathcal{I}(\mathcal{D}) \subset \mathcal{I}(\mathcal{D})$ on $\omega^{k+1}, \ldots, \omega^n$, since $d\omega^j \in \mathcal{I}(\mathcal{D})$ implies that $\theta^i \wedge d\omega^j \in \mathcal{I}(\mathcal{D})$, for all differential forms θ^i on U.

The exterior derivative of the one-forms $\omega^{k+1}, \dots, \omega^n$ on U can be computed using the formula (6.50). More precisely, for $j = k+1, \dots, n$ and $l, m = 1, \dots, n$ we have

$$d\omega^j(X_l, X_m) = d\big(\omega^j(X_m)\big) \cdot X_l - d\big(\omega^j(X_l)\big) \cdot X_m - \omega^j\big([X_l, X_m]\big) .$$

If $1 \leq l, m \leq k$, it follows from (8.17) that the first two terms on the right-hand side vanish, hence

$$d\omega^j(X_l, X_m) = -\omega^j\big([X_l, X_m]\big) . \tag{8.19}$$

The assertion then follows from the last equation. In fact, assuming that $\mathcal{D}$ is involutive, i.e., $[X_l, X_m] \in \mathcal{D}$, the right-hand side of (8.19) vanishes. This shows that $d\omega^j(X_l, X_m) = 0$, for $l, m = 1, \dots, k$, and therefore $d\omega^j \in \mathcal{I}(\mathcal{D})$.

Conversely, assume that $d\omega^j \in \mathcal{I}(\mathcal{D})$, for $j = k+1, \dots, n$. Then the left-hand side of (8.19) vanishes, showing that $\omega^j\big([X_l, X_m]\big) = 0$, for $1 \leq l, m \leq k$. This implies that (see (8.18))

$$[X_l, X_m](p) \in \operatorname{span}\big\{X_1(p), \dots, X_k(p)\big\}$$

for all $p \in U$, which is equivalent to

$$[X_l, X_m] = \sum_{i=1}^{k} C_{lm}^k \, X_i \, ,$$

for some $C_{lm}^k \in C^\infty(U, \mathbb{R})$, implying that the distribution $\mathcal{D}$ is involutive. □

Next, we establish that the condition $d\mathcal{I}(\mathcal{D}) \subset \mathcal{I}(\mathcal{D})$ can be expressed differently.

Proposition 8.7. *Let $\mathcal{I}(\mathcal{D})$ be the annihilator of a distribution $\mathcal{D}$ on M being locally generated by the smooth, linearly independent one-forms $\omega^{k+1}, \dots, \omega^n$ on U. Moreover, let ω be the smooth differential $(n-k)$-form on U defined by $\omega = \omega^{k+1} \wedge \ldots \wedge \omega^n$. Then the following statements are equivalent:*

(i) the annihilator $\mathcal{I}(\mathcal{D})$ is a differential ideal, i.e., $d\mathcal{I}(\mathcal{D}) \subset \mathcal{I}(\mathcal{D})$;
(ii) the exterior derivative of the generating one-forms $\omega^{k+1}, \dots, \omega^n$ equals

$$d\omega^j = \sum_{i=k+1}^{n} \omega^j_{\ i} \wedge \omega^i \, ,$$

for some differential one-form $\omega^j_{\ i}$ on U, $j = k+1 \dots, n$;
(iii) $\omega \wedge d\omega^j = 0$, for $j = k+1 \dots, n$;
(iv) and $d\omega = \theta \wedge \omega$, for some differential one-form θ on U.

Proof. For (i) implies (ii): By assumption, we have that $d\omega^j \in \mathcal{I}(\mathcal{D})$, for $j = k+1 \ldots, n$. Thus the representation (8.12) holds for $d\omega^j$. Setting $\theta^i = \omega^j_i$ we obtain (ii).

For (i) implies (iii): Obviously, it follows from (i) that $\omega \wedge d\omega^j = 0$.

For (i) implies (iv): This implication can easily be checked.

For (ii) implies (iii): We simply insert (ii) in $\omega \wedge d\omega^i$ obtaining that the last expression vanishes.

For (iv) implies (iii): Note that (iv) means that

$$d\omega = \sum_{i=k+1}^{n} (-1)^i d\omega^i \wedge \omega^{k+1} \wedge \ldots \wedge \widehat{\omega^i} \wedge \ldots \wedge \omega^n = \theta \wedge \omega^{k+1} \wedge \ldots \wedge \omega^n \,.$$

Multiplying with ω^i both sides of the last equation, we obtain (iii).

For (iii) implies (ii): As in the proof of Proposition 8.6, let $\{\omega^1, \ldots, \omega^n\}$ be a coframe on U such that $\omega^{k+1}, \ldots, \omega^n$ locally generate the annihilator $\mathcal{I}(\mathcal{D})$. Thus, for $d\omega^h$, $h = 1, \ldots, n$, we have the following representation in the above coframe:

$$d\omega^h = \sum_{l<m=1}^{n} f^h_{lm}\, \omega^l \wedge \omega^m \,, \tag{8.20}$$

where $f^h_{lm} \in C^\infty(U, \mathbb{R})$. This implies, for $j = k+1, \ldots, n$, using (iii),

$$d\omega^j \wedge \omega = \sum_{l<m=1}^{n} f^j_{lm}\, \omega^l \wedge \omega^m \wedge \omega = \sum_{l<m=1}^{k} f^j_{lm}\, \omega^l \wedge \omega^m \wedge \omega^{k+1} \wedge \ldots \wedge \omega^n = 0 \,.$$

Hence, the functions in (8.20) satisfy $f^i_{lm} = 0$, for $j = k+1, \ldots, n$ and $1 \le l, m \le k$, showing that (8.20) can be written as in (ii).

For (iii) implies (i): This implication is straightforward. □

We collect the previous results concerning the annihilator of a distribution in the following version of Frobenius' theorem.

Theorem 8.8 (Frobenius' Theorem (Differential Form Version)). *Let $\mathcal{D}$ be a k-dimensional distribution on a manifold M with annihilator $\mathcal{I}(\mathcal{D})$. Then $\mathcal{D}$ is integrable if and only if one of the following equivalent statements hold:*

(i) We have $d\mathcal{I}(\mathcal{D}) \subset \mathcal{I}(\mathcal{D})$.

(ii) For every point $p \in M$, there exists a neighborhood U of p and smooth, linearly independent one-forms $\omega^{k+1}, \ldots, \omega^n$ on U which generate locally $\mathcal{I}(\mathcal{D})$ and such that

$$d\omega^j = \sum_{i=k+1}^{n} \omega^j_i \wedge \omega^i \,, \tag{8.21}$$

for some smooth one-forms ω^j_i on U.

(iii) *Let* $\omega = \omega^{k+1} \wedge \ldots \wedge \omega^n$ *be a smooth* $(n-k)$*-form on* U*. Then, we have*

$$d\omega^j \wedge \omega = 0\,, \tag{8.22}$$

for $j = k+1, \ldots, n$*.*

(iv) *Let* $\omega = \omega^{k+1} \wedge \ldots \wedge \omega^n$ *be a smooth* $(n-k)$*-form on* U*. Then we have*

$$d\omega = \theta \wedge \omega\,, \tag{8.23}$$

for some smooth one-form θ *on* U*.*

Example 8.9. We consider the case of a 2-dimensional distribution $\mathcal{D}$ on $\mathbb{R}^3$. – Assume that the annihilator $\mathcal{I}(\mathcal{D})$ is locally generated by the smooth one-form $\omega = x_1\, dx^2 + dx^3$ on $U \subset \mathbb{R}^3$ open. It follows that

$$\mathcal{D}(p) = \ker\,(x_1\, dx^2 + dx^3)_p\,,$$

for all $p \in U$. Since $d\omega = dx^1 \wedge dx^2$, we deduce that

$$d\omega \wedge \omega = (dx^1 \wedge dx^2) \wedge (x_1\, dx^2 + dx^3) = dx^1 \wedge dx^2 \wedge dx^3\,.$$

The last expression being different from zero, condition (iii) of Frobenius' Theorem 8.8 implies that $\mathcal{D}$ is *not* integrable.

Application of Frobenius' Theorem

As an application of Frobenius' theorem, we discuss the so-called *coframe equivalence problem*: Given a Lie group and a manifold of same dimension we want to determine when their coframes can be mapped to each other. For a more general setting of the equivalence problem of coframes we refer to [].

Theorem 8.10. *Let* G *be an* n*-dimensional Lie group whose Lie algebra* $\mathfrak{g}$ *has structure coefficients* $c_{ij}^k \in \mathbb{R}$ *with respect to a basis* $\{E_1, \ldots, E_n\}$ *of* $\mathfrak{g}$*, and let* $\{\theta^1, \ldots, \theta^n\}$ *denote the Maurer-Cartan coframe. On the other hand, let* $\{\bar\theta^1, \ldots, \bar\theta^n\}$ *be a coframe on an* n*-dimensional manifold* M *satisfying, for* $k = 1, \ldots, n$*,*

$$d\bar\theta^k = -\sum_{i<j=1}^{n} c_{ij}^k\, \bar\theta^i \wedge \bar\theta^j\,,$$

where c_{ij}^k *are the above structure coefficients. Then, for every* $p \in M$*, there exists an open neighborhood* $U \subset M$ *of* p *and a local diffeomorphism* $\varphi : U \longrightarrow G$ *with* $\varphi(p) = e$ *such that the coframe on* M *is mapped to the Maurer-Cartan coframe on* G*, i.e.,*

$$\varphi^*\theta^i = \bar\theta^i\,, \tag{8.24}$$

for $i = 1, \ldots, n$*.*

Remark. Suppose that there exist two such diffeomorphisms φ_1 and φ_2 satisfying (8.24). Then, we observe that $(\varphi_2^{-1})^*\bar{\theta}^i = \theta^i$, for $i = 1, \ldots, n$. Thus, inserting (8.24) for φ_1, it follows

$$\theta^i = (\varphi_2^{-1})^*\bar{\theta}^i = (\varphi_2^{-1})^*\big(\varphi_1^*\theta^i\big) = (\varphi_1 \circ \varphi_2^{-1})^*\theta^i \,.$$

Since $\theta^i \in \Omega^1(G)$ are left invariant, we then deduce that $\varphi_1 \circ \varphi_2^{-1} = L_g$, for some $g \in G$, showing that the two local diffeomorphisms φ_1 and φ_2 differ by a left translation on the Lie group on G.

Proof. Let $p \in M$ with open neighborhood $U \subset M$. Consider the canonical projections

$$\pi_1 : U \times G \longrightarrow U \qquad \text{and} \qquad \pi_2 : U \times G \longrightarrow G\,,$$

where $U \times G$ is an $2n$-dimensional manifold and introduce $\bar{\vartheta}^j = \pi_1^*\bar{\theta}^j$, $\vartheta^i = \pi_2^*\theta^i$ being one-forms on $U \times G$, for $i, j = 1, \ldots, n$.

Next, we claim that $\{\bar{\vartheta}^1, \ldots, \bar{\vartheta}^n, \vartheta^1, \ldots, \vartheta^n\}$ is a coframe on $U \times G$. – In order to show the claim, let $(p_0, g_0) \in U \times G$ and consider the maps

$$\begin{aligned} \sigma_1 : U &\longrightarrow U \times G\ , \qquad & \sigma_2 : G &\longrightarrow U \times G\,, \\ q &\longmapsto (q, g_0)\,, & g &\longmapsto (p_0, g)\,. \end{aligned}$$

We then see that

$$\pi_1 \circ \sigma_1 = id_U\,, \quad \pi_1 \circ \sigma_2 = p_0\,, \quad \pi_2 \circ \sigma_2 = id_G\,, \quad \pi_2 \circ \sigma_1 = g_0\,.$$

This implies that

$$\sigma_1^*\bar{\vartheta}^j = \sigma_1^*(\pi_1^*\bar{\theta}^j) = (\pi_1 \circ \sigma_1)^*\bar{\theta}^j = \bar{\theta}^j\,,$$

and

$$\sigma_1^*\vartheta^i = \sigma_1^*(\pi_2^*\theta^i) = (\pi_2 \circ \sigma_1)^*\theta^i = 0\,.$$

Similarly, we compute $\sigma_2^*\bar{\vartheta}^j$ and $\sigma_2^*\vartheta^i$.

Now, assume that there exist $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ and $\mu_1, \ldots, \mu_n \in \mathbb{R}$ such that

$$\sum_{j=1}^n \lambda_j\, \bar{\vartheta}^j(p_0, g_0) + \sum_{i=1}^n \mu_i\, \vartheta^i(p_0, g_0) = 0\,. \tag{8.25}$$

Taking the pull-back by σ_1 of the left-hand side, leads to

$$\sum_{j=1}^n \lambda_j\, \big(\sigma_1^*\bar{\vartheta}^j\big)(p_0) + \sum_{i=1}^n \mu_i\, \big(\sigma_1^*\vartheta^i\big)(p_0) = \sum_{j=1}^n \lambda_j\, \bar{\theta}^j(p_0) = 0\,.$$

Since $\{\bar{\theta}^1(p_0), \ldots, \bar{\theta}^n(p_0)\}$ are linearly independent by assumption, we conclude that $\lambda_j = 0$, for $j = 1, \ldots, n$. On the other-hand, taking the pull-back by σ_2 of the left-hand side of (8.25), we deduce that

$$\sum_{j=1}^{n} \lambda_j \left(\sigma_2^* \bar{\vartheta}^j\right)(g_0) + \sum_{i=1}^{n} \mu_i \left(\sigma_2^* \vartheta^i\right)(g_0) = \sum_{i=1}^{n} \mu_i \, \theta^i(g_0) = 0 \, .$$

Hence, it follows by assumption that $\mu_i = 0$, for $i = 1, \ldots, n$. In summary, we get that

$$\left\{ \bar{\vartheta}^1(p_0, g_0), \ldots, \bar{\vartheta}^n(p_0, g_0), \vartheta^1(p_0, g_0), \ldots, \vartheta^n(p_0, g_0) \right\}$$

is a basis of $T^*_{(p_0,g_0)} U \times G$. Since, this holds for every $(p_0, g_0) \in U \times G$ the claim follows.

In order to apply Frobenius' theorem, we first observe that $\bar{\vartheta}^1 - \vartheta^1, \ldots, \bar{\vartheta}^n - \vartheta^n$ are smooth, linearly independent one-forms on $U \times G$ and consider the n-dimensional distribution $\mathcal{D}$ on $U \times G$ defined by (see Lemma 8.2)

$$\mathcal{D}(q, g) = \bigcap_{i=1}^{n} \ker\left(\bar{\vartheta}^i(q, g) - \vartheta^i(q, g)\right) , \tag{8.26}$$

where $(q, g) \in U \times G$. We want to show that $\mathcal{D}$ is integrable. Applying Frobenius' Theorem 8.8 this can be done by establishing that the annihilator $\mathcal{I}(\mathcal{D})$ of $\mathcal{D}$ generated by $\bar{\vartheta}^1 - \vartheta^1, \ldots, \bar{\vartheta}^n - \vartheta^n$ is a differential ideal, i.e., $d\mathcal{I}(\mathcal{D}) \subset \mathcal{I}(\mathcal{D})$. – For this purpose, note that by assumption we have, for $k = 1, \ldots, n$,

$$\begin{aligned} d\bar{\vartheta}^k = d(\pi_1^* \bar{\theta}^k) = \pi_1^*(d\bar{\theta}^k) &= \pi_1^* \left(\sum_{i<j=1}^{n} c_{ij}^k \, \bar{\theta}^i \wedge \bar{\theta}^j \right) \\ &= \sum_{i<j=1}^{n} c_{ij}^k \, \pi_1^* \bar{\theta}^i \wedge \pi_1^* \bar{\theta}^j = \sum_{i<j=1}^{n} c_{ij}^k \, \bar{\vartheta}^i \wedge \bar{\vartheta}^j \, . \end{aligned} \tag{8.27}$$

Similarly, we have that

$$d\vartheta^k = \sum_{i<j=1}^{n} c_{ij}^k \, \vartheta^i \wedge \vartheta^j \, . \tag{8.28}$$

From (8.27) and (8.28), it then follows that

$$d(\bar{\vartheta}^k - \vartheta^k) = \sum_{i<j=1}^{n} c_{ij}^k (\bar{\vartheta}^i \wedge \bar{\vartheta}^j - \vartheta^i \wedge \vartheta^j) \, .$$

The right-hand side can be written differently and we obtain

$$d(\bar{\vartheta}^k - \vartheta^k) = \sum_{i<j=1}^{n} c_{ij}^k \left((\bar{\vartheta}^i - \vartheta^i) \wedge \bar{\vartheta}^j + \vartheta^i \wedge (\bar{\vartheta}^j - \vartheta^j) \right) .$$

Since $\mathcal{I}(\mathcal{D})$ is generated by $\bar{\vartheta}^1-\vartheta^1,\ldots,\bar{\vartheta}^n-\vartheta^n$, we conclude that $d(\bar{\vartheta}^k-\vartheta^k)\in\mathcal{I}(\mathcal{D})$, for $k=1,\ldots,n$. Hence, the n-dimensional distribution $\mathcal{D}$ on $U\times G$ is integrable. Considering $(p,e)\in U\times G$, this means by definition that there exists a n-dimensional submanifold N^n of $U\times G$ passing through (p,e) such that

$$(d\iota_N)_{(q,g)}(T_{(q,g)}N^n)=\mathcal{D}(q,g)\,, \tag{8.29}$$

for all $(q,g)\in N^n$. Note that in order to simplify the notation we write N^n instead of $N^n(p,e)$.

