NOTES ON MANIFOLDS

ALBERTO S. CATTANEO

Contents

1. Introduction 2
2. Manifolds 2
2.1. Coordinates 6
2.2. Dimension 7
2.3. The implicit function theorem 7
3. Maps 8
3.1. The pullback 10
3.2. Submanifolds 11
4. Topological manifolds 12
4.1. Manifolds by local data 14
5. Bump functions and partitions of unity 15
6. Differentiable manifolds 19
6.1. The tangent space 20
6.2. The tangent bundle 25
6.3. Vector fields 26
6.4. Integral curves 28
6.5. Flows 30
7. Derivations 32
7.1. Vector fields as derivations 37
7.2. The Lie bracket 39
7.3. The push-forward of derivations 40
7.4. The Lie derivative 43
7.5. Plane distributions 47
8. Vector bundles 53
8.1. General definitions 53
8.2. Densities and integration 62
8.3. The cotangent bundle and 1-forms 77
8.4. The tensor bundle 81
8.5. Digression: Riemannian metrics 84
9. Differential forms, integration and Stokes theorem 91
9.1. Differential forms 92

Date: April 24, 2018.

9.2. The de Rham differential 94
9.3. Graded linear algebra and the Cartan calculus 99
9.4. Orientation and the integration of differential forms 107
9.5. Manifolds with boundary and Stokes theorem 121
9.6. Singular homology 131
9.7. The nonorientable case 135
9.8. Digression: Symplectic manifolds 138
10. Lie groups 145
10.1. The Lie algebra of a Lie group 146
10.2. The exponential map 150
10.3. Morphisms 153
10.4. Actions of Lie groups 155
10.5. Left invariant forms 157
Appendix A. Topology 158
Appendix B. Multilinear algebra 161
B.1. Tensor powers 166
B.2. Exterior algebra 169
References 176
Index 177

1. Introduction

Differentiable manifolds are sets that locally look like some $\mathbb{R}^n$ so that we can do calculus on them. Examples of manifolds are open subsets of $\mathbb{R}^n$ or subsets defined by constraints satisfying the assumptions of the implicit function theorem (example: the n-sphere S^n). Also in the latter case, it is however more practical to think of manifolds intrinsically in terms of charts.

The example to bear in mind are charts of Earth collected in an atlas, with the indications on how to pass from one chart to another. Another example that may be familiar is that of regular surfaces.

2. Manifolds

Definition 2.1. A chart on a set M is a pair (U, ϕ) where U is a subset of M and ϕ is an injective map from U to $\mathbb{R}^n$ for some n.

The map ϕ is called a chart map or a coordinate map. One often refers to ϕ itself as a chart, for the subset U is part of ϕ as its definition domain.

If (U, ϕ_U) and (V, ϕ_V) are charts on M, we may compose the bijections $(\phi_U)_{|U\cap V}\colon U\cap V \to \phi_U(U\cap V)$ and $(\phi_V)_{|U\cap V}\colon U\cap V \to \phi_V(U\cap V)$

and get the bijection

$$\phi_{U,V} := (\phi_V)_{|U\cap V} \circ (\phi_U|_{U\cap V})^{-1} \colon \phi_U(U\cap V) \to \phi_V(U\cap V)$$

called the transition map from (U, ϕ_U) to (V, ϕ_V) (or simply from U to V).

Definition 2.2. An atlas on a set M is a collection of charts $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$, where I is an index set, such that $\cup_{\alpha\in I} U_\alpha = M$.

Remark 2.3. We usually denote the transition maps between charts in an atlas $(U_\alpha, \phi_\alpha)_{\alpha\in I}$ simply by $\phi_{\alpha\beta}$ (instead of ϕ_{U_α,U_β}).

One can easily check that, if $\phi_\alpha(U_\alpha)$ is open $\forall\alpha \in I$ (in the standard topology of the target), then the atlas $\mathcal{A} = \{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ defines a topology[1] on M:

$$\mathcal{O}_{\mathcal{A}}(M) := \{V \subset M \mid \phi_\alpha(V \cap U_\alpha) \text{ is open } \forall\alpha \in I\}.$$

We may additionally require that all U_α be open in this topology or, equivalently, that $\phi_\alpha(U_\alpha \cap U_\beta)$ is open $\forall\alpha, \beta \in I$. In this case we speak of an open atlas. All transition maps in an open atlas have open domain and codomain, so we can require them to belong to a class $\mathcal{C} \subset \mathcal{C}^0$ of maps (e.g., $\mathcal{C}^k$ for $k = 0, 1, \ldots, \infty$, or analytic, or complex analytic, or Lipschitz).