Next, let $\tilde{\pi}_1$ and $\tilde{\pi}_2$ be the restrictions of π_1 and π_2 to N^n, i.e.,

$$\tilde{\pi}_1=\pi_1\circ\iota_N : N^n\longrightarrow U \qquad \text{and} \qquad \tilde{\pi}_2=\pi_2\circ\iota_N : N^n\longrightarrow G\,.$$

It is left as an exercise to show that

$$(d\tilde{\pi}_1)_{(p,e)} : T_{(p,e)}N^n\longrightarrow T_pU \qquad \text{and} \qquad (d\tilde{\pi}_2)_{(p,e)} : T_{(p,e)}N^n\longrightarrow T_eG$$

are isomorphisms. Note that the statement is a direct consequence of the fact that $\Theta^i:=\iota_N^*\vartheta^i=\tilde{\pi}_2^*\theta_i$ respectively, $\bar{\Theta}^i:=\iota_N^*\bar{\vartheta}^i=\tilde{\pi}_1^*\bar{\theta}_i$ are coframes on the integral manifold N^n, for $i=1,\ldots,n$. In other words, the statement says that at the point $(p,e)\in N^n$ the tangent space $T_{(p,e)}N^n$ contains *no* horizontal or vertical tangent vectors (see Fig. 8.1).

We deduce from the Local Inversion Theorem 1.10 the existence of an open neighborhood U' of $(p,e)\in N^n$ and an open neighborhood $W\subset M$ of $\tilde{\pi}_1(p,e)=p$ such that $\tilde{\pi}_1|_{U'}:U'\longrightarrow W$ is invertible. We denote the inverse by $\tilde{\sigma}_1:W\longrightarrow U'$. Similarly, we deduce by the local inversion theorem for $\tilde{\pi}_2$. Then, we define

$$\varphi=\tilde{\pi}_2\circ\tilde{\sigma}_1 : W\subset M\longrightarrow G\,.$$

It is clear that φ is a diffeomorphism between W and $\varphi(W)$ with $\varphi(p)=e$. It remains to check that (8.24) holds for φ. For this purpose, we compute, for $i=1,\ldots,n$,

$$\begin{aligned}\varphi^*\theta^i&=(\tilde{\pi}_2\circ\tilde{\sigma}_1)^*\theta^i=\tilde{\sigma}_1^*(\tilde{\pi}_2^*\theta^i)\\&=\tilde{\sigma}_1^*\Theta^i=\tilde{\sigma}_1^*\bar{\Theta}^i\\&=\tilde{\sigma}_1^*(\tilde{\pi}_1^*\bar{\theta}^i)=(\tilde{\pi}_1\circ\tilde{\sigma}_1)^*\bar{\theta}^i=\bar{\theta}^i\,.\end{aligned}$$

This completes the proof of the theorem. $\square$

Fig. 8.1. Setting for the proof of the theorem concerning the equivalence problem of coframes.

9 Fiber Bundles

In the previous chapters, we encountered the tangent bundle, cotangent bundle and the tensor bundle. These are examples of manifolds that possess some additional structure. Namely, taking the tangent bundle TM of an n-dimensional manifold M, we know that each point of TM has a neighborhood diffeomorphic to $U \times \mathbb{R}^n$, where U is an open subset of M. Of course, the tangent bundle itself need *not* to be diffeomorphic to $M \times \mathbb{R}^n$. – In the following, we will give a detailed definition for such manifolds, that is, manifolds which look *locally* like a product.

9.1 Basic Definitions and Examples

Definition 9.1. *Let F, M, B denote manifolds and G a Lie group acting effectively on F. A **coordinate bundle** over the **base space** B with **total space** M, **fiber** F and **structure group** G is a submersion*

$$\pi : M \longrightarrow B\,,$$

*called the **bundle projection**, together with a **bundle atlas** $\big\{\big(\pi^{-1}(U_i), (\pi, \varphi_i)\big)\big\}_{i\in I}$ on M. This means the following:*

(i) The family $\{U_i\}_{i\in I}$ is an open covering of B.

(ii) The map

$$(\pi, \varphi_i) : \pi^{-1}(U_i) \longrightarrow U_i \times F$$

is a diffeomorphism. For $b \in U_i$, note that

$$\varphi_i|_{\pi^{-1}(b)} : \pi^{-1}(b) \longrightarrow F$$

is also a diffeomorphism. If b belongs also to U_j, then $\varphi_j|_{\pi^{-1}(b)} : \pi^{-1}(b) \longrightarrow F$ do not necessarily coincide with φ_i. However, they must differ by the action of some element in G. More precisely, we have

(iii) For $i, j \in I$ with $U_i \cap U_j \neq \emptyset$, there is a smooth map

$$f_{i,j} : U_i \cap U_j \longrightarrow G\,,$$

*called the **transition function** from φ_i to φ_j, given by*

$$f_{i,j}(b) = \varphi_j \circ (\varphi_i|_{\pi^{-1}(b)})^{-1} : F \longrightarrow F\,. \tag{9.1}$$

Remark. a) Roughly speaking, the total space M consists of an union of products $U_i \times F$, where the open sets U_i, $i \in I$, cover the base space B and the copies of F belonging to intersecting U_i and U_j are identified by means of elements of G.

b) The set $\pi^{-1}(b)$ is called the fiber over b. Note also that the bundle atlas will be often called a *trivialization*, in order to emphasize the local triviality property of the coordinate bundle.

c) For the transition functions the following holds for $b \in U_i \cap U_j \neq \emptyset$ and $m \in F$:

$$\big(\pi, \varphi_j\big) \circ \big(\pi, \varphi_i\big)^{-1}(b, m) = \big(b, f_{i,j}(b)m\big)\,. \tag{9.2}$$

Note also that (iii) implies that $f_{i,i}(b) = e$, $\big(f_{i,j}(b)\big)^{-1} = f_{j,i}(b)$ and the transition functions satisfy the *cocycle condition*

$$f_{i,k}(b) = f_{j,k}(b) \circ f_{i,j}(b)\,, \qquad b \in U_i \cap U_j \cap U_k \neq \emptyset\,. \tag{9.3}$$

Fig. 9.1. Fiber bundle.

There is an important particular class of coordinate bundles.

Definition 9.2. *A (real)* ***coordinate vector bundle*** *of rank n is a coordinate bundle $\pi : M \longrightarrow B$ with fiber $\mathbb{R}^n$ and structure group $GL_n(\mathbb{R})$ (or a subgroup of $GL_n(\mathbb{R})$). A (complex)* ***coordinate vector bundle*** *of rank n is defined analogously. – In the following, we will always denote the total space of a vector bundle by E.*

Now, we define in a detailed manner the notion of differentiable maps between coordinate fiber bundles which will allow us to introduce in Definition 9.4 below equivalence classes of coordinate fiber bundles, simply called *fiber bundles*.

Definition 9.3. *Let $\pi_i : M_i \longrightarrow B_i$, for $i = 1, 2$, be two coordinate bundles with fiber F and structure group G. A differentiable map*

$$h : M_1 \longrightarrow M_2$$

is said to be a ***bundle map*** *if it maps each fiber $\pi_1^{-1}(b_1)$ diffeomorphically onto a fiber $\pi_2^{-1}(b_2)$, thereby inducing a differentiable map $\bar{h} : B_1 \longrightarrow B_2$ such that $\pi_2 \circ h = \bar{h} \circ \pi_1$. Moreover, for every bundle charts $\big(\pi_1^{-1}(U_i), (\pi_1, \varphi_i)\big)$ and $\big(\pi_2^{-1}(V_j), (\pi_2, \psi_j)\big)$ of π_1, respectively π_2, the* ***local representation*** *of h at $b \in U_i \cap \bar{h}^{-1}(V_j)$ given by*

$$\psi_j \circ h \circ (\varphi_i|_{\pi_1^{-1}(b)})^{-1} : F \longrightarrow F\,,$$

coincides with the action of an element of G, and the resulting map

$$\begin{aligned} h_{i,j} : U_i \cap \bar{h}^{-1}(V_j) &\longrightarrow G\,, \\ b &\longmapsto \psi_j \circ h \circ \left(\varphi_i|_{\pi_1^{-1}(b)}\right)^{-1} \end{aligned} \tag{9.4}$$

is smooth.

Fig. 9.2. Bundle map.

Definition 9.4. *Let $h : M_1 \longrightarrow M_2$ be a bundle map between two coordinate bundles $\pi_i : M_i \longrightarrow B$ with the same structure group, fiber and base space B. Then, if the induced map $\bar{h} : B \longrightarrow B$ is the identity map on B, the coordinate bundles π_1 and π_2 are called **equivalent** and h a **bundle isomorphism**. A **fiber bundle** is then defined to be an equivalence class of coordinate bundles. Moreover, a fiber bundle is called **trivial** if the product bundle is an element of its equivalence class.*

Remark. Note that, alternatively, a fiber bundle can be defined in terms of *equivalent* system of bundle charts as in the case of manifolds (see Definition 2.17).

Next, we define a special type of fiber bundles whose fiber and structure group coincide and whose total space will be denoted by P.

Definition 9.5. *A fiber bundle $\pi : P \longrightarrow B$ with fiber and structure group G is a called a **principal G-bundle** if there exists a free right action on P and a bundle atlas such that for each bundle chart $\big(\pi^{-1}(U), (\pi, \varphi)\big)$, the map $\varphi : \pi^{-1}(U) \longrightarrow G$ is **G-equivariant**, i.e.,*

$$\big(\pi, \varphi\big)(pg) = \big(\pi(p), \varphi(p)g\big)\,, \tag{9.5}$$

for $p \in \pi^{-1}(U)$ and all $g \in G$.

We claim that for a principal G-bundle the base space B is the quotient space P/G. – Since from (9.5) we have that $\pi(pg) = \pi(p)$, the orbit $G_p = \{pg \ : \ g \in G\} \subset P$ of $p \in P$ is contained in $\pi^{-1}\big(\pi(p)\big)$. Conversely, if $\big(\pi^{-1}(U), (\pi, \varphi)\big)$ is a bundle chart about p, then it follows, for $q \in \pi^{-1}\big(\pi(p)\big)$,

$$\begin{aligned} q &= (\pi, \varphi)^{-1}\big(\pi(q), \varphi(q)\big) = (\pi, \varphi)^{-1}\big(\pi(q), \varphi(p)\varphi(p)^{-1}\varphi(q)\big) \\ &= (\pi, \varphi)^{-1}\big(\pi(q), \varphi(p)g\big) \overset{(9.5)}{=} pg\,, \end{aligned}$$

where $g = \varphi(p)^{-1}\varphi(q) \in G$. This shows the claim.

Moreover, note that for a principal G-bundle the Lie group G acts on itself by *left translations*. Indeed, the transition functions for $p \in P$ read as[1]

$$f_{i,j}\big(\pi(p)\big) = \varphi_j(p)\varphi_i(p)^{-1} : G \longrightarrow G\,, \tag{9.6}$$

being independent of the choice of $q \in \pi^{-1}\big(\pi(p)\big)$, since for $q = pg$ we have

$$\varphi_j(pg)\varphi_i(pg)^{-1} = \varphi_j(p)g\big(\varphi_i(p)g\big)^{-1} = \varphi_j(p)gg^{-1}\varphi_i(p)^{-1} = \varphi_j(p)\varphi_i(p)^{-1}\,.$$

In the following theorem, we characterize bundle isomorphisms between principal bundles.

Theorem 9.6. *Let $\pi_i : P_i \longrightarrow B$ be two principal G-bundles over B and assume that $h : P_1 \longrightarrow P_2$ is a G-equivariant smooth map inducing the identity on B. Then, we have that h is a bundle isomorphism implying that the two bundles are equivalent.*

Proof. Let $\big(\pi_1^{-1}(U_i), (\pi_1, \varphi_i)\big)$ and $\big(\pi_2^{-1}(V_j), (\pi_1, \psi_j)\big)$ be bundle charts of π_1 and π_2, respectively. Then the maps

$$\begin{aligned} \pi_1^{-1}(U_i) &\longrightarrow G\,, & \pi_1^{-1}(V_j) &\longrightarrow G\,,\\ p &\longmapsto \varphi_i(p)\,, & p &\longmapsto \big(\psi_j \circ h\big)(p) \end{aligned}$$

are smooth and well-defined since h induces the identity map on B by assumption. Thus the assignment

$$\begin{aligned} f_{i,j} : U_i \cap V_j &\longrightarrow G\,,\\ b &\longmapsto \big(\psi_j \circ h\big)(p)\varphi_i(p)^{-1}\,, \end{aligned}$$

where $p \in \pi_1^{-1}(b)$, is also smooth.

Now, we show that $f_{i,j}(b)$ equals the local representation $h_{i,j}(b)$ of h at b given by $\psi_j \circ h \circ (\varphi_i|_{\pi_1^{-1}(b)})^{-1}$. – Let $g \in G$ and let $a := \varphi_i(p) \in G$. Then, we have

$$g = aa^{-1}g = \varphi_i(p)a^{-1}g \overset{(9.5)}{=} \varphi_i(pa^{-1}g)\,.$$

For $p \in \pi_1^{-1}(b)$, this implies that

$$\begin{aligned} \psi_j \circ h \circ (\varphi_i|_{\pi_1^{-1}(b)})^{-1}(g) &= \psi_j \circ h(pa^{-1}g) = \psi_j\big(h(p)\big)a^{-1}g\\ &= \big(\psi_j \circ h\big)(p)\varphi_i(p)^{-1}g\,, \end{aligned} \tag{9.7}$$

where we used that the map h is G-equivariant by assumption. Thus the theorem follows using Definition 9.4. $\square$

[1] Let $g \in G$ and let $a := \varphi_i(p) \in G$. Then, we can write $g = aa^{-1}g = \varphi_i(p)a^{-1}g = \varphi_i(pa^{-1}g)$, implying for $p \in \pi^{-1}(b)$ that (see (9.1))

$$\varphi_j \circ (\varphi_i|_{\pi^{-1}(b)})^{-1}(g) = \varphi_j(pa^{-1}g) = \varphi_j(p)\varphi_i(p)^{-1}g\,.$$

Fig. 9.3. Equivalent G-principal bundles.

Examples of Fiber Bundles

Example 9.7. The *trivial* fiber bundle with base space B and fiber F is the projection $\pi : B \times F \longrightarrow B$ onto the first factor. The structure group of the trivial bundle is simply the identity transformation on F. – Note that in general the size of the structure group measures how twisted the bundle is.

Example 9.8. The *trivial* principal G-bundle over B is the projection π : $B \times G \longrightarrow B$ onto the first factor. The free action of G on $P = B \times G$ is defined by right multiplication on G, i.e., $(b, g_0)g = (b, g_0 g) \in P$.

Example 9.9. The tangent bundle TM of an n-dimensional manifold M is the total space of a rank n vector bundle over M with bundle projection $\pi : TM \longrightarrow M$ (see Section 2.5.3). Recall that if $\{(U_i, \varphi_i)\}_{i\in I}$ is an atlas on M, then $\big\{\big(\pi^{-1}(U_i), \Phi_i\big)\big\}_{i\in I}$, where $\Phi_i = (\pi, d\varphi_i)$, is an bundle atlas on TM. Moreover, the transition functions are given by (see (2.21))

$$f_{i,j} = d(\varphi_j \circ \varphi_i^{-1}) \circ \varphi_i : U_i \cap U_j \longrightarrow GL_n(\mathbb{R}) .$$

Note that similar arguments hold for the exterior bundle $\bigwedge^p T^*M$ and the tensor bundle $T^r_s(M)$.

Example 9.10 (Canonical Line Bundle over $\mathbb{C}P^n$). There is a canonical or "tautological" (complex) vector bundle of rank 1 over $\mathbb{C}P^n$, where the *complex projective space* $\mathbb{C}P^n$ is defined similarly to $\mathbb{R}P^n$ (see Example 2.21). – The total space E_L of this bundle is given by

$$E_L = \big\{\big([z], \xi\big) \in \mathbb{C}P^n \times \mathbb{C}^{n+1} \ : \ \xi \in [z]\big\} ,$$

with bundle projection

$$\begin{aligned} \pi_L : E_L &\longrightarrow \mathbb{C}P^n , \\ \big([z], \xi\big) &\longmapsto [z] . \end{aligned}$$

In other words, the fiber $\pi_L^{-1}([z])$ over $[z] \in \mathbb{C}P^n$ is simply $[z]$ itself. For the open covering

$$U_i = \{[z] \in \mathbb{C}P^n \ : \ z_i \neq 0\} ,$$

of the complex projective space, with $i = 1, \dots, n+1$, we take the bundle charts

$$\begin{aligned} (\pi_L, \varphi_i) : \pi_L^{-1}(U_i) &\longrightarrow U_i \times \mathbb{C} , \\ \big([z], \xi\big) &\longmapsto \big([z], \lambda_i\big) , \end{aligned} \tag{9.8}$$

where $\lambda_i \in \mathbb{C}$ is such that $(\xi_1, \ldots, \xi_{n+1}) = \lambda_i \left(\frac{z_1}{z_i}, \ldots, 1, \ldots, \frac{z_{n+1}}{z_i}\right) \in \mathbb{C}^{n+1}$. Moreover, the transition functions on $U_i \cap U_j \neq \emptyset$ are given by

$$f_{i,j}\big([z]\big) = \frac{z_j}{z_i} \,. \tag{9.9}$$

Since $z_i \neq 0$ and $z_j \neq 0$ on $U_i \cap U_j$, we deduce that the transition functions are elements of $GL_1(\mathbb{C})$ as required by Definition 9.2.

Sections of Bundles

It is clear that the trivial bundle $B \times F \longrightarrow B$ has the property that through any point $(b, m) \in B \times F$ there is a copy $B \times \{m\}$ of B. This fact can be reformulated in the following way: For a trivial bundle, there exists a map $s : B \longrightarrow B \times F$ defined by $s(b) = (b, m)$ such that it is a lift of id_B, i.e., $\pi \circ s = id_B$, through (b, m). However, such maps do *not* exist for all fiber bundles.

Definition 9.11. *Let $\pi : M \longrightarrow B$ be a fiber bundle. A differentiable map $s : B \longrightarrow M$ is called a **section** if $\pi \circ s = id_B$.*

Note that we already encountered the concept of sections in the case of vector bundles. For example, we know that vector fields are sections of the tangent bundle and that differential one-forms are sections of the cotangent bundle. More generally, every vector bundle admits a section, namely the *zero section* given by $s(b) = 0 \in E_b$, where E_b denotes the fiber $\pi^{-1}(b)$ over $b \in B$ which is a vector space. Note that from Exercise 9.13 we know that $0 \in \mathrm{e}_b$ is independent of the trivialization. – For principal bundles, however, we have

Theorem 9.12. *Let $\pi : P \longrightarrow M$ be a principal G-bundle. Then it admits a section if and only if it is the trivial principal G-bundle.*

Remark. It is important to note that we mean the existence of a section defined on B and not a (local) section $s : U \longrightarrow P$ defined only on $U \subset B$ open, since – by local triviality – the map

$$s(b) = \big(\pi, \varphi\big)^{-1}(b, g) \,,$$

where $g \in G$ is fixed and $b \in U$, defines automatically a (local) section for every principal bundle. If a (local) section $s : U \longrightarrow P$ satisfies

$$s(b) = \big(\pi, \varphi\big)^{-1}(b, e) \,, \tag{9.10}$$

then it will be called the (local) section *associated* to the bundle chart (π, φ).

Proof. Assume that $h : P \longrightarrow B \times G$ is the bundle map making π to a trivial bundle. Then, for any $g \in G$, the map $s(b) := h^{-1}(b, g)$ defines a section of the trivial principal G-bundle.

Conversely, assume that there exists a section $s : B \longrightarrow P$. Then consider the map $\varphi : P \longrightarrow G$ such that

$$p = s\big(\pi(p)\big)\varphi(p) \,. \tag{9.11}$$

Since $p \in P$ and $s\big(\pi(p)\big)$ belong by definition to the same fiber, this map is well-defined. Moreover, it is equivariant. Indeed, observe that (9.11) can be written as

$$\varphi_i\big(s(\pi(p))\big)^{-1}\varphi_i(p) = \varphi(p) \,,$$

implying that

$$\begin{aligned}\varphi(pg) &= \varphi_i\big(s(\pi(pg))\big)^{-1}\varphi_i(pg) \\ &= \varphi_i\big(s(\pi(p))\big)^{-1}\varphi_i(p)g = \varphi(p)g \,.\end{aligned}$$

Note that the smoothness of $\varphi : P \longrightarrow G$ is clear.

Next, we define the bundle map

$$h := (\pi, \varphi) : P \longrightarrow B \times G \,,$$

which obviously induces the identity map on B. Then we conclude using Theorem 9.6. □

* * *

Exercises.

Exercise 9.13. Show that for a coordinate vector bundle $\pi_E : E \longrightarrow B$ the fiber $\pi^{-1}(b)$ over each point b of the base space B is a vector space with vector space structure independent of the chosen bundle atlas.

9.2 The Tangent Bundle of a Sphere as Principal Bundle

In this section, as detailed illustration of some concepts in the previous section, we consider the tangent bundle TS^2 of the two-sphere S^2 which can be made to a principal $U(1)$-bundle. – It is left as an exercise to translate the results of this section to the case of the n-dimensional sphere.

From Example 2.34, we know that the tangent bundle TS^2 of the two-dimensional unit sphere can be written as

$$TS^2 = \bigcup_{p \in S^2} T_pS^2 = \big\{(p, v) \in S^2 \times \mathbb{R}^3 \;:\; \langle p, v\rangle = 0\big\} \,,$$

where $\langle \cdot, \cdot \rangle$ means the usual scalar product in $\mathbb{R}^3$. Considering only tangent vectors with unit length we define the set

$$P = \{(p, v) \in S^2 \times S^2 \ : \ \langle p, v \rangle = 0\} . \tag{9.12}$$

a) In a first step, we show that P is an one-dimensional submanifold of $S^2 \times S^2$. For this purpose, let

$$\begin{aligned} f : \mathbb{R}^3 \times \mathbb{R}^3 &\longrightarrow \mathbb{R} , \\ (p, v) &\longmapsto \langle p, v \rangle \end{aligned}$$

be a smooth map, whose restriction to $S^2 \times S^2$ – also denoted by f – remains smooth. Obviously, we have that $P = f^{-1}(0)$. Next, we claim that $f : S^2 \times S^2 \longrightarrow \mathbb{R}$ is a submersion. – Let $(p, v) \in P$ be given by $p = (0, 0, 1) \in \mathbb{R}^3$ and $v = (v_1, v_2, 0) \in \mathbb{R}^3$ with $v_1^2 + v_2^2 = 1$. Note that the tangent space $T_{(p,v)} S^2 \times S^2 \subset \mathbb{R}^3 \times \mathbb{R}^3$ is then generated by

$$\begin{aligned} \tilde{e}_1 &= (1, 0, 0, 0, 0, 0) , & \tilde{e}_2 &= (0, 1, 0, 0, 0, 0) , \\ \tilde{e}_3 &= (0, 0, 0, -v_2, v_1, 0) , & \tilde{e}_4 &= (0, 0, 0, 0, 0, 1) . \end{aligned}$$

From this we easily deduce that

$$\begin{aligned} df_{(p,v)} \cdot \tilde{e}_1 &= \frac{\partial f}{\partial x_1}(p, v) = v_1 , & df_{(p,v)} \cdot \tilde{e}_2 &= \frac{\partial f}{\partial x_2}(p, v) = v_2 , \\ df_{(p,v)} \cdot \tilde{e}_3 &= 0 , & df_{(p,v)} \cdot \tilde{e}_4 &= \frac{\partial f}{\partial x_6}(p, v) = 1 . \end{aligned}$$

Thus it follows that the differential $df_{(p,v)}$ is of maximal rank 1. Using Proposition 2.2, we conclude that P is a submanifold of $S^2 \times S^2$.

b) In a second step, we note that the map

$$\begin{aligned} \pi : P &\longrightarrow S^2 , \\ (p, v) &\longmapsto p \end{aligned} \tag{9.13}$$

defines a submersion from P into S^2.

c) Next, we want to introduce a free right action of $U(1)$ on P. Recall that the abelian Lie group $U(1) = \{e^{i\theta} \in \mathbb{C} \ : \ \theta \in \mathbb{R}\}$ is the unit circle in the complex plane. – We define the map

$$\begin{aligned} U(1) \times P &\longrightarrow P , \\ \big(e^{i\theta}, (p, v)\big) &\longmapsto (p, v)e^{i\theta} = \big(p, \cos\theta\, v + \sin\theta\, p \times v\big) , \end{aligned} \tag{9.14}$$

which describes in fact a rotation of v by the angle θ. It is left as an exercise to check that this map defines a right action of $U(1)$ on P in the sense of Definition 7.27. Moreover, this action is free: Let $(p, v) \in P$ such that $(p, v)e^{i\theta} = (p, v)$. This implies that $v = \cos\theta\, v + \sin\theta\, p \times v$. Since $\{v, p \times v\}$ is a basis of $T_p S^2$, we deduce that $\cos\theta = 1$ and $\sin\theta = 0$, i.e., $e^{i\theta} = 1$.

d) In a next step, we construct a bundle atlas for $\pi : P \longrightarrow S^2$. – Let

$$U_i^+ = \{p \in S^2 \ : \ \langle p, e_i \rangle > 0\} \qquad U_i^- = \{p \in S^2 \ : \ \langle p, e_i \rangle < 0\} ,$$

with $\{e_i\}_{1\leq i\leq 3}$ the canonical basis of $\mathbb{R}^3$, denote an open covering of S^2. For U_3^+, for example, we then consider the smooth map

$$\begin{aligned} v_3 : U_3^+ &\longrightarrow \mathbb{R}^3 , \\ p &\longmapsto \frac{e_1 - \langle e_1, p\rangle p}{\|e_1 - \langle e_1, p\rangle p\|} . \end{aligned} \tag{9.15}$$

Obviously, we have $\|v_3(p)\| = 1$ and $\langle p, v_3(p)\rangle = 0$, i.e., $(p, v_3(p)) \in P$. Since $p \neq e_1$, for $p \in U_3^+$, the map v_3 is also well-defined. Now, we are ready to define the following smooth map, which will turn out to be a bundle chart:

$$\begin{aligned} (\pi, \varphi_3^+) : \pi^{-1}(U_3^+) &\longrightarrow U_3^+ \times U(1) , \\ (p, v) &\longmapsto \big(p, \langle v, v_3(p)\rangle + i \,\langle v, p \times v_3(p)\rangle\big) . \end{aligned} \tag{9.16}$$

Since $\{v_3(p), p\times v_3(p)\}$ is a basis of T_pS^2, we observe that indeed $\langle v, v_3(p)\rangle + i\,\langle v, p \times v_3(p)\rangle \in U(1)$ and (π, φ_3^+) is a smooth diffeomorphism. Moreover, we have that

$$\begin{aligned} (\pi, \varphi_3^+)\big((p, v)e^{i\theta}\big) &= (\pi, \varphi_3^+)\big(p, \cos\theta\, v + \sin\theta\, p \times v\big) \\ &= \Big(p, \big\langle(\cos\theta\, v + \sin\theta\, p \times v), v_3(p)\big\rangle \\ &\qquad + i\,\big\langle(\cos\theta\, v + \sin\theta\, p \times v), p \times v_3(p)\big\rangle\Big) . \end{aligned}$$

A straightforward calculation then gives that $\varphi_3^+ : \pi^{-1}(U_3^+) \longrightarrow U(1)$ is $U(1)$-equivariant, i.e.,

$$(\pi, \varphi_3^+)\big((p, v)e^{i\theta}\big) = \Big(p, \big(\langle v, v_3(p)\rangle + i\,\langle v, p \times v_3(p)\rangle\big)e^{i\theta}\Big) . \tag{9.17}$$

e) We now determine the transition functions. – First, we observe that the map

$$\begin{aligned} s_3^+ : U_3^+ &\longrightarrow \pi^{-1}(U_3^+) \subset P , \\ p &\longmapsto \big(p, v_3(p)\big) , \end{aligned} \tag{9.18}$$

gives the (local) section associated to the bundle chart (π, φ_3^+), i.e., it satisfies

$$s_3^+(p) = \big(\pi, \varphi_3^+\big)^{-1}(p, 1) . \tag{9.19}$$

Using formula (9.6) for the transition functions of a principal bundle and (9.16), we obtain for $p \in U_3^+ \cap U_2^+ \neq \emptyset$ that

$$f_{3,2}(p) = \varphi_2^+\big(p, v_3(p)\big)\varphi_3^+\big(p, v_3(p)\big)^{-1} = \varphi_2^+\big(p, v_3(p)\big) .$$

Remark. Summarizing the previous results, we have shown that $\pi : P \longrightarrow S^2$ is a principal $U(1)$-bundle in the sense of Definition 9.5.

Using Theorem 9.12, we now want to decide if this principal bundle is trivial or not. – Assume that there exists a smooth section $s : S^2 \longrightarrow P$ with $\pi \circ s = id_{S^2}$. One can check that this is equivalent to the existence of a smooth map $\tilde{s} : S^2 \longrightarrow S^2$ such that $\big\langle \tilde{s}(p), p \big\rangle = 0$, for all $p \in S^2$. Moreover, we consider the flow of a vector field $X \in \mathcal{X}(S^2)$ given by

$$\begin{aligned} \varGamma_t^X : S^2 &\longrightarrow S^2 \,, \\ p &\longmapsto \cos(\pi t)p + \sin(\pi t)\tilde{s}(p) \,. \end{aligned} \tag{9.20}$$

Note that $\big\|\cos(\pi t)p + \sin(\pi t)\tilde{s}(p)\big\| = 1$, since $\big\langle \tilde{s}(p), p \big\rangle = 0$. We also observe that $\varGamma_0^X = id_{S^2}$ and $\varGamma_1^X = -id_{S^2}$. For the volume form $\omega_{S^2} = \iota_{S^2}^* \varOmega$ of S^2 (see Example 3.54), we claim that

$$\frac{d}{dt}\bigg|_{t=s} \int_{S^2} (\varGamma_t^X)^* \omega_{S^2} = 0 \,. \tag{9.21}$$

For a fixed $s \in \mathbb{R}$, we compute, using Definition 6.35 for the Lie derivative,

$$L_X\big((\varGamma_s^X)^*\omega_{S^2}\big) = \lim_{t \longrightarrow 0} \frac{1}{t}\Big((\varGamma_s^X)^*\omega_{S^2} - (\varGamma_{-t}^X)^*\big((\varGamma_s^X)^*\omega_{S^2}\big)\Big) \,.$$

Since $(\varGamma_{-t}^X)^*(\varGamma_s^X)^*\omega_{S^2} = (\varGamma_s^X \circ \varGamma_{-t}^X)^*\omega_{S^2} = (\varGamma_{s-t}^X)^*\omega_{S^2}$, we obtain that

$$\begin{aligned} L_X\big((\varGamma_s^X)^*\omega_{S^2}\big) &= \lim_{t \longrightarrow 0} \frac{1}{t}\Big((\varGamma_s^X)^*\omega_{S^2} - (\varGamma_{s-t}^X)^*\omega_{S^2}\Big) \\ &= -\frac{d}{dt}\bigg|_{t=s} (\varGamma_t^X)^*\omega_{S^2} \,. \end{aligned}$$

On the other hand, using Cartan's formula (6.42), the left-hand side of the last equation can be written as

$$L_X\big((\varGamma_s^X)^*\omega_{S^2}\big) = d\big(\mathrm{int}_X(\varGamma_s^X)^*\omega_{S^2}\big) + \mathrm{int}_X\big(d(\varGamma_s^X)^*\omega_{S^2}\big) \,.$$

The second term on the right-hand side vanishes, since the pull-back and the exterior differential commute. Thus, we deduce that

$$-\frac{d}{dt}\bigg|_{t=s} (\varGamma_t^X)^*\omega_{S^2} = d\big(\mathrm{int}_X(\varGamma_s^X)^*\omega_{S^2}\big) \,.$$

Integrating the last equation over S^2, it follows

$$-\frac{d}{dt}\bigg|_{t=s} \int_{S^2} (\varGamma_t^X)^*\omega_{S^2} = \int_{S^2} d\big(\mathrm{int}_X(\varGamma_s^X)^*\omega_{S^2}\big) \,.$$

Stokes' theorem implies that the right-hand side vanishes and we arrive at the claim (9.21).

A straightforward calculation shows that

$$\int_{S^2} (\Gamma_0^X)^* \omega_{S^2} = \int_{S^2} \omega_{S^2} = \int_{S^2} \iota^*_{\partial B^3} \Omega = \int_{B^3} d\Omega$$
$$= 3 \int_{B^3} dx^1 \wedge dx^2 \wedge dx^3 = 3\,\mathrm{Vol}(B^3)\,. \qquad (9.22)$$

On the hand hand, a similar calculation gives

$$\int_{S^2} (\Gamma_1^X)^* \omega_{S^2} = -\int_{S^2} \omega_{S^2} = -3\,\mathrm{Vol}(B^3)\,.$$

However, the last result must be equal to (9.22), as a direct consequence of (9.21). Thus, we end up with a contradiction to the assumption of the existence of a section $s : S^2 \longrightarrow P$. Finally, we conclude with Theorem 9.12 that the principal $U(1)$-bundle $\pi : P \longrightarrow S^2$ is *not* trivial.

9.3 Associated and Reducible Principal Bundles

In this section, we analyze the structure of fiber bundles in more detail.

Associated Principal Bundles

We have seen that a bundle atlas $\{(\pi^{-1}(U_i), (\pi, \varphi_i))\}_{i \in I}$ of a given fiber bundle $\pi : M \longrightarrow B$ with structure group G induces transition functions $f_{i,j} : U_i \cap U_j \longrightarrow G$ which satisfy the cocycle condition (9.3). We will see that these transition functions are the main ingredients for the reconstruction of the given fiber bundle. More generally, we have

Proposition 9.14. *Let $\{U_i\}_{i \in I}$ be an open covering of a manifold B and G a Lie group acting effectively on a manifold F. Moreover, assume that there exists for all $i, j \in I$ maps $f_{i,j} : U_i \cap U_j \longrightarrow G$ such that*

$$f_{i,k}(p) = f_{j,k}(p) \circ f_{i,j}(p)\,, \qquad p \in U_i \cap U_j \cap U_k \neq \emptyset\,. \qquad (9.23)$$

Then there exists a fiber bundle $\pi : M \longrightarrow B$ with fiber F, structure group G and a bundle atlas whose transition functions are given by $f_{i,j}$. In the case of $F = G$ and if G acts on itself by left translations, then the atlas determines a principal G-bundle.

Proof. Consider the disjoint union

$$X = \bigcup_{i \in I} (U_i \times F)\,,$$

and define on this union the following equivalence relation: For $(p_i, q_1) \in U_i \times F$, $(p_j, q_2) \in U_j \times F$, we set $(p_i, q_1) \sim (p_j, q_2)$ if and only if $p_i = p_j \in U_i \cap U_j$ and $q_2 = f_{i,j}(p_i) q_1$, for some $i, j \in I$.

We check that $\sim$ is an equivalence relation on X. – Since $q = eq = f_{i,i}(p_i)q$, we have that $(p_i, q) \sim (p_i, q)$. Now, assume that $(p_i, q_1) \sim (p_j, q_2)$, i.e., that $p_i = p_j$ and $q_2 = f_{i,j}(p_i)q_1$. Hence, we get $q_1 = f_{i,j}(p_i)^{-1}q_2 = f_{j,i}(p_j)q_2$ showing that $(p_j, q_2) \sim (p_i, q_1)$. The transitivity is also a direct consequence of the cocycle condition (9.23). This shows that $\sim$ defines indeed an equivalence relation.

From a set point of view, we now define

$$M = X/\sim = \bigcup_{i\in I}(U_i \times F)/\sim\,,$$

and consider the projection

$$\begin{aligned} \pi : M &\longrightarrow B\,, \\ [(p,q)] &\longmapsto p\,, \end{aligned} \tag{9.24}$$

where $[(p, q)]$ denotes the equivalence class of $(p, q) \in U_i \times F$, for some $i \in I$. Note that π is clearly well-defined meaning that it is independent of the choice of representative. If $\rho : X \longrightarrow M$ denotes the projection on the quotient space, we claim that each restriction

$$\begin{aligned} \rho_i : U_i \times F &\longrightarrow \rho_i(U_i \times F) \subset M\,, \\ (p_i, q) &\longmapsto [(p_i, q)]\,, \end{aligned} \tag{9.25}$$

defines a bijection. Each equivalence class in $\rho_i(U_i \times F)$ possesses a representant in $U_i \times F$. But by definition of the equivalence relation on X this representant is unique and the claim follows. – We will take the inverse of (9.25) as a bundle chart (π, φ_i), i.e.,

$$\begin{aligned} (\pi, \varphi_i) : \pi^{-1}(U_i) &\longrightarrow U_i \times F\,, \\ [(p,q)] &\longmapsto (p,q)\,. \end{aligned} \tag{9.26}$$

By construction the transition functions of the corresponding atlas are just the given $f_{i,j}$.