Definition 2.4. A $\mathcal{C}$-atlas is an open[2] atlas such that all transition maps are $\mathcal{C}$-maps.

Notice that, by definition, a $\mathcal{C}$-atlas is also in particular a $\mathcal{C}^0$-atlas.

Example 2.5. Let $M = \mathbb{R}^n$. Then $\mathcal{A} = \{(\mathbb{R}^n, \phi)\}$ is a $\mathcal{C}$-atlas for any structure $\mathcal{C}$ if ϕ is an injective map with open image. Notice that M has the standard topology iff ϕ is a homeomorphism with its image. If ϕ is the identity map Id, this is called the standard atlas for $\mathbb{R}^n$.

Example 2.6. Let M be an open subset of $\mathbb{R}^n$ with its standard topology. Then $\mathcal{A} = \{(U, \iota)\}$, with ι the inclusion map, is a $\mathcal{C}$-atlas for any structure $\mathcal{C}$.

Example 2.7. Let $M = \mathbb{R}^n$. Let $\mathcal{A} = \{(\mathbb{R}^n, \mathrm{Id}), (\mathbb{R}^n, \phi)\}$. Then $\mathcal{A}$ is a $\mathcal{C}$-atlas iff ϕ and its inverse are $\mathcal{C}$-maps.

Example 2.8. Let M be the set of lines (i.e., one-dimensional affine subspaces) of $\mathbb{R}^2$. Let U_1 be the subset of nonvertical lines and U_2 the subset of nonhorizontal lines. Notice that every line in U_1 can be

[1]For more on topology, see Appendix A.

[2]Notice that to define a $\mathcal{C}^0$-atlas we would not need the condition that the atlas be open, but we will need this condition for the proof of several important properties.

uniquely parametrized as $y = m_1x + q_1$ and every line in U_2 can be uniquely parametrized as $x = m_2y + q_2$. Define $\phi_i\colon U_i \to \mathbb{R}^2$ as the map that assigns to a line the corresponding pair (m_i, q_i), for $i = 1, 2$. Then $\mathcal{A} = \{(U_1, \phi_1), (U_2, \phi_2)\}$ is a $\mathcal{C}^k$-atlas for $k = 0, 1, 2, \dots, \infty$.

Example 2.9. Define S^1 in terms of the angle that parametrizes it (i.e., by setting $x = \cos\theta$, $y = \sin\theta$). The angle θ is defined modulo 2π. The usual choice of thinking of S^1 as the closed interval $[0, 2\pi]$ with 0 and 2π identified does not give an atlas. Instead, we think of S^1 as the quotient of $\mathbb{R}$ by the equivalence relation $\theta \sim \widetilde{\theta}$ if $\theta - \widetilde{\theta} = 2\pi k$, $k \in \mathbb{Z}$. We then define charts by taking open subsets of $\mathbb{R}$ and using shifts by multiple of 2π as transition functions. A concrete choice is the following. Let E denote the class of 0 (equivalently, for S^1 in $\mathbb{R}^2$, E is the eastward point $(1, 0)$). We set $U_E = S^1 \setminus \{E\}$ and denote by $\phi_E\colon U_E \to \mathbb{R}$ the map that assigns the angle in $(0, 2\pi)$. Analogously, we let W denote the westward point $(-1, 0)$ (i.e., the equivalence class of π) and set $U_W = S^1 \setminus \{W\}$. We denote by $\phi_W\colon U_W \to \mathbb{R}$ the map that assigns the angle in $(-\pi, \pi)$. We have $S^1 = U_E \cup U_W$, $\phi_E(U_E \cap U_W) = (0, \pi) \cup (\pi, 2\pi)$ and $\phi_W(U_E \cap U_W) = (-\pi, 0) \cup (0, \pi)$. Finally, we have

$$\phi_{EW}(\theta) = \begin{cases} \theta & \text{if } \theta \in (0, \pi), \\ \theta - 2\pi & \text{if } \theta \in (\pi, 2\pi). \end{cases}$$

Hence $\mathcal{A} = \{(U_E, \phi_E), (U_W, \phi_W)\}$ is a $\mathcal{C}^k$-atlas for $k = 0, 1, 2, \dots, \infty$.

Example 2.10 (Regular surfaces)**.** Recall that a regular surface is a subset S of $\mathbb{R}^3$ such that for every $p \in S$ there is an open subset U of $\mathbb{R}^2$ and a map $\mathbf{x}\colon U \to \mathbb{R}^3$ with $p \in \mathbf{x}(U) \subset S$ satisfying the following properties:

(1) $\mathbf{x}\colon U \to \mathbf{x}(U)$ is a homeomorphism (i.e., $\mathbf{x}$ is injective, continuous and open),
(2) $\mathbf{x}$ is $\mathcal{C}^\infty$, and
(3) the differential $\mathrm{d}_u\mathbf{x}\colon \mathbb{R}^2 \to \mathbb{R}^3$ is injective for all $u \in U$.

A map $\mathbf{x}$ satisfying these properties is called a regular parametrization.[3] The first property allows one to define a chart $(\mathbf{x}(U), \mathbf{x}^{-1})$ and all charts arising this way form an open atlas. The second and third properties make this into a $\mathcal{C}^\infty$-atlas, so a regular surface is an example of $\mathcal{C}^\infty$-manifold.

Example 2.11. Let $M = S^n := \{x \in \mathbb{R}^{n+1} \mid \sum_{i=1}^{n+1}(x^i)^2 = 1\}$ be the n-sphere. Let $N = (0, \dots, 0, 1)$ and $S = (0, \dots, 0, -1)$ denote its north and south poles, respectively. Let $U_N := S^n \setminus \{N\}$ and $U_S := S^n \setminus \{S\}$.

[3] In the terminology of Definition 6.2, $\mathbf{x}$ is an embedding of U into $\mathbb{R}^3$.

Let $\phi_N\colon U_N \to \mathbb{R}^n$ and $\phi_S\colon U_S \to \mathbb{R}^n$ be the stereographic projections with respect to N and S, respectively: ϕ_N maps a point y in S^n to the intersection of the plane $\{x^{n+1} = 0\}$ with the line passing through N and y; similarly for ϕ_S. A computation shows that $\phi_{SN}(\mathbf{x}) = \phi_{NS}(\mathbf{x}) = \frac{\mathbf{x}}{||\mathbf{x}||^2}$, $\mathbf{x} \in \mathbb{R}^n \setminus \{0\}$. Then $\mathcal{A} = \{(U_N, \phi_N), (U_S, \phi_S)\}$ is a $\mathcal{C}^k$-atlas for $k = 0, 1, 2, \dots, \infty$.

Example 2.12 (Constraints)**.** Let M be a subset of $\mathbb{R}^n$ defined by $\mathcal{C}^k$-constraints satisfying the assumptions of the implicit function theorem. Then locally M can be regarded as the graph of a $\mathcal{C}^k$-map. Any open cover of M with this property yields a $\mathcal{C}^k$-atlas. We will give more details on this in subsection 2.3.

As we have seen in the examples above, the same set may occur with different atlases. The main point, however, is to consider different atlases just as different description of the same object, at least as long as the atlases are compatible. By this we mean that we can decide to consider a chart from either atlas. This leads to the following

Definition 2.13. Two $\mathcal{C}$-atlases on the same set are $\mathcal{C}$-equivalent if their union is also a $\mathcal{C}$-atlas.

Notice that the union of two atlases has in general more transition maps and in checking equivalence one has to check that also the new transition maps are $\mathcal{C}$-maps. In particular, this first requires checking that the union of the two atlases is open.

Example 2.14. Let $M = \mathbb{R}^n$, $\mathcal{A}_1 = \{(\mathbb{R}^n, \mathrm{Id})\}$ and $\mathcal{A}_2 = \{(\mathbb{R}^n, \phi)\}$ for an injective map ϕ with open image. These two atlases are $\mathcal{C}$-equivalent iff ϕ and its inverse are $\mathcal{C}$-maps.

We finally arrive at the

Definition 2.15. A $\mathcal{C}$-manifold is an equivalence class of $\mathcal{C}$-atlases.

Remark 2.16. Usually in defining a $\mathcal{C}$-manifold we explicitly introduce one atlas and tacitly consider the corresponding $\mathcal{C}$-manifold as the equivalence class containing this atlas. Also notice that the union of all atlases in a given class is also an atlas, called the maximal atlas, in the same equivalence class. Thus, we may equivalently define a manifold as a set with a maximal atlas. This is not very practical as the maximal atlas is huge.

Working with an equivalence class of atlases instead of a single one also has the advantage that whatever definition we want to give requires choosing just a particular atlas in the class and we may choose the most convenient one.

Example 2.17. The standard $\mathcal{C}$-manifold structure on $\mathbb{R}^n$ is the $\mathcal{C}$-equivalence class of the atlas $\{(\mathbb{R}^n, \mathrm{Id})\}$.