In order to make M to a differentiable manifold, we first introduce the following topology on M: We say that $\Omega \subset M$ is *open* in M if and only if the set $(\pi, \varphi_i)\big(\Omega \cap \pi^{-1}(U_i)\big)$ is open in $U_i \times F$, for all $i \in I$. Verify that this indeed defines a topology on M. Next, let $\{(U_i, \psi_i)\}_{i\in I}$ and $\{(V_\alpha, \xi_\alpha)\}_{\alpha\in A}$ be an atlas for B, respectively F. On the open sets $W_{i,\alpha} = (\pi, \varphi_i)^{-1}(U_i \times V_\alpha) \subset M$ we then consider the map $\Phi_{i,\alpha}\big([(p, q)]\big) = \big(\psi_i(p), \xi_\alpha(q)\big)$. It is left as an exercise to show that $\{(W_{i,\alpha}, \Phi_{i,\alpha})\}_{i\in I,\alpha\in A}$ defines a differentiable structure on M. Moreover, one can show that for this differentiable structure on M the bundle projection $\pi : M \longrightarrow B$ and the bundle charts (π, φ_i) are smooth maps. – In summary, $\pi : M \longrightarrow B$ is the desired fiber bundle.

In the case of $F = G$, we define a right action of G on M by $[(p, g)]g_0 = [(p, gg_0)]$, for all $g_0 \in G$. It is easy to see that this definition is independent of the choice of representative in the class $[(p,g)]$ and that the action is free. Moreover, the G-equivariance of φ_i follows from

$$\begin{aligned}(\pi,\varphi_i)\big([(p,g)]g_0\big) &= (\pi,\varphi_i)\big([(p,gg_0)]\big)\\ &\overset{(9.26)}{=} (p,gg_0) = \big(p,\varphi_i\big([(p,g)]\big)g_0\big)\,.\end{aligned}$$

We then conclude the existence of a principal G-bundle $\pi : P \longrightarrow B$ with the desired properties. □

This result leads to the following definition:

Definition 9.15. *Let $\pi : M \longrightarrow B$ be a fiber bundle with structure group G. The principal G-bundle constructed from the transition functions of π as in the previous proposition is called the principal G-bundle* ***associated*** *to the given fiber bundle $\pi : M \longrightarrow B$.*

There is also a canonical way of associating a fiber bundle to a given principal bundle. – Let $\pi_P : P \longrightarrow B$ be a principal G-bundle and let F be a manifold on which the Lie group G acts effectively on the left. On $P \times F$, we then define the equivalence relation $(p,m) \sim (p',m')$ if and only if there exists $g \in G$ such that $(p',m') = (pg, g^{-1}m)$. The quotient space $(P \times F)/\sim$ is denoted by $P \times_G F$ and there is a well-defined map

$$\begin{aligned}\tilde{\pi} : P \times_G F &\longrightarrow B\,,\\ [p,m] &\longmapsto \pi_P(p)\,.\end{aligned}\tag{9.27}$$

Theorem 9.16. *Let $\pi_P : P \longrightarrow B$ be a principal G-bundle and let F be a manifold on which the Lie group G acts effectively on the left. Then $\tilde{\pi} : P \times_G F \longrightarrow B$ defined in (9.27) is a fiber bundle over B with fiber F and structure group G. Moreover, the principal G-bundle π_P is associated to $\tilde{\pi}$ in the sense of Definition 9.15.*

Proof. Let $\big(\pi_P^{-1}(U), (\pi_P,\varphi)\big)$ be a chart for the given principal bundle. We then define

$$\begin{aligned}(\tilde{\pi},\tilde{\varphi}) : \tilde{\pi}^{-1}(U) &\longrightarrow U \times F\,,\\ [b,m] &\longmapsto \big(\pi_P(p), \varphi(p)m\big)\,.\end{aligned}\tag{9.28}$$

This map is a diffeomorphism. Indeed, let $s : U \longrightarrow \pi_P^{-1}(U)$ be the (local) section associated to the above bundle chart of the principal bundle and let

$$\begin{aligned}f : U \times F &\longrightarrow \tilde{\pi}^{-1}(U)\,,\\ (p,m) &\longmapsto [s(b),m]\,.\end{aligned}\tag{9.29}$$

In order to show that f is the inverse of $(\tilde{\pi},\tilde{\varphi})$ defined in (9.28), we observe that

$$(\tilde{\pi},\tilde{\varphi})\circ f(b,m) = (\tilde{\pi},\tilde{\varphi})\big([s(b),m]\big) = \Big(\pi_P\big(s(b)\big),\varphi\big(s(b)\big)m\Big) = (b,m)\,.$$

On the other hand, given $p \in \pi_P^{-1}(U)$ and noting that the (local) section can be written as $s\big(\pi_P(p)\big) = p\varphi(p)^{-1}$, we have

$$\begin{aligned} f\circ(\tilde{\pi},\tilde{\varphi})\big([p,m]\big) &= f\big(\pi_P(p),\varphi(p)m\big) = \big[s\big(\pi_P(p)\big),\varphi(p)m\big] \\ &= \big[p\varphi(p)^{-1},\varphi(p)m\big] = [p,m]\,. \end{aligned}$$

Next, let $\big(\pi_P^{-1}(V),(\pi_P,\psi)\big)$ be another chart of the principal bundle with $U\cap V \neq \emptyset$ and let $b \in B$ be an element of this intersection. Defining $(\tilde{\pi},\tilde{\psi}) : \tilde{\pi}^{-1}(V) \longrightarrow V\times F$ as in (9.28), we obtain

$$\begin{aligned} \tilde{\psi}\circ(\tilde{\pi},\tilde{\varphi})^{-1}(b,m) &\overset{(9.29)}{=} \tilde{\psi}\big([s(b),m]\big) = \tilde{\psi}\big([(\pi_P,\varphi)^{-1}(b,e),m]\big) \\ &\overset{(9.28)}{=} \big(\psi\circ(\pi_P,\varphi)^{-1}(b,e)\big)\,m = \big(f_{\varphi,\psi}(b)e\big)\,m = f_{\varphi,\psi}(b)\,m\,, \end{aligned}$$

where (9.2) for the transition function $f_{\varphi,\psi}$ of π_P is also used. In summary, we have constructed a bundle atlas for the fiber bundle $\tilde{\pi} : P\times_G F \longrightarrow B$ with fiber F and structure group G. – The fact that the transition functions of the fiber bundle $\tilde{\pi} : P\times_G F \longrightarrow B$ coincide with those of $\pi_P : P \longrightarrow B$ implies directly that π_P is associated to $\tilde{\pi}$. □

Remark. a) Note that the fiber bundle $\tilde{\pi} : P\times_G F \longrightarrow B$ is also often said to be *associated* to the principal bundle $\pi_P : P \longrightarrow B$.

b) In the case of $F = \mathfrak{g}$ with left action of G on $\mathfrak{g}$ given by the adjoint action $Ad : G\times\mathfrak{g} \longrightarrow \mathfrak{g}$ defined in (7.37), we obtain the so-called *adjoint bundle* $P\times_{Ad}\mathfrak{g}$. This vector bundle will be important in the following.

Reducible Principal Bundles

Definition 9.17. *Let H be a Lie subgroup of G and P_H, P_G two manifolds with $P_H \subset P_G$. Moreover, let $\pi_G : P_G \longrightarrow B$ and $\pi_H : P_H \longrightarrow B$ be a principal G-bundle, respectively, a principal H-bundle over B such that the injection $f : P_H \hookrightarrow P_G$ is a fiber preserving map and induces the identity transformation on B (see Fig. ??). We then say that the structure group G of π_G is* ***reducible*** *to H and we call π_H the* ***reduced subbundle*** *of π_G.*

Fig. 9.4. Reduction of the structure group. The map f satisfies $\pi_G\circ f = \pi_H$.

Proposition 9.18. *The structure group G of a principal bundle $\pi_G : P_G \longrightarrow B$ is reducible to e – neutral element of G – if and only if π_G is trivial.*

Proof. First, we assume that $\pi_G : P_G \longrightarrow B$ is trivial. By definition this means that there exists a bundle isomorphism $h : B \times G \longrightarrow P_G$. Considering the map

$$(id_B, \iota_e) : B \times \{e\} \longrightarrow B \times G\,,$$

we conclude that $f = h \circ (id_B, \iota_e)$ reduces G to e.

Conversely, assume that $f : B \times \{e\} \longrightarrow P$ is a fiber preserving map such that $\pi_G \circ f = id_B$. Then it follows that the map

$$\begin{aligned} s : B &\longrightarrow P\,,\\ b &\longmapsto f(b,e) \end{aligned}$$

is a section of P. Indeed, we have $\pi_G\big(s(b)\big) = \pi_G\big(f(b,e)\big) = b$. Theorem 9.12 then implies that P is trivial. $\square$

Proposition 9.19. *The structure group G of a principal bundle $\pi_G : P_G \longrightarrow B$ is reducible to a Lie subgroup H if and only if there exists a bundle atlas $\big\{\big(\pi_G^{-1}(U_i), (\pi_G, \varphi_i)\big)\big\}_{i\in I}$ of π_G with transition functions $f_{i,j}$ taking their values in H.*

Proof. Suppose first that the structure group G is reducible to H and denote by $\pi_H : P_H \longrightarrow B$ the reduced subbundle, where P_H is considered as submanifold of P_G. Moreover, let $\big\{\big(\pi_H^{-1}(V_i), (\pi_H, \psi_i)\big)\big\}_{i\in I}$ be a bundle atlas for π_H whose corresponding transition functions clearly take their values in H. – We now look for a bundle atlas of $\pi_G : P_G \longrightarrow B$ whose transition functions take their values only in H.

Take again $\{V_i = U_i\}_{i\in I}$ as open covering for B. By assumption every $p \in \pi_G^{-1}(U_i)$ may be represented in the form $p = qg$, for some $q \in \pi_H^{-1}(U_i)$ and $g \in G$. Then, we set

$$\varphi_i(p) = \psi_i(q)g\,.$$

Note that the map $\varphi : \pi_G^{-1}(U_i) \longrightarrow G$ is well-defined meaning that it is independent of the representation for $p \in \pi_G^{-1}(U_i)$. Indeed, let $p = \tilde{q}\tilde{g}$ be another representation, then we have

$$\varphi(p) = \psi_i(\tilde{q})\tilde{g} = \psi_i(p\tilde{g}^{-1})\tilde{g} = \psi_i(p) = \psi(qg) = \psi(q)g\,.$$

It is also clear that

$$(\pi_G, \varphi_i) : \pi_G^{-1}(U_i) \longrightarrow U_i \times G \tag{9.30}$$

defines a diffeomorphism. Using (9.6), we find for the corresponding transition functions

$$f_{i,j}\big(\pi_G(p)\big) = \varphi_j(p)\varphi_i(p)^{-1} = \psi_j(q)\psi_i(q)^{-1}\,,$$

taking their values in H. Thus (9.30) is the desired bundle atlas for π_G.

For the converse, we only give a sketch of the proof. – Assume that there exists a bundle atlas $\big\{\big(\pi_G^{-1}(U_i), (\pi_G, \varphi_i)\big)\big\}_{i\in I}$ of $\pi_G : P_G \longrightarrow B$ with corresponding transition functions $f_{i,j}$ taking their values in H. By Proposition

9.14, we can construct a principal H-bundle from the covering $\{U_i\}_{i\in I}$ and the transition functions $f_{i,j}$ which we will denote by $\pi_H : P_H \longrightarrow B$. Next, we define the map $f_i : \pi_H^{-1}(U_i) \longrightarrow \pi_G^{-1}(U_i)$ to be the composition of the following three maps:

$$\pi_H^{-1}(U_i) \longrightarrow U_i \times H \hookrightarrow U_i \times G \overset{(\pi_G,\varphi_i)^{-1}}{\longrightarrow} \pi_G^{-1}(U_i)\,.$$

Obviously, we have $f_i = f_j$ on $U_i \cap U_j \neq \emptyset$. It is left as an exercise to check that $f : P_H \longrightarrow P_G$ given by $\{f_i\}_{i\in I}$ reduces G to H in the sense of Definition 9.17. □

9.4 Connections in a Principal Fiber Bundle

Let $\pi : P \longrightarrow B$ be a principal G-bundle over B. For each $p \in P$, we denote by $\mathcal{V}_p$ the subspace of the tangent space T_pP of P at p consisting of vectors *tangent* to the fiber over $\pi(p)$. To be more precise, elements in $\mathcal{V}_p$ are those being tangent to the orbit $G_p = \{pg \ : \ g \in G\}$ of p – recall that for a principal G-bundle we have $B = P/G$. In other words, we have $\mathcal{V}_p = \ker d\pi_p$. We will call $\mathcal{V}_p$ the *vertical subspace* of T_pP.

The right action on P induces a homomorphism σ of the Lie algebra $\mathfrak{g}$ of G into the Lie algebra $\mathcal{X}(P)$ of vector fields on P by Proposition 7.29. For every $A \in \mathfrak{g}$, we call $\sigma(A) = A^*$ the **fundamental vector field** corresponding to A. Then we make the following important observation: As a direct consequence of (7.27) the fundamental vector field $A^*(p)$ is an element of $\mathcal{V}_p \subset T_pP$, for all $p \in P$. We also know from Proposition 7.29 that $\sigma : \mathfrak{g} \longrightarrow \mathcal{X}(P)$ is an injective map because the action of G is free. Since the dimension of each vertical subspace equals that of the Lie algebra $\mathfrak{g}$, we then conclude that the linear map

$$\begin{aligned} d\sigma_p : \mathfrak{g} &\longrightarrow \mathcal{V}_p\,, \\ A &\longmapsto A^*(p) \end{aligned} \tag{9.31}$$

is an isomorphism. – Recall the computation in (7.29), where we showed that the fundamental vector field can also be expressed in terms of the differential $d\sigma_p$ of σ_p.

Proposition 9.20. *Let A^* be the fundamental vector field corresponding to $A \in \mathfrak{g}$. Then, for each $g \in G$, we have that $(R_g)_*A^*$ is the fundamental vector field for $Ad_{g^{-1}}(A) \in \mathfrak{g}$.*

Proof. By Definition 2.47 of the push-forward we have $\big((R_g)_*A^*\big)(p) = (dR_g)_{pg^{-1}} \cdot A^*(pg^{-1})$ and (7.32) then shows that $\big((R_g)_*A^*\big)(p) = d\sigma_p \cdot Ad_{g^{-1}}\big(X_A(e)\big) = d\sigma_p \cdot Ad_{g^{-1}}(A)$. We conclude the proposition using (9.31). □

Definition 9.21. *Let M^n denote an n-dimensional manifold. A **connection** on a principal G-bundle $\pi : P \longrightarrow M$ is an n-dimensional differentiable distribution $\mathcal{H}$ on P satisfying the following conditions:*

(i) $\pi_* \mathcal{H} = TM$ *;*
(ii) $(R_g)_* \mathcal{H} = \mathcal{H}$, *for all* $g \in G$.

Remark. The following remarks concerning the previous proposition are important:

a) By (i) there is a splitting $T_pP = \mathcal{H}_p \oplus \ker d\pi_p = \mathcal{H}_p \oplus \mathcal{V}_p$, for every $p \in P$. In particular, every vector field $X \in \mathcal{X}(P)$ can be written as
$$X_p = X_p^H + X_p^V \ ,$$
where $X_p^H \in \mathcal{H}_p$ and $X_p^V \in \mathcal{V}_p$ are called the *horizontal*, respectively, the *vertical* component of X at p. Since $\mathcal{H}$ is an n-dimensional differentiable distribution on P, we obviously have that X_p^H induces a differentiable horizontal vector field X^H on P.
b) Note also that condition (ii) means that $\mathcal{H}$ is invariant by the action of G on P.
c) We emphasize that there is *no* canonical differential distribution $\mathcal{H}$. In other words, by a connection we mean the choice of a differential distribution such that (i) and (ii) hold.

Given a connection $\mathcal{H}$ on P, it is possible to define an one-form on P – which we will denote by ω – with values in the Lie algebra $\mathfrak{g}$ as follows:

Definition 9.22. *The **connection form** ω of a connection $\mathcal{H}$ on a principal G-bundle $\pi : P \longrightarrow M$ is the $\mathfrak{g}$-valued one-form on P given by*
$$\omega_p(X_p) = (d\sigma_p)^{-1} \cdot X_p^V \ , \tag{9.32}$$
where $X_p \in T_pP$. In other words, for each $X_p \in T_pP$, we define $\omega_p(X_p)$ to be the unique $A \in \mathfrak{g}$ such that $A^*(p) = X_p^V$.

Note that because $d\sigma_p$ defined in (9.31) is an isomorphism, for every $p \in P$, the definition makes sense. In the following, we will often write $\omega \in \Omega^1(P, \mathfrak{g})$.

Proposition 9.23. *The connection form $\omega \in \Omega^1(P, \mathfrak{g})$ of a connection $\mathcal{H}$ on a principal G-bundle $\pi : P \longrightarrow M$ satisfies the following:*

(i) $\omega(A^*) = A$, *for all* $A \in \mathfrak{g}$*;*
(ii) $\omega|_{\mathcal{H}} = 0$*;*
(iii) $(R_g)^*\omega = Ad_{g^{-1}} \circ \omega$, *for all* $g \in G$.

Conversely, if $\omega \in \Omega^1(P, \mathfrak{g})$ satisfies (i) and (iii), then $\ker\omega$ is a connection on $\pi : P \longrightarrow M$.