Remark 2.18. Notice that the same set can be given different manifold structures. For example, let $M = \mathbb{R}^n$. On it we have the the standard $\mathcal{C}$-structure of the previous example. For any injective map ϕ with open image we also have the $\mathcal{C}$-structure given by the equivalent class of the the $\mathcal{C}$-atlas $\{(\mathbb{R}^n, \phi)\}$. The two structures define the same $\mathcal{C}$-manifold iff ϕ and its inverse are $\mathcal{C}$-maps. Notice that if ϕ is not a homeorphism, the two manifolds are different also as topological spaces. Suppose that ϕ is a homeomorphism but not a $\mathcal{C}^k$-diffeomorphism; then the two structures define the same topological space and the same $\mathcal{C}^0$-manifold, but not the same $\mathcal{C}^k$-manifold.[4]

Example 2.19. Let $\mathcal{A} = \{(U_\alpha, \phi_\alpha)\}_{\alpha \in I}$ be a $\mathcal{C}$-atlas on M. Let V be an open subset of U_α for some α. Define $\psi_V := \phi_\alpha|_V$. Then $\mathcal{A}' := \mathcal{A} \cup \{(V, \psi_V)\}$ is also a $\mathcal{C}$-atlas and moreover $\mathcal{A}$ and $\mathcal{A}'$ are $\mathcal{C}$-equivalent, so they define the same manifold. This example shows that in a manifold we can always shrink a chart to a smaller one.

Example 2.20 (Open subsets)**.** Let U be an open subset of a $\mathcal{C}$-manifold M. If $\{(U_\alpha, \phi_\alpha)\}_{\alpha \in I}$ is an atlas for M, then $\{(U_\alpha \cap U, \phi_\alpha|_{U_\alpha \cap U})\}_{\alpha \in I}$ is a $\mathcal{C}$-atlas for U. This makes U into a $\mathcal{C}$-manifold with the relative topology.

Example 2.21 (Cartesian product)**.** Let M and N be $\mathcal{C}$-manifolds. We can make $M \times N$ into a $\mathcal{C}$-manifold as follows. Let $\{(U_\alpha, \phi_\alpha)\}_{\alpha \in I}$ be a $\mathcal{C}$-atlas for M and $\{(V_j, \psi_j)\}_{j \in J}$ a $\mathcal{C}$-atlas for N. Then $(U_\alpha \times V_j, \phi_\alpha \times \psi_j)_{(\alpha,j) \in I \times J}$ is a $\mathcal{C}$-atlas for $M \times N$, called the product atlas. Note that the topology it induces is the product topology.

2.1. **Coordinates.** Recall that an element of an open subset V of $\mathbb{R}^n$ is an n-tuple $(x^1, \ldots, x^n)$ of real numbers called coordinates. We also have maps $\pi^i \colon V \to \mathbb{R}$, $(x^1, \ldots, x^n) \mapsto x^i$ called coordinate functions. One often writes x^i instead of π^i to denote a coordinate function. Notice that x^i has then both the meaning of a coordinate (a real number) and of a coordinate function (a function on V), but this ambiguity causes no problems in practice.

If (U, ϕ_U) is a chart with codomain $\mathbb{R}^n$, the maps $\pi^i \circ \phi_U \colon U \to \mathbb{R}$ are also called coordinate functions and are often denoted by x^i. One usually calls U together with its coordinate functions a coordinate neighborhood.

[4]We will see in Example 3.10, that these two $\mathcal{C}^k$-manifolds are anyway $\mathcal{C}^k$-diffeomorphic.

2.2. **Dimension.** Recall that the existence of a $\mathcal{C}^k$-diffeomorphism between an open subset of $\mathbb{R}^m$ and an open subset of $\mathbb{R}^n$ implies $m = n$ since the differential at any point is a linear isomorphism of $\mathbb{R}^m$ and $\mathbb{R}^n$ as vector spaces (the result is also true for homeomorphisms, though the proof is more difficult). So we have the

Definition 2.22. A connected manifold has dimension n if for any (and hence for all) of its charts the target of the chart map is $\mathbb{R}^n$. In general, we say that a manifold has dimension n if all its connected components have dimension n. We write $\dim M = n$.

2.3. **The implicit function theorem.** As mentioned in Example 2.12, a typical way of defining manifolds is by the implicit function theorem which we recall here.