Proof. The conditions (i) and (ii) follow immediately from Definition 9.22. Since every $X \in \mathcal{X}(P)$ can be decomposed into a horizontal and vertical vector field X^H and X^V, respectively, it suffices to show (iii) in the following two particular cases:

a) If X is horizontal, then the right-hand side of (iii) vanishes because of (ii). On the other hand, since $(R_g)_* X$ is also horizontal by (ii) of Definition 9.21, the left-hand side of (iii) also vanishes. This establishes (iii) in the case of X being horizontal.
b) If X is vertical, we can further assume that X is a fundamental vector field A^*. Then Proposition 9.20 implies that $(R_g)_* A^*$ is the fundamental vector field of $Ad_{g^{-1}}(A) \in \mathfrak{g}$. Thus by definition of the connection form ω, it follows

$$\begin{aligned}\big((R_g)^*\omega\big)_p(A_p^*) &= \omega_{pg}(dR_g \cdot A_p^*) \\ &= Ad_{g^{-1}}(A) = Ad_{g^{-1}} \circ \omega_p(A_p^*)\,.\end{aligned}$$

Thus (iii) is proved.

For the converse, we define

$$\mathcal{H}_p = \ker \omega_p = \big\{X_p \in T_pP \ : \ \omega_p(X_p) = 0\big\}\,. \tag{9.33}$$

It is left as an exercise to check that the differential distribution $\mathcal{H}$ thus defined gives a connection for $\pi : P \longrightarrow M$ with connection form ω. □

Connection Form in Local Coordinates

To a given connection form $\omega \in \Omega^1(P, \mathfrak{g})$, we now associate a family of $\mathfrak{g}$-valued one-forms defined on open subsets of the base space M. – Let $\big\{\big(\pi^{-1}(U_i), (\pi, \varphi_i)\big)\big\}_{i\in I}$ be an atlas for a principal G-bundle $\pi : P \longrightarrow M$ with corresponding transition functions $f_{i,j} : U_i \cap U_j \longrightarrow G$. Moreover, for all $i \in I$, let $s_i : U_i \longrightarrow \pi^{-1}(U_i)$ be the (local) section on U_i associated to the chart $\big(\pi^{-1}(U_i), (\pi, \varphi_i)\big)$, i.e.,

$$s_i(b) = \big(\pi, \varphi_i\big)^{-1}(b, e)\,,$$

and let $\theta \in \Omega^1(G, \mathfrak{g})$ be the Maurer-Cartan form of Definition 7.8. We then define a $\mathfrak{g}$-valued one-form A_i on U_i by

$$A_i = s_i^* \omega\,, \tag{9.34}$$

and a $\mathfrak{g}$-valued one-form $\theta_{i,j}$ on $U_i \cap U_j \neq \emptyset$ by

$$\theta_{i,j} = f_{i,j}^* \theta\,. \tag{9.35}$$

Note that in the physics literature the (local) section $s_i : U_i \longrightarrow \pi^{-1}(U_i)$ and the $\mathfrak{g}$-valued one-form A_i are often called **local gauge**, respectively, (local) **gauge potential** in the local gauge s_i.

Proposition 9.24. *Let* $\{(\pi^{-1}(U_i),(\pi,\varphi_i))\}_{i\in I}$ *be an atlas and* ω *a connection form for a principal* G*-bundle* $\pi : P \longrightarrow M$. *Then the one-forms* $A_i \in \Omega^1(U_i,\mathfrak{g})$ *satisfy the following compatibility condition on* $U_i \cap U_j$*:*

$$A_j(b) = Ad_{f_{j,i}(b)^{-1}} \circ A_i(b) + \theta_{j,i}(b)\,. \tag{9.36}$$

Proof. First, for $b \in U_i \cap U_j \neq \emptyset$, we note that

$$s_j(b) = s_i(b) f_{j,i}(b)\,. \tag{9.37}$$

Indeed, we have

$$\begin{aligned}(\pi,\varphi_i)\big(s_j(b)\big) &= \big(\pi,\varphi_i\big)\circ\big(\pi,\varphi_j\big)^{-1}(b,e)\\ &\overset{(9.2)}{=} \big(b, f_{j,i}(b)\big) \overset{(9.5)}{=} (\pi,\varphi_i)\big(s_i(b)f_{j,i}(b)\big)\,.\end{aligned}$$

Consider now the right action $\mathcal{A} : G\times P \longrightarrow P$ of the Lie group G on P. By Leibniz's formula (see []) the image of its differential $d\mathcal{A}$ at the point $(g,p)\in G\times P$ applied to $Z \in T_{(g,p)}G\times P$ equals

$$d\mathcal{A}_{(g,p)}\cdot Z = (d\sigma_p)_g \cdot \tilde{X}_g + (dR_g)_p \cdot \tilde{Y}_p\,, \tag{9.38}$$

where $\sigma_p : G \longrightarrow P$, $R_g : P \longrightarrow P$ are defined as usual and $(\tilde{X}_g,\tilde{Y}_p) \in T_gG \oplus T_pP$ correspond to Z. In the particular case of $(df_{j,i})_b \cdot X_b \in T_{f_{j,i}(b)}G$ and $(ds_i)_b \cdot X_b \in T_{s_i(b)}P$, where $X_b \in T_bM$, the differential in (9.38) reads as

$$d\mathcal{A}_{\left(f_{j,i}(b),s_i(b)\right)}\cdot Z = d\sigma_{s_i(b)}\cdot (df_{j,i})_b \cdot X_b + dR_{f_{j,i}(b)}\cdot (ds_i)_b \cdot X_b\,.$$

Taking the differential of (9.37), we thus arrive at

$$(ds_j)_b \cdot X_b = d\sigma_{s_i(b)}\cdot (df_{j,i})_b \cdot X_b + dR_{f_{j,i}(b)}\cdot (ds_i)_b \cdot X_b\,. \tag{9.39}$$

In a next step, we apply the given connection one-form ω on both sides of (9.39). – Since

$$\omega_{s_j(b)}\big((ds_j)_b\cdot X_b\big) = \big(s_j^*\omega\big)_b(X_b) = A_j(X_b) \tag{9.40}$$

by (9.34), we get $A_j(X_b)$ for the left-hand side of (9.39). For the second term on the right-hand side of (9.39), we have

$$\begin{aligned}\omega_{s_i(b)f_{j,i}(b)}\big(dR_{f_{j,i}(b)}\cdot (ds_i)_b\cdot X_b\big) &= \big(R^*_{f_{j,i}(b)}\omega\big)_{s_i(b)}\big((ds_i)_b\cdot X_b\big)\\ &= Ad_{f_{j,i}(b)^{-1}}\circ \omega_{s_i(b)}\big((ds_i)_b\cdot X_b\big)\\ &\overset{(9.40)}{=} Ad_{f_{j,i}(b)^{-1}}\circ A_i(X_b)\,,\end{aligned}$$

where we used (iii) of Proposition 9.23. Now, we apply ω on the first term of the right-hand side of (9.39). For this purpose, we first consider $X_A \in \mathcal{X}_L(G)$ generated by $A\in\mathfrak{g}$ such that $X_A\big(f_{j,i}(b)\big) = (df_{j,i})_b \cdot X_b$. Hence by definition of the Maurer-Cartan form we have $\theta_{f_{j,i}(b)}\big((df_{j,i})_b \cdot X_b\big) = A$. Using (7.29), we deduce that

$$d\sigma_{s_i(b)} \cdot X_A\big(f_{j,i}(b)\big) = A^*\big(s_i(b)f_{j,i}(b)\big)\,. \tag{9.41}$$

We then conclude that

$$\begin{aligned}\omega_{s_i(b)f_{j,i}(b)}\big(A^*\big(s_i(b)f_{j,i}(b)\big)\big) &= A \\ &= \theta_{f_{j,i}(b)}\big((df_{j,i})_b \cdot X_b\big) \\ &= \big(f_{j,i}^*\theta\big)_b(X_b) \overset{(9.35)}{=} \big(\theta_{j,i}\big)_b(X_b)\,.\end{aligned}$$

Finally, putting the previous results together we arrive at the compatibility condition (9.36). □

Remark. a) Considering the case of a matrix Lie group G the compatibility condition (9.36) takes the simple form

$$A_j(b) = f_{j,i}(b)^{-1}A_i(b)f_{j,i}(b) + f_{j,i}(b)^{-1}(df_{j,i})_b\,. \tag{9.42}$$

b) A change in the choice of the local gauge will be called a **local gauge transformation**. The compatibility condition (9.36) then describes the effect of a local gauge transformation on the gauge potentials representing a fixed connection on P.

The converse of the previous proposition also holds:

Proposition 9.25. *Let $\pi : P \longrightarrow M$ be a principal G-bundle and $\{U_i\}_{i\in I}$ an open covering of M. Moreover, assume that there exist $\mathfrak{g}$-valued one-forms A_i on U_i satisfying (9.36) for every $i \in I$. Then there is a unique connection form ω on P which gives rise to the family $\{A_i\}_{i\in I}$ in the above described manner.*

Gauge transformations

Bundle maps from a principal bundle into itself inducing the identity on the base space play an important role. For the next definition we recall Theorem 9.6.

Definition 9.26. *Let $\pi : P \longrightarrow M$ be a principal G-bundle. Denote by $h : P \longrightarrow P$ a bundle automorphism, i.e., a G-equivariant smooth map inducing the identity on M. The set of all bundle automorphisms form a group $\mathcal{G}$ under composition called the* ***group of gauge transformations*** *for $\pi : P \longrightarrow M$.*

We observe that every bundle automorphism $h : P \longrightarrow P$ can be represented for all $p \in P$ as

$$h(p) = p\,u(p)\,, \tag{9.43}$$

where $u : P \longrightarrow G$ is a smooth map that must satisfy

$$u(pg) = g^{-1}u(p)g\,, \tag{9.44}$$

since h is assumed to be G-equivariant. In the chart $\bigl(\pi^{-1}(U_i), (\pi, \varphi_i)\bigr)$ the local representation of h at $b \in U_i$ is given by

$$\bigl(\varphi_i \circ h\bigr)(p)\varphi_i(p)^{-1} = \varphi_i\bigl(p\,u(p)\bigr)\varphi_i(p)^{-1} = \varphi_i(p)u(p)\varphi_i(p)^{-1}\,,$$

where $p \in \pi^{-1}(b)$ and (9.7) is used. On U_i, we then define the map

$$u_i(b) = \varphi_i\bigl(\pi^{-1}(b)\bigr)u\bigl(\pi^{-1}(b)\bigr)\varphi_i\bigl(\pi^{-1}(b)\bigr)^{-1}\,. \tag{9.45}$$

Note that right-hand side is independent of $q \in \pi^{-1}(b)$, since for $q = pg$ we have

$$\varphi_i(pg)u(pg)\varphi_i(pg)^{-1} = \varphi_i(p)g\,u(pg)g^{-1}\varphi_i(p)^{-1} \stackrel{(9.44)}{=} \varphi_i(p)u(p)\varphi_i(p)^{-1}\,.$$

Note also that if $u_j : U_j \longrightarrow G$ is defined as in (9.45) by another chart $\bigl(\pi^{-1}(U_j), (\pi, \varphi_j)\bigr)$, then the following equation holds for $b \in U_i \cap U_j \neq \emptyset$:

$$u_j(b) = f_{i,j}(b)u_i(b)f_{i,j}(b)^{-1}\,. \tag{9.46}$$

Next, we give two other interpretations of the map $u : P \longrightarrow G$.

a) Let $s_i : U_i \longrightarrow \pi^{-1}(U_i)$ denote the (local) section associated the the bundle chart $\bigl(\pi^{-1}(U_i), (\pi, \varphi_i)\bigr)$ of the principal G-bundle π. Noting that $s_i\bigl(\pi(p)\bigr) = p\varphi_i(p)^{-1}$, we obtain for $p \in \pi^{-1}(U_i)$ that

$$\begin{aligned}\bigl(s_i^*u\bigr)\bigl(\pi(p)\bigr) &= u \circ s_i\bigl(\pi(p)\bigr) = u\bigl(p\varphi_i(p)^{-1}\bigr)\\ &\stackrel{(9.44)}{=} \varphi_i(p)u(p)\varphi_i(p)^{-1} \stackrel{(9.45)}{=} u_i(b)\,.\end{aligned} \tag{9.47}$$

b) The map $u : P \longrightarrow G$ can also be seen as section of the associated bundle $\tilde{\pi} : P \times_c G \longrightarrow M$, where c denotes the action by conjugation on G defined in (7.36). – We consider the (local) section of the associated bundle $\tilde{\pi}$ given by

$$\begin{aligned}\tilde{s}_i : U_i &\longrightarrow P \times_c G\,,\\ \tilde{\pi}(p) &\longmapsto \bigl[p, u(p)\bigr]\,.\end{aligned} \tag{9.48}$$

Note that he right-hand side is independent of $p \in P$ by definition of the associated bundle in Section 9.3. Moreover, using the bundle chart $\bigl(\tilde{\pi}^{-1}(U_i), (\tilde{\pi}, \tilde{\varphi}_i)\bigr)$ for $\tilde{\pi}$ constructed in the proof of Theorem 9.16, we see that the map $\tilde{\varphi}_i \circ \tilde{s}_i : U_i \longrightarrow G$ agrees with (9.45).

Now, we want to determine how a connection form transforms under a bundle automorphism. – Let $\omega \in \Omega^1(P, \mathfrak{g})$ be a connection form on the principal G-bundle $\pi : P \longrightarrow M$. Clearly, the pull-back $h^*\omega$ by the bundle automorphism $h : P \longrightarrow P$ is again a connection form on P. For $X_p \in T_pP$, we have

$$\big(h^*\omega\big)_p(X_p) = \omega_{h(p)}(dh_p \cdot X_p)\,.$$

In order to compute the right-hand side, we use (9.43) for the differential dh and then proceed as in the proof of Proposition 9.24 to obtain[2]

$$\big(h^*\omega\big)_p(X_p) = u(p)^{-1}\omega_p(X_p)u(p) + u(b)^{-1}du_p \cdot X_p\,. \tag{9.49}$$

With the (local) section $s_i : U_i \longrightarrow \pi^{-1}(U_i)$ associated the bundle chart, we obtain using (9.47) that

$$\begin{aligned}\big(s_i^*(u^{-1}\omega u)\big)_{\pi(p)}(X_b) &= u^{-1} \circ s_i\big(\pi(p)\big)\ (s_i^*\omega)_{\pi(p)}(X_b)\ u \circ s_i\big(\pi(p)\big)\\ &= u_i(b)^{-1}\ (s_i^*\omega)_{\pi(p)}(X_b)\ u_i(b)\,,\end{aligned}$$

where $b = \pi(p) \in U_i$. Applying the pull-back of s_i on both sides of (9.49), we thus arrive at

$$s_i^*\big(h^*\omega\big)(b) = u_i(b)^{-1}A_i(b)u_i(b) + u_i(b)^{-1}(du_i)_b\,, \tag{9.50}$$

where $A_i = s_i^*\omega$ denote the gauge potentials introduced in (9.34).

Remark. Comparing with (9.42), we conclude that the above transformation formula can be interpreted equivalently as the effect of a local gauge transformation on the gauge potentials representing a fixed connection form, or as effect of a gauge transformation on a connection form, viewed in a fixed local gauge.

Horizontal Lift and Parallel Transport

Given a connection on a principal G-bundle $\pi : P \longrightarrow M$, we can now define the concept of parallel transport of fibers along any given curve in the base space M.

Definition 9.27. *The* ***horizontal lift*** *of a vector field $X \in \mathcal{X}(M)$ on M is a vector field $\tilde{X} \in \mathcal{X}(P)$ on P such that $(d\pi)_p \cdot \tilde{X}_p = X_{\pi(p)}$, for every $p \in P$.*

We observe that – as a direct consequence of Definition 9.21 – a connection makes the linear map $(d\pi)_p : \mathcal{H}_p \longrightarrow T_{\pi(p)}M$ to an isomorphism. This leads to the following result.

Proposition 9.28. *Let $X \in \mathcal{X}(M)$ and $\mathcal{H}$ be a connection on a principal G-bundle $\pi : P \longrightarrow M$. Then there exists a unique horizontal lift $\tilde{X} \in \mathcal{X}(P)$ of X. Moreover, the horizontal lift $\tilde{X}$ is invariant by R_g, for every $g \in G$. Conversely, for every horizontal vector field $Y \in \mathcal{X}(P)$ being invariant by R_g, there exists a vector field $X \in \mathcal{X}(M)$ whose horizontal lift equals Y.*

[2] The computation is completely analogous to the one in the proof of Proposition 9.24 except that (9.37) is replaced by (9.43).

Proof. The existence and uniqueness of the horizontal lift $\tilde{X}$ is clear from the above observation. It remains to show that $\tilde{X} \in \mathcal{X}(P)$, i.e., that $\tilde{X}$ is *smooth.*

For this purpose, let $b \in M$ and consider a chart $\big(\pi^{-1}(U), (\pi, \varphi)\big)$ of $\pi : P \longrightarrow M$ about $\pi^{-1}(b)$ which is by definition a diffeomorphism between $\pi^{-1}(U)$ and $U \times G$. Using this diffeomorphism, we deduce the existence of a smooth vector field Z on $\pi^{-1}(U)$ such that $X_c = (d\pi)_{\pi^{-1}(c)} \cdot Z_{\pi^{-1}(c)}$, for all $c \in U$. Then due to the uniqueness we have $\tilde{X} = Z^H$ and the horizontal vector field Z^H on $\pi^{-1}(U)$ remains smooth. – The fact that $\tilde{X}$ is invariant by R_g, for every $g \in G$, is a direct consequence of the R_g-invariance of the connection in (ii) of Definition 9.21.