Theorem 2.23 (Implicit function theorem)**.** *Let W be an open subset of $\mathbb{R}^{m+n}$, $F\colon W \to \mathbb{R}^m$ a $\mathcal{C}^k$-map ($k > 0$) and $c \in \mathbb{R}^n$. We define $M := F^{-1}(c)$. If for every $q \in M$ the linear map $\mathrm{d}_q F$ is surjective, then M has the structure of an m-dimensional $\mathcal{C}^k$-manifold with topology induced from $\mathbb{R}^{m+n}$.*

The proof of this theorem relies on another important theorem in analysis:

Theorem 2.24 (Inverse function theorem)**.** *Let W be an open subset of $\mathbb{R}^s$ and $G\colon W \to \mathbb{R}^s$ a $\mathcal{C}^k$-map ($k > 0$). If $\mathrm{d}_q G$ is an isomorphism at $q \in W$, then there is an open neighborhood V of q in W, such that $G|_V$ is a $\mathcal{C}^k$-diffeomorphism $V \to G(V)$.*

The inverse function theorem is a nice application of Banach's fixed point theorem. We do not prove it here (see e.g. [3, Appendix 10.1]).

Sketch of a proof of the implicit function theorem. Let $q \in M$. The matrix with entries $\frac{\partial F^i}{\partial x^j}(q)$, $i = 1, \dots, m$, $j = 1, \dots, n+m$ has by assumption rank m. This implies that we can rearrange its rows so that its left $m \times m$ block is invertible. More precisely, we can find a permutation σ of $\{1, \dots, m+n\}$ such that $(\frac{\partial \widetilde{F}^i}{\partial x^j}(q))_{i,j=1,\dots,m+n}$ is invertible, where $\widetilde{F} = F \circ \Phi_\sigma$ and Φ_σ is the diffeomorphism of $\mathbb{R}^{m+n}$ that sends $(x^1, \dots, x^{m+n})$ to $(x^{\sigma(1)}, \dots x^{\sigma(m+n)})$. We then define a new map $G\colon W \to \mathbb{R}^{m+n}$, $(x^1, \dots, x^{m+n}) \mapsto (\widetilde{F}^1, \dots, \widetilde{F}^m, x^{m+1}, \dots x^{m+n})$. Now $\mathrm{d}_q G$ is invertible, so we can apply the inverse function theorem to it. This means that there is a neighborhood V of q in W such that $G|_V$ is a $\mathcal{C}^k$-diffeomorphism $V \to G(V)$. We then define $U_q := V \cap M$ and $\phi_{U_q} := \pi \circ G|_{U_q}$ as a chart around q, where $\pi\colon \mathbb{R}^{m+n} \to \mathbb{R}^n$ is the projection to the last n coordinates. Repeating this for all $q \in M$, or just enough of them for the U_qs to cover M, we get an atlas for M. One can finally check that this atlas is $\mathcal{C}^k$. Since the maps G are, in particular, homeomorphisms, the atlas topology is the same as the induced topology. □

3. Maps

Let $F\colon M \to N$ be a map of sets. Let (U, ϕ_U) be a chart on M and (V, ψ_V) be a chart on N with $V \cap F(U) \neq \emptyset$. The map

$$F_{U,V} := \psi_{V|_{V\cap F(U)}} \circ F_{|_U} \circ \phi_U^{-1}\colon \phi_U(U) \to \psi_V(V)$$

is called the representation of F in the charts (U, ϕ_U) and (V, ψ_V). Notice that a map is completely determined by all its representations in a given atlas.

Definition 3.1. A map $F\colon M \to N$ between $\mathcal{C}$-manifolds is called a $\mathcal{C}$-map or $\mathcal{C}$-morphism if all its representations are $\mathcal{C}$-maps.

In Proposition 5.3 we will give a handier characterization of $\mathcal{C}$-maps in the case when the target N has a Hausdorff topology.

Remark 3.2. If we pick another chart $(U', \phi_{U'})$ on M and another chart $(V', \psi_{V'})$ on N, we get

$$F_{U',V'}|_{\phi_{U'}(U\cap U')} = \psi_{V,V'} \circ F_{U,V}|_{\phi_U(U\cap U')} \circ \phi_{U,U'}^{-1}. \tag{3.1}$$

This has two consequences. The first is that it is enough to choose one atlas in the equivalence class of the source and one atlas in the equivalence class of the target and to check that all representations are $\mathcal{C}$-maps for charts of these two atlases: the condition will then automatically hold for any other atlases in the same class. The second is that a collection of maps between chart images determines a map between manifolds only if equation (3.1) is satisfied for all transition maps. More precisely, fix an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ of M and an atlas $\{(V_j, \psi_j)\}_{j\in J}$ of N. Then a collection of $\mathcal{C}$-maps $F_{\alpha,j}\colon \phi_\alpha(U_\alpha) \to \psi_j(V_j)$ determines a $\mathcal{C}$-map $F\colon M \to N$ only if

$$F_{\alpha' j'}|_{\phi_{\alpha'}(U_\alpha\cap U_{\alpha'})} = \psi_{jj'} \circ F_{\alpha j}|_{\phi_\alpha(U_\alpha\cap U_{\alpha'})} \circ \phi_{\alpha\alpha'}^{-1}, \quad \forall \alpha, \alpha' \in I\ \forall j, j' \in J.$$