Conversely, let $Y \in \mathcal{X}(P)$ be a horizontal R_g-invariant vector field on P. For every $b \in M$ we then choose a point $p \in \pi^{-1}(b)$ and set $X_b = (d\pi)_p \cdot Y_p$. Note that $X_b \in T_bM$ is independent of the choice of p in the fiber $\pi^{-1}(b)$ over b. Indeed, if $q = pg \in \pi^{-1}(b)$ and using the properties of Y, we have

$$(d\pi)_q \cdot Y_q = (d\pi)_{pg} \cdot \big((dR_g)_p \cdot Y_p\big) = (d\pi)_p \cdot Y_p\,.$$

The smoothness of the resulting vector field X on M is clear. □

The notion of horizontal lift of a vector field X on M naturally leads to the following concept of horizontal lift of a *curve* in M: The integral curve of the horizontal lift $\tilde{X}$ through a point $p_0 \in P$ will be defined to be the horizontal lift of the integral curve of X through the point $b_0 = \pi(p_0) \in M$. More precisely, we have

Definition 9.29. *Let $\gamma : [0,1] \longrightarrow M$ be a C^1-curve in M. A* ***horizontal lift*** *of γ is a horizontal curve $\tilde{\gamma} : [0,1] \longrightarrow P$ in P such that $\pi \circ \tilde{\gamma} = \gamma$.*

Remark. Note that *horizontal* curve means that $\dot{\tilde{\gamma}}(t) \in \mathcal{H}_{\tilde{\gamma}(t)}$, for all $t \in [0,1]$, where $\dot{\tilde{\gamma}}(t)$ denotes the tangent vector of $\tilde{\gamma}$ at the point $\tilde{\gamma}(t)$.

Theorem 9.30. *Let $\gamma : [0,1] \longrightarrow M$ be a C^1-curve in M and $p_0 \in P$ with $\pi(p_0) = \gamma(0) \in M$. Then there exists a unique horizontal lift $\tilde{\gamma} : [0,1] \longrightarrow P$ of γ through p_0 meaning that $\tilde{\gamma}(0) = p_0$.*

Proof. Let $t_0 \in [0,1]$ and choose $I \subset [0,1]$ containing t_0 such that $\gamma(I) \subset U$ with U open subset of M. By local triviality of the principal G-bundle $\pi : P \longrightarrow M$, we deduce the existence of a C^1-curve $c : I \longrightarrow P$ such that $c(t_0) = q_0$ with $q_0 \in P$ satisfying $\pi(q_0) = \gamma(t_0) \in M$ and $\pi\big(c(t)\big) = \gamma(t)$, for all $t \in I$. Repeating this argument for an open covering of M, there exists a C^1-curve $c : [0,1] \longrightarrow P$ such that $c(0) = p_0$ and $\pi\big(c(t)\big) = \gamma(t)$, for all $t \in [0,1]$.

Next, we observe that if the horizontal lift $\tilde{\gamma} : [0,1] \longrightarrow P$ of γ exists, it must be of the form

$$\tilde{\gamma}(t) = c(t)a(t) \qquad \text{for } t \in [0,1]\,, \tag{9.51}$$

where $a : [0,1] \longrightarrow G$ is a curve in the structure group G with $a(0) = e$.

In order to determine the curve $a : [0,1] \longrightarrow G$ making $\tilde{\gamma}$ to a horizontal one, we proceed again as in the proof of Proposition 9.24. First, we apply Leibniz's formula to the right action $\mathcal{A} : G \times P \longrightarrow P$ and obtain by differentiating (9.51) that

$$\dot{\tilde{\gamma}}(t) = d\sigma_{c(t)} \cdot \dot{a}(t) + dR_{a(t)} \cdot \dot{c}(t)\,.$$

Applying the connection form ω to both sides, we then deduce

$$\omega\big(\dot{\tilde{\gamma}}(t)\big) = Ad_{a(t)^{-1}} \circ \omega\big(\dot{c}(t)\big) + a(t)^{-1}\dot{a}(t)\,,$$

where $a(t)^{-1}\dot{a}(t) : [0,1] \longrightarrow \mathfrak{g}$ denotes the curve $\theta_{a(t)}\big(\dot{a}(t)\big) = \big(dL_{a(t)^{-1}}\big)_{a(t)} \cdot \dot{a}(t)$ in the Lie algebra $\mathfrak{g}$. Thus we arrive at the following condition for $\tilde{\gamma}$ to be horizontal:

$$0 = Ad_{a(t)^{-1}} \circ \omega\big(\dot{c}(t)\big) + a(t)^{-1}\dot{a}(t)\,,$$

which is equivalent to

$$\dot{a}(t)a(t)^{-1} = -\omega\big(\dot{c}(t)\big)\,, \tag{9.52}$$

where $\dot{a}(t)a(t)^{-1}$ denotes the curve $\big(dR_{a(t)^{-1}}\big)_{a(t)} \cdot \dot{a}(t)$ in $\mathfrak{g}$. Lemma 9.31 below then gives a solution $a : [0,1] \longrightarrow G$ for (9.52). – Thus, we have constructed a horizontal lift $\tilde{\gamma}$ of γ. □

Lemma 9.31. *Let $\mathfrak{g}$ be the Lie algebra of a Lie group G. Moreover, assume that $C : [0,1] \longrightarrow \mathfrak{g}$ is a continuous curve in $\mathfrak{g}$. Then, there exists a unique C^1-curve $a : [0,1] \longrightarrow G$ in G such that*

$$\dot{a}(t)a(t)^{-1} = C(t)\,, \qquad a(0) = e\,, \tag{9.53}$$

for all $t \in [0,1]$.

Proof. In the particular case of a constant curve $C(t) = A \in \mathfrak{g}$, it follows for (9.53) that

$$\big(dR_{a(t)^{-1}}\big)_{a(t)} \cdot \dot{a}(t) = \dot{a}(t)a(t)^{-1} = A\,.$$

This equation is satisfied by $a(t) = \Gamma_t^{X_A}(e) = R_{\exp(tA)}\, e$.

Returning to the general case, suppose that $C(t)$ is defined for all $t \in \mathbb{R}$. We then consider a vector field X on $\mathbb{R} \times G$ having the following value at $(t,g) \in \mathbb{R} \times G$:

$$X(t,g) = \left(\frac{\partial}{\partial x}(t), (dR_g)_e \cdot C(t)\right) \in T_t\mathbb{R} \times T_gG\,,$$

where x denotes the canonical coordinate function on $\mathbb{R}$. It is clear that there exists a unique (local) integral curve starting at $(0,e)$ which we write as $t \longmapsto \big(t, a(t)\big) \in \mathbb{R} \times G$. Since $\dot{a}(t) = (dR_{a(t)})_e \cdot C(t)$, we conclude that $a(t)$ is the desired C^1-curve. It remains to show that $a(t)$ is defined for all $t \in [0,1]$. □

With the help of the previous theorem we are now ready to define the parallel transport of fibers as follows:

Definition 9.32. *Let $\gamma : [0,1] \longrightarrow M$ be a C^1-curve. We call the map*

$$\begin{aligned} \tau_\gamma : \pi^{-1}\big(\gamma(0)\big) &\longrightarrow \pi^{-1}\big(\gamma(1)\big)\,, \\ p &\longmapsto \tilde{\gamma}_p(1)\,. \end{aligned}$$

where $\tilde{\gamma}_p$ denotes the horizontal lift of γ through p, the ***parallel transport*** *along γ.*

Remark. Recall that the horizontal lift depends on the connection $\mathcal{H}$ for $\pi : P \longrightarrow M$ (see (9.52)). To be more precise we should call τ_γ the parallel transport of γ *with respect to $\mathcal{H}$.*

The next proposition shows that τ_γ is an isomorphism between $\pi^{-1}\big(\gamma(0)\big)$ and $\pi^{-1}\big(\gamma(1)\big)$.

Proposition 9.33. *Let $\gamma : [0,1] \longrightarrow M$ be a C^1-curve. Then the parallel transport commutes with the right action on G, i.e.,*

$$\tau_\gamma \circ R_g = R_g \circ \tau_\gamma\,,$$

for all $g \in G$.

Proof. Let $p \in \pi^{-1}\big(\gamma(0)\big)$ and observe that $R_g \circ \tilde{\gamma}_p(1) = \tilde{\gamma}_{pg}(1)$ as a consequence of the fact that R_g maps horizontal curves in P into horizontal ones (see (ii) of Definition (9.21)). By uniqueness of $\tilde{\gamma}_{pg}(1)$ the assertion follows. □

9.5 Curvature of a Connection

Consider again a principal G-bundle $\pi : P \longrightarrow M$. Let α be a Lie algebra-valued s-form on P, i.e., $\alpha \in \Omega^s(P, \mathfrak{g})$. We say that α is (*right*) *equivariant* if it satisfies

$$(R_g)^*\alpha = Ad_{g^{-1}} \circ \alpha\,, \tag{9.54}$$

for all $g \in G$, and α is said to be *horizontal* if

$$\alpha(X_1, \ldots, X_s) = 0\,,$$

whenever at least one of the vector fields $X_1, \ldots, X_s \in \mathcal{X}(P)$ is vertical.

Example 9.34. A connection form ω on $\pi : P \longrightarrow M$ is an equivariant one-form on P.

There is an interesting interpretation of such forms: Every equivariant and horizontal Lie algebra-valued s-form $\alpha \in \Omega^s(P, \mathfrak{g})$ on P can be regarded as smooth section of $\bigwedge^s T^*M \otimes P \times_{Ad} \mathfrak{g}$. In other words, α is an s-form on M with values in $P \times_{Ad} \mathfrak{g}$ defined in Section 9.3. Indeed, let $\alpha \in \Omega^s(P, \mathfrak{g})$ be equivariant and horizontal. Then we define, for all $X_1, \ldots, X_s \in T_pP$, the following s-form $\tilde{\alpha}$ on M:

$$\tilde{\alpha}_{\pi(p)}(d\pi_p \cdot X_1, \ldots, d\pi_p \cdot X_s) = \big[p, \alpha_p(X_1, \ldots, X_s)\big]\,. \tag{9.55}$$

It is left as an exercise to check that this definition is independent of p and $X_1, \ldots, X_s$. – Note that this construction is completely similar to the one in (9.48).

Remark. Since a connection form ω is only equivariant and *not* horizontal, there is no interpretation as a section of $\bigwedge^1 T^*M \otimes P \times_{Ad} \mathfrak{g}$. However, if we fix a connection $\tilde{\omega}$, then the difference $\alpha := \omega - \tilde{\omega}$ – which is horizontal – can be regarded as section of $\bigwedge^1 T^*M \otimes P \times_{Ad} \mathfrak{g}$. Conversely, given a connection ω and α a section of $\bigwedge^1 T^*M \otimes P \times_{Ad} \mathfrak{g}$, then $\tilde{\omega} := \omega + \alpha$ is again a connection form on P. In other words, the space of connections on P is an affine space.

Definition 9.35. *Let $\pi : P \longrightarrow M$ be a principal G-bundle with connection $\mathcal{H}$ and $\alpha \in \Omega^s(P, \mathfrak{g})$. We define the* ***covariant derivative*** *$D\alpha$ of α with respect to $\mathcal{H}$ to be the following element of $\Omega^{s+1}(P, \mathfrak{g})$:*

$$D\alpha(X_1, \ldots, X_{s+1}) = d\alpha(X_1^H, \ldots, X_{s+1}^H)\,,$$

where $X_1^H, \ldots, X_{s+1}^H \in \mathcal{H}$ are the horizontal components of the vector fields $X_1, \ldots, X_{s+1} \in \mathcal{X}(P)$.

Proposition 9.36. *Let $\alpha \in \Omega^s(P, \mathfrak{g})$ be an equivariant Lie algebra-valued s-form on P. Then $D\alpha \in \Omega^{s+1}(P, \mathfrak{g})$ is an equivariant and horizontal $(s+1)$-form on P.*

Proof. As a direct consequence of Definition 9.21, we observe that $\big((R_g)_*Y\big)^H = (R_g)_*Y^H$, for every $Y \in \mathcal{X}(P)$. This implies that

$$\begin{aligned}(R_g)^*D\alpha(X_1, \ldots, X_{s+1}) &= D\alpha\big((R_g)_*X_1, \ldots, (R_g)_*X_{s+1}\big)\\ &= d\alpha\Big(\big((R_g)_*X_1\big)^H, \ldots, \big((R_g)_*X_{s+1}\big)^H\Big)\\ &= d\alpha\big((R_g)_*X_1^H, \ldots, (R_g)_*X_{s+1}^H\big)\\ &= (R_g)^*d\alpha(X_1^H, \ldots, X_{s+1}^H)\,.\end{aligned}$$

Since pull-back and exterior derivative commute, it follows that

$$(R_g)^*D\alpha(X_1, \ldots, X_{s+1}) = d\big((R_g)^*\alpha\big)(X_1^H, \ldots, X_{s+1}^H)\,.$$

By assumption α is an equivariant s-form. Hence, we obtain

$$\begin{aligned}(R_g)^* D\alpha(X_1,\ldots,X_{s+1}) &= Ad_{g^{-1}} \circ d\alpha(X_1^H,\ldots,X_{s+1}^H)\\ &= Ad_{g^{-1}} \circ D\alpha(X_1,\ldots,X_{s+1}),\end{aligned}$$

showing that $D\alpha \in \Omega^{s+1}(P,\mathfrak{g})$ is also equivariant. – The fact that $D\alpha$ is horizontal is obvious. $\square$

Definition 9.37. *Let $\mathcal{H}$ be a connection for a principal G-bundle $\pi : P \longrightarrow M$ with connection form ω. We define the **curvature form** Ω to be the covariant derivative $D\omega \in \Omega^2(P,\mathfrak{g})$ of ω with respect to $\mathcal{H}$.*

Remark. Note that because of the previous proposition the curvature form is an equivariant and horizontal two-form on P. Explicitly, we have

$$(R_g)^*\Omega = Ad_{g^{-1}} \circ \Omega. \tag{9.56}$$

Moreover, the curvature form $\Omega \in \Omega^2(P,\mathfrak{g})$ can be interpreted as smooth section of $\bigwedge^2 T^*M \otimes P \times_{Ad} \mathfrak{g}$ as mentioned before.

Theorem 9.38 (Cartan's Structure Equation). *Let ω be a connection form on a principal G-bundle $\pi : P \longrightarrow M$ and $\Omega \in \Omega^2(P,\mathfrak{g})$ its connection form. Then, we have*

$$\Omega_p(X_p,Y_p) = (d\omega)_p\big(X_p,Y_p\big) + \big[\omega_p(X_p),\omega_p(Y_p)\big], \tag{9.57}$$

for all $X_p, Y_p \in T_pP$.

Proof. First note that by assumption $X_p, Y_p \in T_pP$ can be decomposed uniquely into their vertical and horizontal components. Moreover, since both sides of (9.57) are bilinear and skew-symmetric in X_p and Y_p, it is sufficient to show Cartan's structure equation in the following three special cases:

a) If X_p and Y_p are both horizontal, then $\omega_p(X_p) = \omega_p(Y_p) = 0$. Thus, we see that (9.57) reduces to the definition of the curvature form Ω.

b) Let X_p and Y_p be both vertical. Then we consider $A, B \in \mathfrak{g}$ such that their corresponding fundamental vector fields A^* and B^* satisfy $A^*(p) = X_p$, respectively, $B^*(p) = Y_p$. Using formula (6.50) for the exterior derivative of ω, we obtain

$$d\omega(A^*,B^*) = d\big(\omega(B^*)\big)\cdot A^* - d\big(\omega(A^*)\big)\cdot B^* - \omega\big([A^*,B^*]\big). \tag{9.58}$$

Since $\omega(A^*) = A$ and $\omega(B^*) = B$ are independent of $p \in P$, we deduce that the first and second term on the right-hand side of (9.58) vanish. From Proposition (7.29) we also deduce that

$$d\omega(A^*,B^*) = -\omega\big([A^*,B^*]\big) = -\omega\big([A,B]^*\big).$$

This can be written as

$$d\omega(A^*,B^*) = -[A,B] = -\big[\omega(A^*),\omega(B^*)\big].$$

Since $A^*(p) = X_p$ and $B^*(p) = Y_p$, we arrive at

$$(d\omega)_p(X_p, Y_p) = -\big[\omega_p(X_p), \omega_p(Y_p)\big], .$$

On the other hand, we see that $\Omega_p(X_p, Y_p) = 0$ for vertical X_p and Y_p by definition of the curvature form.

c) Let X_p be horizontal and Y_p vertical. We extend X_p to a horizontal vector field X on P and consider a fundamental vector field B^* with $B^*(p) = Y_p$ as before. Using again (6.50), we obtain

$$d\omega(X, B^*) = d\big(\omega(B^*)\big) \cdot X - d\big(\omega(X)\big) \cdot B^* - \omega\big([X, B^*]\big) . \tag{9.59}$$

Since $\omega(X) = 0$ and $d\big(\omega(B^*)\big) \cdot X$ vanishes as before, (9.59) simplifies to

$$d\omega(X, B^*) = -\omega\big([X, B^*]\big) .$$

From Lemma 9.39 below we conclude that the right-hand side of the last equation vanishes. On the other hand, we clearly have $\Omega_p(X_p, Y_p) = 0$. – Thus we have established that Cartan's formula also holds in the third case and therefore (9.57) is proved. □

Lemma 9.39. *Let $A^* \in \mathcal{X}(P)$ be the fundamental vector field to $A \in \mathfrak{g}$ and $X \in \mathcal{X}(P)$ a horizontal vector field. Then, the bracket $[X, A^*]$ is a horizontal vector field on P.*

Proof. Recall from the proof of Proposition 7.29 that $R_{a_t} : P \longrightarrow P$ is the flow of A^*. Hence, by definition of the bracket of vector fields

$$[X, A^*] = \lim_{t \longrightarrow 0} \frac{1}{t}\big((R_{a_t})_* X - X\big) .$$

From the assumption that X is horizontal and Definition 9.21, it follows that $(R_{a_t})_* X$ is also horizontal. Thus, we deduce the proposition. □

Corollary 9.40. *Let $\mathcal{H}$ be a connection on a principal G-bundle $\pi : P \longrightarrow M$ with connection form ω and $\Omega \in \Omega^2(P, \mathfrak{g})$ the curvature form with respect to $\mathcal{H}$. Assume also that $X, Y \in \mathcal{X}(P)$ are horizontal vector fields. Then, we have*

$$\Omega_p(X_p, Y_p) = -\omega_p\big([X, Y](p)\big) . \tag{9.60}$$

In particular, the differentiable distribution $\mathcal{H}$ on P is integrable if and only if Ω vanishes identically on P.