Definition 3.3. A $\mathcal{C}$-map from a $\mathcal{C}$-manifold M to $\mathbb{R}$ with its standard manifold structure is called a $\mathcal{C}$-function. We denote by $\mathcal{C}(M)$ the vector space of $\mathcal{C}$-functions on M.

Remark 3.4. In the case of a function, we always choose the standard atlas for the target $\mathbb{R}$. Therefore, we may simplify the notation: we simply write

$$f_U := f|_U \circ \phi_U^{-1}\colon \phi_U(U) \to \mathbb{R}.$$

If $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ is an atlas on M, a collection of $\mathcal{C}$-functions f_α on $\phi_\alpha(U_\alpha)$ determines a $\mathcal{C}$-function f on M with $f_{U_\alpha} = f_\alpha$ $\forall\alpha \in I$ if and only if

$$\boxed{f_\beta(\phi_{\alpha\beta}(x)) = f_\alpha(x)} \tag{3.2}$$

for all $\alpha, \beta \in I$ and for all $x \in \phi_\alpha(U_\alpha \cap U_\beta)$.

Remark 3.5. Notice that a $\mathcal{C}^k$-map between open subsets of Cartesian powers of $\mathbb{R}$ is also automatically $\mathcal{C}^l$ $\forall l \leq k$, so a $\mathcal{C}^k$-manifold can be regarded also as a $\mathcal{C}^l$-manifold $\forall l \leq k$. As a consequence, $\forall l \leq k$, we have the notion of $\mathcal{C}^l$-maps between $\mathcal{C}^k$-manifolds and of $\mathcal{C}^l$-functions on a $\mathcal{C}^k$-manifold.

Definition 3.6. An invertible $\mathcal{C}$-map between $\mathcal{C}$-manifolds whose inverse is also a $\mathcal{C}$-map is called a $\mathcal{C}$-isomorphism. A $\mathcal{C}^k$-isomorphism, $k \geq 1$, is usually called a $\mathcal{C}^k$-diffeomorphism (or just a diffeomorphism).

Example 3.7. Let M and N be open subsets of Cartesian powers of $\mathbb{R}$ with the standard $\mathcal{C}$-manifold structure. Then a map is a $\mathcal{C}$-map of $\mathcal{C}$-manifolds iff it is a $\mathcal{C}$-map in the standard sense.

Example 3.8. Let M be a $\mathcal{C}$-manifold and U an open subset thereof. We consider U as a $\mathcal{C}$-manifold as in Example 2.20. Then the inclusion map $\iota\colon U \to M$ is a $\mathcal{C}$-map.

Example 3.9. Let M and N be $\mathcal{C}$-manifolds and $M \times N$ their Cartesian product as in Example 2.21. Then the two canonical projections $\pi_M\colon M \times N \to M$ and $\pi_N\colon M \times N \to N$ are $\mathcal{C}$-maps.

Example 3.10. Let M be $\mathbb{R}^n$ with the equivalence class of the atlas $\{(\mathbb{R}^n, \phi)\}$, where ϕ is an injective map with open image. Let N be $\mathbb{R}^n$ with its standard structure. Then $\phi\colon M \to N$ is a $\mathcal{C}$-map for any $\mathcal{C}$ (since its representation is the identity map on open subset of $\mathbb{R}^n$). If in addition ϕ is also surjective, then $\phi\colon M \to N$ is a $\mathcal{C}$-isomorphism.[5]

Remark 3.11. Let M and N be as in the previous example with ϕ a bijection. Assume that $\phi\colon \mathbb{R} \to \mathbb{R}$ is a homeomorphism but not a $\mathcal{C}^k$-diffeomorphism. Then the given atlases are $\mathcal{C}^0$-equivalent but not $\mathcal{C}^k$-equivalent. As a consequence, M and N are the same $\mathcal{C}^0$-manifold but different $\mathcal{C}^k$-manifolds. On the other hand, $\phi\colon M \to N$ is always a $\mathcal{C}^k$-diffeomorphism of $\mathcal{C}^k$-manifolds. More difficult is to find examples of two $\mathcal{C}^k$-manifolds that are the same $\mathcal{C}^0$-manifold (or $\mathcal{C}^0$-isomorphic to each other), but are different, non $\mathcal{C}^k$-diffeomorphic $\mathcal{C}^k$-manifolds. Milnor constructed a $\mathcal{C}^\infty$-manifold structure on the 7-sphere that is not