Proof. Since X and Y are horizontal, Cartan's structure equation (9.57) implies

$$\Omega_p(X_p, Y_p) = (d\omega)_p\big(X_p, Y_p\big) .$$

Using (6.50) for the right-hand side, it follows

$$\Omega_p(X_p, Y_p) = -\omega_p\big([X, Y](p)\big) .$$

Recall that due to Frobenius' Theorem 8.5 the distribution $\mathcal{H}$ on P is integrable if and only if it is involutive in the sense of Definition 8.4. By (9.60) we deduce that the involutivity of $\mathcal{H}$ is equivalent to the fact of vanishing curvature on P. $\square$

In the following, we will often write Cartan's structure equation differently. More precisely, in a basis $\{E_i\}_{i=1,\dots,n}$ of $\mathfrak{g}$ with $\omega = \sum_{i=1}^n \omega^i E_i$, we have (see (7.21))

$$\begin{aligned}\omega \wedge \omega(X, Y) &= \sum_{i,j=1}^n \omega^i \wedge \omega^j (X, Y)\, [E_i, E_j] \\ &= \sum_{i,j=1}^n \big(\omega^i(X)\omega^j(Y) - \omega^j(X)\omega^i(Y)\big)\, [E_i, E_j] = 2\,\big[\omega(X), \omega(Y)\big] ,\end{aligned}$$

implying that Cartan's structure equation becomes

$$\Omega = d\omega + \frac{1}{2}\,\omega \wedge \omega . \tag{9.61}$$

Moreover, using the structure coefficients c_{ij}^k of $\mathfrak{g}$ defined in (7.22) and writing $\Omega = \sum_{i=1}^n \Omega^i E_i$, Cartan's structure equation can also be written as

$$\Omega^k = d\omega^k + \sum_{i<j=1}^n c_{ij}^k\, \omega^i \wedge \omega^j , \tag{9.62}$$

for $k = 1, \dots, n$.

Theorem 9.41 (Bianchi's Identity). *Let $\mathcal{H}$ be a connection on a principal G-bundle $\pi : P \longrightarrow M$ with curvature form $\Omega \in \Omega^2(P, \mathfrak{g})$. Then the covariant derivative of Ω with respect to $\mathcal{H}$ vanishes identically on P, i.e., we have*

$$D\Omega = 0 . \tag{9.63}$$

Proof. By Definition 9.35 of the covariant derivative D, it suffices to show that

$$d\Omega(X, Y, Z) = 0 ,$$

where $X, Y, Z \in \mathcal{X}(P)$ are all horizontal. For this purpose, we apply the exterior derivation d on Cartan's structure equation in the form (9.62) and obtain

$$d\Omega^k(X, Y, Z) = +\sum_{i<j=1}^n c_{ij}^k\, d\omega^i \wedge \omega^j (X, Y, Z) - \sum_{i<j=1}^n c_{ij}^k\, \omega^i \wedge d\omega^j (X, Y, Z) ,$$

Since $\omega^i(X) = \omega^i(Y) = \omega^i(Z) = 0$, for $i = 1, \dots, n$, the right-hand side of the last equation vanishes implying the result. $\square$

Curvature Form in Local Coordinates

Proceeding as in the case of a connection form, we associate to the curvature form $\Omega \in \Omega^2(P, \mathfrak{g})$ a family of $\mathfrak{g}$-valued two-forms defined on open subsets of the base space M. – Starting with the same setting as for a connection form in Section 9.4, we define

$$F_i = s_i^* \Omega \,. \tag{9.64}$$

Note that in the physics literature F_i is often called the (local) **field strength**. From Cartan's structure equation (9.61), we get

$$s_i^* \Omega = s_i^* \left(d\omega + \frac{1}{2}\, \omega \wedge \omega \right) = d(s_i^* \omega) + \frac{1}{2} (s_i^* \omega) \wedge (s_i^* \omega) \,,$$

and in terms of the gauge potential $A_i = s_i^* \omega$ defined (9.34) this becomes

$$F_i = dA_i + \frac{1}{2}\, A_i \wedge A_i \,. \tag{9.65}$$

There is also a compatibility condition for the field strength.

Proposition 9.42. *Let $\left\{ \left(\pi^{-1}(U_i), (\pi, \varphi_i) \right) \right\}_{i \in I}$ be an atlas and ω a connection form for a principal G-bundle $\pi : P \longrightarrow M$ with curvature form Ω. Then the two-forms $F_i \in \Omega^2(U_i, \mathfrak{g})$ satisfy the following compatibility condition on $U_i \cap U_j \neq \emptyset$:*

$$F_j(b) = Ad_{f_{j,i}(b)^{-1}} \circ F_i(b) \,. \tag{9.66}$$

Proof. Let $b \in U_i \cap U_j$ and $X_b, Y_b \in T_b M$. Then, we have

$$(F_j)_b \big(X_b, Y_b \big) = (s_j^* \Omega)_b \big(X_b, Y_b \big) = \Omega_{s_j(b)} \big((ds_j)_b \cdot X_b, (ds_j)_b \cdot Y_b \big) \,.$$

Now, we use the formula (9.39) for $(ds_j)_b \cdot X_b$ and $(ds_j)_b \cdot Y_b$. Note that the first term on the right-hand side of (9.39) is vertical because of (9.41). For the curvature being horizontal, we thus obtain

$$\begin{aligned} (F_j)_b \big(X_b, Y_b \big) &= \Omega_{s_j(b)} \big(dR_{f_{j,i}(b)} \cdot (ds_i)_b \cdot X_b, dR_{f_{j,i}(b)} \cdot (ds_i)_b \cdot Y_b \big) \\ &= \big((R_{f_{j,i}(b)})^* \Omega \big)_{s_i(b)} \big((ds_i)_b \cdot X_b, (ds_i)_b \cdot Y_b \big) \,. \end{aligned}$$

Since Ω is equivariant (see (9.56)), it then follows

$$\begin{aligned} (F_j)_b \big(X_b, Y_b \big) &= Ad_{f_{j,i}(b)^{-1}} \circ \Omega_{s_i(b)} \big((ds_i)_b \cdot X_b, (ds_i)_b \cdot Y_b \big) \\ &= Ad_{f_{j,i}(b)^{-1}} \circ (F_i)_b \big(X_b, Y_b \big) \,, \end{aligned}$$

showing the proposition. □

Remark. In Exercise 9.44 below, we show the compatibility condition when interpreting the curvature form as a section of $\bigwedge^2 T^* M \otimes P \times_{Ad} \mathfrak{g}$.

Flat Connections

Because of their integrability proved in Corollary 9.40, connections with vanishing curvature are very important.

Definition 9.43. *A connection ω on a principal G-bundle $\pi : P \longrightarrow M$ is said to be **flat** if its corresponding curvature form Ω vanishes identically.*

On trivial principal bundles such connections are easy to produce. – Consider the trivial principal G-bundle $\pi : M \times G \longrightarrow G$ with projection on the second factor and let θ be the Maurer-Cartan form on G. Then we define a $\mathfrak{g}$-valued one-form ω on P by

$$\omega = \pi^* \theta\,. \tag{9.67}$$

We first claim that ω is a connection form on $P = M \times G \longrightarrow G$. – Note that for $A \in T_eG$, we have

$$\begin{aligned}\big((R_g)^*\theta\big)_e(A) &= \theta_g\big((dR_g)_e \cdot A\big) = (dL_{g^{-1}})_g \cdot (dR_g)_e \cdot A \\ &= d(L_{g^{-1}} \circ R_g)_e \cdot A = Ad_{g^{-1}}(A) = Ad_{g^{-1}} \circ \theta_e(A)\,,\end{aligned}$$

showing that the Maurer-Cartan form θ is equivariant and also the equivariance of ω. Moreover, using (7.34) and Proposition 7.11, we have that

$$\begin{aligned}\omega_{(p,g)}\big(A^*(p,g)\big) &= (\pi^*\theta)_{(p,g)}\big(A^*(p,g)\big) = \theta_g\big(d\pi_{(p,g)} \cdot A^*(p,g)\big) \\ &= \theta_g\big(A^*(g)\big) = \theta_g\big(X_A(g)\big) = A\,.\end{aligned}$$

Proposition 9.23 then implies that ω is a connection form as claimed with connection (see (9.33))

$$\mathcal{H}_{(p,g)} = \ker \omega_{(p,g)} = \big\{X_{(p,g)} \in T_pM \oplus T_gG \ : \ \omega_{(p,g)}(X_{(p,g)}) = 0\big\}\,,$$

being as direct consequence of (9.67) the tangent space to the submanifold $M \times \{g\}$ of $M \times G$. In order to show that ω is flat, we use Maurer-Cartan's structure equation (7.23) for the computation of

$$d\omega = d(\pi^*\theta) = \pi^*(d\omega) = \pi^*\left(-\frac{1}{2}\,\theta \wedge \theta\right) = -\frac{1}{2}\,\omega \wedge \omega\,.$$

This gives

$$\Omega \overset{(9.61)}{=} d\omega + \frac{1}{2}\,\omega \wedge \omega = 0\,,$$

showing the flatness of ω.

* * *

Exercises.

Exercise 9.44. At the beginning of this section, we have seen that the curvature form can be regarded as section of $\bigwedge^2 T^*M \otimes P \times_{Ad} \mathfrak{g}$. Using the bundle atlas for $\pi : P \times_{Ad} \mathfrak{g} \longrightarrow M$ constructed in the proof of Theorem 9.16, define $F_i \in \Omega^2(U_i, \mathfrak{g})$ in a suitable way and show the compatibility condition (9.66) with the help of the corresponding transition functions.

Hint. Proceed similarly to the case of local gauge transformations studied in Section 9.4.

9.6 The Hopf Bundle

The *complex projective space* $\mathbb{C}P^n$ is defined similarly to $\mathbb{R}P^n$ (see Example 2.21). We will study the case $n = 1$ in more detail. – On $\mathbb{C}^2 \setminus \{0\}$, we consider the equivalence relation $z = (z_1, z_2) \sim \xi = (\xi_1, \xi_2)$ if and only if there exists $\lambda \in \mathbb{C}$ such that $z = \lambda\xi$. Let

$$\begin{aligned} \pi : \mathbb{C}^2 \setminus \{0\} &\longrightarrow \mathbb{C}P^1 \,, \\ z = (z_1, z_2) &\longmapsto [z] = [z_1, z_2], \,, \end{aligned} \tag{9.68}$$

denote the canonical projection which gives the quotient $\mathbb{C}P^1 = \mathbb{C}^2 \setminus \{0\}/\sim$ the quotient topology. Next, we define the open sets

$$U_1 = \big\{[z] \in \mathbb{C}P^1 \ : \ z_1 \neq 0\big\} \,, \qquad U_2 = \big\{[z] \in \mathbb{C}P^1 \ : \ z_2 \neq 0\big\} \,,$$

in $\mathbb{C}P^1$ with $U_1 \cup U_2 = \mathbb{C}P^1$ and the homeomorphisms

$$\begin{aligned} \varphi_1 : U_1 &\longrightarrow \mathbb{C} = \mathbb{R}^2 \,, & \qquad \varphi_2 : U_2 &\longrightarrow \mathbb{C} = \mathbb{R}^2 \,, \\ [z] &\longmapsto \frac{z_2}{z_1} \,, & [z] &\longmapsto \frac{z_1}{z_2} \,. \end{aligned}$$

Moreover, the transition function $\varphi_2 \circ \varphi_1^{-1} : \varphi_1(U_1 \cap U_2) \longrightarrow \varphi_2(U_1 \cap U_2)$ given by

$$\varphi_2 \circ \varphi_1^{-1}(z) = \varphi_2\big([1, z]\big) = \frac{1}{z} \,, \tag{9.69}$$

is a C^∞-diffeomorphism. The same holds for the other transition function $\varphi_1 \circ \varphi_2^{-1}$. In summary, the complexe projective space $\mathbb{C}P^1$ becomes a C^∞-differentiable manifold of (real) dimension two for this differentiable structure.

Now, we consider the two-dimensional unit sphere S^2 in $\mathbb{R}^3$ for which we construct a differentiable structure via the stereographic projection. More precisely, let $N = (0, 0, 1)$ and $S = (0, 0, -1)$ be the north and south pole of S^2 and define the open sets

$$U_S = S^2 \setminus \{N\} \,, \qquad U_N = S^2 \setminus \{S\} \,.$$

The stereographic projection maps are then given by

$$\begin{aligned} \varphi_S : U_S &\longrightarrow \mathbb{R}^2 \,, & \qquad \varphi_N : U_N &\longrightarrow \mathbb{R}^2 \,, \\ x &\longmapsto \left(\frac{x_1}{1 - x_3}, \frac{x_2}{1 - x_3}\right) \,, & x &\longmapsto \left(\frac{x_1}{1 + x_3}, \frac{x_2}{1 + x_3}\right) \,. \end{aligned} \tag{9.70}$$

In words, the values of φ_S and φ_N correspond to the intersection with the (x_1, x_2)-plane of the straight line joining x and the north pole N, respectively the south pole S. Moreover, the transition functions from $\mathbb{R}^2 \setminus \{0, 0\}$ into $\mathbb{R}^2 \setminus \{0, 0\}$ read as

$$\varphi_S \circ \varphi_N^{-1}(x) = \frac{1}{\|x\|^2}\, x = \varphi_N \circ \varphi_S^{-1}(x)\,. \tag{9.71}$$

Identifying $\mathbb{R}^2$ with $\mathbb{C}$ this becomes

$$\varphi_S \circ \varphi_N^{-1}(z) = \frac{1}{\bar{z}} = \varphi_N \circ \varphi_S^{-1}(z)\,.$$

Note that these transition functions coincide up to complex conjugation with (9.69).

By composition of the transition functions (9.71) and (9.69) for S^2, respectively $\mathbb{C}P^1$, and a gluing argument, it is now possible to construct a diffeomorphism

$$\begin{aligned} f : \mathbb{C}P^1 &\longrightarrow S^2\,,\\ [z] &\longmapsto \frac{1}{|z_1|^2+|z_2|^2}\big(2\, z_1\bar{z}_2, |z_1|^2-|z_2|^2\big)\,, \end{aligned} \tag{9.72}$$

showing that $\mathbb{C}P^1$ and S^2 are diffeomorphic. – For more details, we refer to G.L. Naber [], Section 1.2.

Next, we consider $S^3 \subset \mathbb{R}^4$ viewed as

$$S^3 = \big\{(z_1,z_2) \in \mathbb{C}^2 \ :\ |z_1|^2+|z_2|^2 = 1\big\}\,,$$

and the map

$$\begin{aligned} S^3 \times U(1) &\longrightarrow S^3\,,\\ \big((z_1,z_2),e^{i\theta}\big) &\longmapsto (z_1,z_2)e^{i\theta} = (z_1\, e^{i\theta}, z_2\, e^{i\theta})\,. \end{aligned}$$

It is left as an exercise to check that this map is well-defined and gives a free right action of $U(1) = \{z \in \mathbb{C} \ :\ |z| = 1\} = \{e^{i\theta} \in \mathbb{C} \ :\ \theta \in \mathbb{R}\}$ on S^3. Now, consider the restriction of the projection $\pi : \mathbb{C}^2 \setminus \{0\} \longrightarrow \mathbb{C}P^1$ defined in (9.68) to S^3 – which we will also denote by π – and let $(z_1,z_2) \in S^3$ and $(\xi_1,\xi_2) \in S^3$ such that $\pi(z_1,z_2) = \pi(\xi_1,\xi_2)$. This implies that there exists $\lambda \in U(1)$ such that $(z_1,z_2) = \lambda\,(\xi_1,\xi_2)$. In other words, for $(z_1,z_2) \in S^3$ and all $e^{i\theta} \in U(1)$, we have

$$\pi\big((z_1,z_2)e^{i\theta}\big) = \pi(z_1,z_2)\,. \tag{9.73}$$

Note that the restriction of $\pi : \mathbb{C}^2 \setminus \{0\} \longrightarrow \mathbb{C}P^1$ to S^3 can also be written differently in the following way: Identyfing $\mathbb{C}^2$ with $\mathbb{R}^4$ and using the fact that $\mathbb{C}P^1$ and S^2 are diffeomorphic, we obtain

$$\begin{aligned} \pi : S^3 &\longrightarrow S^2\,,\\ (x_1,x_2,x_3,x_4) &\longmapsto \big(2\, x_1x_3 + 2\, x_2x_4, 2\, x_2x_3 - 2\, x_1x_4,\\ &\qquad (x_1)^2+(x_2)^2-(x_3)^2-(x_4)^2\big)\,. \end{aligned} \tag{9.74}$$

Now, we want to show that $\pi : S^3 \longrightarrow \mathbb{C}P^1$ is a principal $U(1)$-bundle, called the **Hopf bundle**. For this purpose, we have to construct a bundle atlas. – Let U_1 and U_2 be open subsets of $\mathbb{C}P^1$ as defined before. Then, for $i = 1, 2$, we define the following diffeomorphisms:

$$\begin{aligned}(\pi, \varphi_i) : \pi^{-1}(U_i) &\longrightarrow U_i \times U(1)\,, \\ (z_1, z_2) &\longmapsto \left([z_1, z_2], \frac{z_i}{|z_i|} \right)\,. \end{aligned} \tag{9.75}$$

Moreover, for all $e^{i\theta} \in U(1)$, we have

$$\varphi_i\big((z_1, z_2)e^{i\theta}\big) = \varphi_i\big((z_1\, e^{i\theta}, z_2\, e^{i\theta})\big) = \frac{z_i}{|z_i|}\, e^{i\theta} = \varphi_i\big((z_1, z_2)\big)e^{i\theta}\,.$$

Togther with (9.73), we have thus shown that the bundle atlas

$$\big\{\big(\pi^{-1}(U_1), (\pi, \varphi_1)\big), \big(\pi^{-1}(U_2), (\pi, \varphi_2)\big)\big\}$$

satisfies (9.5) implying that $\pi : S^3 \longrightarrow \mathbb{C}P^1$ is a principal $U(1)$-bundle. Equivalently, we have constructed a principal $U(1)$-bundle $\pi : S^3 \longrightarrow S^2$. For completeness, we note that the transition functions $f_{1,2} : U_1 \cap U_2 \longrightarrow U(1)$ and $f_{2,1} : U_2 \cap U_1 \longrightarrow U(1)$ are given by (see (9.6))

$$f_{1,2}\big((z_1, z_2)\big) = \varphi_2(z_1, z_2)\varphi_1(z_1, z_2)^{-1} = \frac{|z_1|\, z_2}{|z_2|\, z_1}\,,$$

respectively, by

$$f_{2,1}\big((z_1, z_2)\big) = \varphi_1(z_1, z_2)\varphi_2(z_1, z_2)^{-1} = \frac{|z_2|\, z_1}{|z_1|\, z_2}\,.$$

Moreover, the (local) sections associated to the above bundle charts read as

$$s_1\big([z_1, z_2]\big) = (\pi, \varphi_1)^{-1}\big([z_1, z_2], 1\big) = \left(|z_1|, \frac{z_2}{z_1}|z_1| \right) \in S^3\,, \tag{9.76}$$

respectively as

$$s_2\big([z_1, z_2]\big) = (\pi, \varphi_2)^{-1}\big([z_1, z_2], 1\big) = \left(\frac{z_1}{z_2}|z_2|, |z_2| \right) \in S^3\,. \tag{9.77}$$

On $\mathbb{R}^4$, there is an addition to the vector space structure an associative and distributive multiplication generalizing that of complex numbers. More precisely, we write elements in $\mathbb{R}^4$ as $x_1 + i\, x_2 + j\, x_3 + k\, x_4$ and define

$$\begin{gathered} i^2 = j^2 = k^2 = -1\,, \\ ij = -ji = k\,, \quad jk = -kj = i\,, \quad ki = -ik = j\,. \end{gathered}$$

We end up with the so-called **quaternionic numbers** $\mathbb{H}$. The imaginary part of $q \in \mathbb{H}$ is given by $\mathrm{Im}(q) = i\,x_2 + j\,x_3 + k\,x_4$, the real part by $\mathrm{Re}(q) = x_1$ and the quaternionic conjugate by $\bar{q} = x_1 - i\,x_2 - j\,x_3 - k\,x_4$. Moreover, the norm $|q|^2 = q\bar{q}$ equals the usual norm on $\mathbb{R}^4$.