[5] In general, ϕ is a $\mathcal{C}$-isomorphism from M to the open subset $\phi(M)$ of $\mathbb{R}^n$.

diffeomorphic to the standard 7-sphere. From the work of Donaldson and Freedman one can derive uncountably many different C^∞-manifold structures on $\mathbb{R}^4$ (called the exotic $\mathbb{R}^4$s) that are not diffeomorphic to each other nor to the standard $\mathbb{R}^4$. In dimension 3 and less, one can show that any two $\mathcal{C}^0$-isomorphic manifolds are also diffeomorphic.

3.1. **The pullback.** If M and N are $\mathcal{C}$-manifold and $F\colon M \to N$ is a $\mathcal{C}$-map, the $\mathbb{R}$-linear map

$$\begin{array}{rccc} F^*\colon & \mathcal{C}(N) & \to & \mathcal{C}(M) \\ & f & \mapsto & f \circ F \end{array}$$

is called pullback by F. If $f, g \in \mathcal{C}(N)$, then clearly

$$F^*(fg) = F^*(f)F^*(g).$$

Moreover, if $G\colon N \to Z$ is also a $\mathcal{C}$-map, then

$$(G \circ F)^* = F^* G^*.$$

Remark 3.12. We can rephrase Remark 3.4 by using pullbacks. Namely, if f is a function on M, then its representation in the chart (U, ϕ_U) is $f_U = (\phi_U^{-1})^* f|_U$. Moreover, if $\{(U_\alpha, \phi_\alpha)\}_{\alpha \in I}$ is an atlas on M, a collection of $\mathcal{C}$-functions f_α on $\phi_\alpha(U_\alpha)$ determines a $\mathcal{C}$-function f on M with $f_{U_\alpha} = f_\alpha$ $\forall \alpha \in I$ if and only if

$$\boxed{f_\alpha = \phi_{\alpha\beta}^* f_\beta} \tag{3.3}$$

for all $\alpha, \beta \in I$, where, by abuse of notation, f_α denotes here the restriction of f_α to $\phi_\alpha(U_\alpha \cap U_\beta)$ and f_β denotes the restriction of f_β to $\phi_\beta(U_\alpha \cap U_\beta)$.

Remark 3.13 (The push-forward)**.** If $F\colon M \to N$ is a $\mathcal{C}$-isomorphism, it is customary to denote the inverse of F^* by F_* and to call it the push-forward. Explicitly,

$$\begin{array}{rccc} F_*\colon & \mathcal{C}(M) & \to & \mathcal{C}(N) \\ & f & \mapsto & f \circ F^{-1} \end{array}$$

By this notation equation (3.3) reads

$$\boxed{f_\beta = (\phi_{\alpha\beta})_* f_\alpha} \tag{3.4}$$

Also note that

$$F_*(fg) = F_*(f)F_*(g)$$

and that, if $G\colon N \to Z$ is also a $\mathcal{C}$-map, then

$$(G \circ F)_* = G_* F_*.$$

3.2. **Submanifolds.** A submanifold is a subset of a manifold that is locally given by fixing some coordinates. More precisely:

Definition 3.14. Let N be an n-dimensional $\mathcal{C}$-manifold. A k-dimensional $\mathcal{C}$-submanifold, $k \leq n$, is a subset M of N such that there is a $\mathcal{C}$-atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha \in I}$ of N with the property that $\forall \alpha$ such that $U_\alpha \cap M \neq \emptyset$ we have $\phi_\alpha(U_\alpha \cap M) = W_\alpha \times \{x\}$ with W_α open in $\mathbb{R}^k$ and x in $\mathbb{R}^{n-k}$. Any chart with this property is called an adapted chart and an atlas consisting of adapted charts is called an adapted atlas. Notice that by a diffeomorphism of $\mathbb{R}^n$ we can always assume that $x = 0$.

Remark 3.15. Let $\{(U_\alpha, \phi_\alpha)\}_{\alpha \in I}$ be an adapted atlas for $M \subset N$. Then $\{(V_\alpha, \psi_\alpha)\}_{\alpha \in I}$, with $V_\alpha := U_\alpha \cap M$ and $\psi_\alpha := \pi \circ \phi_\alpha|_{V_\alpha} \colon V_\alpha \to \mathbb{R}^k$, where $\pi \colon \mathbb{R}^n \to \mathbb{R}^k$ is the projection to the first k coordinates, is a $\mathcal{C}$-atlas for M. Moreover, the inclusion map $\iota \colon M \to N$ is clearly a $\mathcal{C}$-map.