As in the case of complex numbers, we can define the *quaternionic projective space* $\mathbb{H}P^1$ as the quotient $\mathbb{H}^2 \setminus \{0\}/\sim$ with canonical projection

$$\begin{aligned} \pi : \mathbb{H}^2 \setminus \{0\} &\longrightarrow \mathbb{H}P^1\,, \\ (q_1, q_2) &\longmapsto [q_1, q_2]\,. \end{aligned} \tag{9.78}$$

We also endow $\mathbb{H}^2$ with the standard scalar product

$$\big\langle (q_1, q_2), (\tilde{q}_1, \tilde{q}_2) \big\rangle = q_1 \tilde{q}_1 + q_2 \tilde{q}_2\,. \tag{9.79}$$

Note that the real part of the last expression just gives the usual standard scalar product on $\mathbb{R}^8 = \mathbb{H}^2$.

Now, we proceed as in the case of the Hopf bundle: Consider the restriction of the canonical projection $\pi : \mathbb{H}^2 \setminus \{0\} \longrightarrow \mathbb{H}P^1$ to $S^7 \subset \mathbb{R}^8$ viewed as

$$S^7 = \big\{ (q_1, q_2) \in \mathbb{H}^2 \ : \ |q_1|^2 + |q_2|^2 = 1 \big\}\,,$$

and define a free a right action

$$\begin{aligned} S^7 \times SU(2) &\longrightarrow S^7\,, \\ \big((q_1, q_2), g\big) &\longmapsto (q_1, q_2)g = (q_1 g, q_2 g) \end{aligned}$$

of $SU(2) = \{q \in \mathbb{H} \ : \ |q| = 1\}$ on S^7, in order to obtain a principal $SU(2)$-bundle $\pi : S^7 \longrightarrow \mathbb{H}P^1$. Equivalently, since $\mathbb{H}P^1$ is isomorphic to S^4, we have obtained a principal $SU(2)$-bundle $\pi : S^7 \longrightarrow S^4$, the so-called (**generalized**) **Hopf bundle**.

Connection and Curvature for the Hopf Bundle

We now construct a connection on the generalized Hopf bundle by introducing first a Lie algebra-valued one-form on S^7 having the characteristic properites of a connection form. Then we will see that the connection corresponding to this one-form has a precise geometrical interpretation.

Consider the following one-form $\tilde{\omega}$ on $\mathbb{H}^2$:

$$\tilde{\omega}_{(q_1, q_2)} = \mathrm{Im}(\bar{q}_1\, dq_1 + \bar{q}_2\, dq_2)\,.$$

Then, let ω be the restriction of $\tilde{\omega}$ to S^7, i.e.,

$$\omega = \iota_{S^7}^* \tilde{\omega}\,. \tag{9.80}$$

This one-form is clearly $\mathrm{Im}\,\mathbb{H}$-valued. Because of Exercise 9.46, the one-form ω on S^7 can also be seen as taking its values in the Lie algebra $\mathfrak{su}(2)$ of

$SU(2)$. For $p = (q_1, q_2) \in S^7 \subset \mathbb{H}^2$ and $X_p = (X_p^1, X_p^2) \in T_pS^7 \subset T_p\mathbb{H}^2 = T_{q_1}\mathbb{H} \times T_{q_2}\mathbb{H}$, we have

$$\begin{aligned} \omega_p(X_p) &= \tilde{\omega}_p(d\iota_{S^7} \cdot X_p) = \omega_{(q_1,q_2)}(\tilde{X}_{(q_1,q_2)}) \\ &= \mathrm{Im}\big(\bar{q}_1\, dq_1(\tilde{X}_{(q_1,q_2)}) + \bar{q}_2\, dq_2(\tilde{X}_{(q_1,q_2)})\big) \\ &= \mathrm{Im}\big(\bar{q}_1\, X^1_{(q_1,q_2)} + \bar{q}_2\, X^2_{(q_1,q_2)}\big)\,, \end{aligned} \tag{9.81}$$

where, as usual, tangent vectors to $\mathbb{H}$ are identified with elements of $\mathbb{H}$ via the canonical isomorphism and $d\iota_{S^7} \cdot X_p = \tilde{X}_p \in T_p\mathbb{H}^2$.

Next, we claim that

$$(R_g)^*\omega = Ad_{g^{-1}} \circ \omega\,, \tag{9.82}$$

for all $g \in SU(2)$. In order to show the claim, we compute for $X_p \in T_pS^7$

$$\begin{aligned} \big((R_g)^*\omega\big)_p(X_p) &= \omega_{pg}(dR_g \cdot X_p) = \omega_{pg}(X_p^1 g, X_p^2 g) \\ &\overset{(9.81)}{=} \mathrm{Im}\big(\overline{q_1 g}\, X_p^1 g + \overline{q_2 g}\, X_p^2 g\big) \\ &= \mathrm{Im}\big(\bar{g}\bar{q}_1 X_p^1 g + \bar{g}\bar{q}_2 X_p^2 g\big) \\ &= \mathrm{Im}\big(g^{-1}\bar{q}_1\, X_p^1 g + g^{-1}\bar{q}_2\, X_p^2 g\big) \\ &= g^{-1}\mathrm{Im}\big(\bar{q}_1\, X_p^1 + \bar{q}_2\, X_p^2\big)g = Ad_{g^{-1}}\big(\omega_p(X_p)\big)\,, \end{aligned}$$

where we used that $\bar{g} = g^{-1}$, for $g \in SU(2)$.

For all $p \in S^7$, the $\mathrm{Im}\,\mathbb{H}$-valued one-form ω also satisfies

$$\omega_p\big(A^*(p)\big) = A\,, \tag{9.83}$$

where $A^* \in \mathcal{X}(S^7)$ is the fundamental vector field generated by $A \in \mathfrak{su}(2) = \mathrm{Im}\,\mathbb{H}$. Indeed, we have

$$\begin{aligned} \omega_p\big(A^*(p)\big) &= \omega_p\big(q_1 A, q_2 A\big) = \mathrm{Im}\big(\bar{q}_1\, q_1 A + \bar{q}_2\, q_2 A\big) \\ &= \mathrm{Im}\big((|q_1|^2 + |q_2|^2)A\big) = \mathrm{Im}(A) = A\,, \end{aligned}$$

since $p \in S^7$ and $A \in \mathrm{Im}\,\mathbb{H}$. – Note that the previous calculations are carried out in more details in G.L. Naber [], Section 4.8.

Because of (9.82) and (9.83) we can now apply Proposition 9.23 to conclude that ω defines a connection form for the principal $SU(2)$-bundle $\pi : S^7 \longrightarrow S^4$. Moreover, a connection $\mathcal{H}$ on $\pi : S^7 \longrightarrow S^4$ is then given by

$$\mathcal{H}_p = \ker\omega_p = \big\{X_p \in T_pS^7 \;:\; \mathrm{Im}\big(\bar{q}_1 X_p^1 + \bar{q}_2 X_p^2\big) = 0\big\}\,. \tag{9.84}$$

This subspace of T_pS^7 has a simple geometrical interpretation. – Since the vertical subspace $\mathcal{V}_p$ of T_pS^7 contains vectors being tangent to the fiber $\big\{(q_1 g, q_2 g) \;:\; g \in SU(2)\big\}$ through p, we deduce

$$\mathcal{V}_p = \Big\{ X_p \in T_p S^7 \ : \ X_p = \frac{d}{dt}\Big|_{t=0} (q_1, q_2)\, g(t) \Big\}$$
$$= \big\{ (q_1\, A, q_2\, A) \ : \ A \in \operatorname{Im} \mathbb{H} \big\} \,,$$

where $g(t) \in C^1\big([-\delta, \delta], SU(2)\big)$ is a curve in $SU(2)$ with $g(0) = e$. The orthogonal complement of $\mathcal{V}_p$ in $\mathbb{H}^2$ for the standard scalar product in $\mathbb{R}^8$ consists of elements $(\tilde{q}_1, \tilde{q}_2) \in \mathbb{H}^2$ satisfying (see (9.79))

$$\operatorname{Re}\Big(\big\langle (q_1 A, q_2 A), (\tilde{q}_1, \tilde{q}_2) \big\rangle \Big) = \operatorname{Re}\big(\overline{q_1 A}\, \tilde{q}_1 + \overline{q_2 A}\, \tilde{q}_2 \big) = 0 \,,$$

for all $A \in \operatorname{Im} \mathbb{H}$. It is not difficult to check that the last equation is equivalent to

$$\operatorname{Im}\big(\bar{q}_1 \tilde{q}_1 + \bar{q}_2 \tilde{q}_2 \big) = 0 \,.$$

Comparing with (9.84) we then deduce that $\mathcal{H}_p$ is just the subset of the real orthogonal complement of $\mathcal{V}_p$ which lies in $T_p S^7$. This is why $\mathcal{H}$ is often called the *natural* connection on the generalized Hopf bundle (see G.L. Naber [], Section 5.1).

In an analogous manner, we construct the natural connection on the Hopf bundle $\pi : S^3 \longrightarrow S^2$ with connection form given by

$$\begin{aligned} \omega &= \iota^*_{S^3} \big(\operatorname{Im}(\bar{z}_1\, dz_1 + \bar{z}_2\, dz_2) \big) \\ &= \iota^*_{S^3} \big(x_1\, dx_2 - x_2\, dx_1 + x_3\, dx_4 - x_4\, dx_3 \big) \,, \end{aligned} \tag{9.85}$$

where $z_1 = x_1 + i\, x_2$ and $z_2 = x_3 + i\, x_4$. Since the structure group $U(1)$ of the Hopf bundle is abelian, we conclude using Cartan's structure equation (9.57) that the curvature form is simply

$$\begin{aligned} \Omega = d\omega &= \iota^*_{S^3} \big(\operatorname{Im}(d\bar{z}_1 \wedge dz_1 + d\bar{z}_2 \wedge dz_2) \big) \\ &= 2\, \iota^*_{S^3} \big(dx_1 \wedge dx_2 + dx_3 \wedge dx_4 \big) \,. \end{aligned} \tag{9.86}$$

In a next step, we express the natural connection form of the generalized Hopf bundle in local coordinates. – Let $s_1 : U_1 \subset S^4 \longrightarrow S^7$ be the (local) section associated to a bundle chart of $\pi : S^7 \longrightarrow S^4$ and $\varphi_1 : U_1 \longrightarrow \mathbb{H} = \mathbb{R}^4$ the usual coordinate function for S^4 being isomorphic to $\mathbb{H}P^1$. (At the beginning of this section explicit expressions are given in the complex case). In order to make sure that the map $s_1 \circ \varphi_1^{-1} : \varphi_1(U_1) \subset \mathbb{R}^4 \longrightarrow S^7$ takes it values in S^7, the following representation for φ_1^{-1} will be used:

$$\varphi_1^{-1}(q) = \left[\frac{1}{\sqrt{1 + |q|^2}}, \frac{q}{\sqrt{1 + |q|^2}} \right] . \tag{9.87}$$

For $q \in \mathbb{H} \setminus \{0\}$, we then obtain

$$\big(s_1 \circ \varphi_1^{-1} \big)(q) = \frac{1}{\sqrt{1 + |q|^2}}\, (1, q) \in S^7 \subset \mathbb{H}^2 \,.$$

It is left as an exercise (see G.L. Naber [], Equation 4.8.14) to show that the local coordinate expression for the connection form ω is given by

$$\left((s_1 \circ \varphi_1^{-1})^* \omega\right)_q = \mathrm{Im}\left(\frac{\bar{q}}{1+|q|^2}\, dq\right) . \tag{9.88}$$

Moreover, as a consequence of Cartan's structure equation (9.57), we obtain for the local coordinate expression of the curvature form (see G.L. Naber [], Equation 4.10.13)

$$\left((s_1 \circ \varphi_1^{-1})^* \Omega\right)_q = \mathrm{Im}\left(\frac{1}{(1+|q|^2)^2}\, dq \wedge d\bar{q}\right) . \tag{9.89}$$

* * *

Exercise 9.45. Show that the Lie algebra $\mathfrak{u}(1)$ of $U(1)$ is isomorphic to the pure imaginary complex numbers $\mathrm{Im}\,\mathbb{C}$ with trivial bracket.

Exercise 9.46. Show that the Lie algebra $\mathfrak{su}(2)$ of $SU(2)$ is isomorphic to the pure imaginary quaternions $\mathrm{Im}\,\mathbb{H}$ with bracket $[q_1, q_2] = q_1 q_2 - q_2 q_1$.

9.7 Grassmannian Manifolds and Stiefel Bundles

We give a straightforward generalization of the projective spaces encountered in previous chapters. – Let E denote an n-dimensional vector space over the ground field $\mathcal{F}$ and $k \leq n$. The collection of k-dimensional subspaces or k-planes $G_k(E)$ of E can be made to a manifold, the so-called **Grassmannian manifold**. Note that $G_1(\mathbb{R}^n) = \mathbb{R}P^{n-1}$, $G_1(\mathbb{C}^n) = \mathbb{C}P^{n-1}$ and $G_1(\mathbb{H}^n) = \mathbb{H}P^{n-1}$.

Next, we construct a canonical vector bundle over the Grassmannian manifold being a generalization of Example 9.92. – The total space E_k of this vector bundle is given by

$$E_k = \left\{(P, v) \in G_k(E) \times E \;:\; v \in P\right\} , \tag{9.90}$$

with bundle projection

$$\begin{aligned} \pi_k : E_k &\longrightarrow G_k(E)\,, \\ (P, v) &\longmapsto P\,. \end{aligned} \tag{9.91}$$

Thus the fiber over $P \in G_k(E)$ is P itself. Details are left to the reader.

There is also a principal bundle over the Grassmannian manifold. – Considering the complex case, we define the *k-frames* V in $\mathbb{C}^n$ as the set of k orthonormal vectors $(v_1, \ldots, v_k)$ satisfying

$$\langle v_i, v_j \rangle = \delta_{ij} \,,$$

for $i, j = 1, \ldots, k$, where $\langle \cdot, \cdot \rangle$ denotes the standard scalar product in $\mathbb{C}^n$. The collection of all k-frames can be made to a manifold, which we will denote by $V_k(\mathbb{C}^n)$. The right action of $g \in U(k)$ on $V \in V_k(\mathbb{C}^n)$ is denoted by Vg. Then, we define the following bundle projection:

$$\begin{aligned} \pi : V_k(\mathbb{C}^n) &\longrightarrow G_k(\mathbb{C}^n)\,, \\ V &\longmapsto P_V\,, \end{aligned} \tag{9.92}$$

where $P_V \in G_k(\mathbb{C}^n)$ is the k-plane in $\mathbb{C}^n$ passing through the origin generated by the k-frame V. For $V, \tilde{V} \in V_k(\mathbb{C}^n)$ such that $\pi(V) = \tilde{\pi}(V)$, there exists $g \in U(k)$ with $V = \tilde{V}g$. We deduce that the quotient of $V_k(\mathbb{C}^n)$ by this $U(k)$-action is just $G_k(\mathbb{C}^n)$. Thus, we arrive at the principal $U(k)$-bundle $\pi : V_k(\mathbb{C}^n) \longrightarrow G_k(\mathbb{C}^n)$, the so-called **Stiefel bundle**.

Remark. It is important to note that in the case of $n = 2$ and $k = 1$ the Stiefel bundle $\pi : V_1(\mathbb{C}^2) \longrightarrow G_1(\mathbb{C}^2) = \mathbb{C}P^1$ is exactly the Hopf bundle $\pi : S^3 \longrightarrow S^2$ of Section 9.92. Moreover, note also that the Stiefel bundle $\pi : V_1(\mathbb{H}^2) \longrightarrow G_1(\mathbb{H}^2) = \mathbb{H}P^1$ gives the (generalized) Hopf bundle $\pi : S^7 \longrightarrow S^4$.