Remark 3.16. In an adapted chart (U_α, ϕ_α) the k-coordinates of W_α parametrize the submanifold and are called tangential coordinates, while the remaining $n - k$ coordinates are called transversal coordinates and parametrize a transversal neighborhood of a point of the submanifold.

Example 3.17. Any open subset M of a manifold N is a submanifold as any atlas of N is automatically adapted. In this case, there are no transversal coordinates.

Remark 3.18. Notice that a chart (U, ψ_U) such that $\psi_U(U \cap M)$ is the graph of a map immediately leads to an adapted chart. To be precise, assume $\psi_U(U \cap M) = \{(x, y) \in V \times \mathbb{R}^{n-k} \mid y = F(x)\}$ with V open in $\mathbb{R}^k$ and F a $\mathcal{C}$-map from V to $\mathbb{R}^{n-k}$. Then let $\Phi \colon V \times \mathbb{R}^{n-k} \to V \times \mathbb{R}^{n-k}$ be defined by $\Phi(x, y) = (x, y - F(x))$. It is clearly a $\mathcal{C}$-isomorphism. Moreover, (U, ϕ_U), with $\phi_U := \Phi \circ \psi_U$ is clearly an adapted chart (with $\phi_U(U \cap M) = V \times \{0\}$).

As a consequence, we may relax the definition by allowing adapted charts (U_α, ϕ_α) such that $\phi_\alpha(U_\alpha \cap M)$ is the graph of a map. In particulat, we have the

Example 3.19 (Graphs)**.** Let F be a $\mathcal{C}$-map from open subset V of $\mathbb{R}^k$ to $\mathbb{R}^{n-k}$ and consider its graph $M = \{(x, y) \in V \times \mathbb{R}^{n-k} \mid y = F(x)\}$. Then M is a $\mathcal{C}$-submanifold of $N = V \times \mathbb{R}^{n-k}$. As an adapted atlas we may take the one consisting of the single chart (N, ι), where $\iota \colon N \to \mathbb{R}^n$ is the inclusion map.

A further consequence is that a subset of the standard $\mathbb{R}^n$ defined in terms of $\mathcal{C}^k$-constraints satisfying the assumptions of the implicit function theorem is a $\mathcal{C}^k$-submanifold. There is a more general version of this, the implicit function theorem for manifolds, which we will see later as Theorem 6.13 on page 25.

4. Topological manifolds

In this Section we concentrate on $\mathcal{C}^0$-manifolds. Notice however that every $\mathcal{C}$-manifold is by definition also a $\mathcal{C}^0$-manifold.

As we have seen, an atlas whose chart maps have open images defines a topology. In this topology the chart maps are clearly open maps. We also have the

Lemma 4.1. *All the chart maps of a $\mathcal{C}^0$-atlas are continuous, so they are homeomorphisms with their images.*

Proof. Consider a chart (U_α, ϕ_α), $\phi_\alpha \colon U_\alpha \to \mathbb{R}^n$. Let V be an open subset of $\mathbb{R}^n$ and $W := \phi_\alpha^{-1}(V)$. For any chart (U_β, ϕ_β) we have $\phi_\beta(W \cap U_\beta) = \phi_{\alpha\beta}(V)$. In a $\mathcal{C}^0$-atlas, all transition maps are homeomorphisms, so $\phi_{\alpha\beta}(V)$ is open for all β, which shows that W is open. We have thus proved that ϕ_α is continuous. Since we already know that it is injective and open, we conclude that it is a homeomorphism with its image.[6] □

Different atlases in general define different topologies. However,

Lemma 4.2. *Two $\mathcal{C}^0$-equivalent $\mathcal{C}^0$-atlases define the same topology.*

Proof. Let $\mathcal{A}_1 = \{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ and $\mathcal{A}_2 = \{(U_j, \phi_j)\}_{j\in J}$ be $\mathcal{C}^0$-equivalent. First observe that by the equivalence condition $\phi_\alpha(U_\alpha \cap U_j)$ is open for all $\alpha \in I$ and for all $j \in J$.

Let W be open in the $\mathcal{A}_1$-topology. We have $\phi_j(W \cap U_\alpha \cap U_j) = \phi_{\alpha j}(\phi_\alpha(W \cap U_\alpha \cap U_j))$. Moreover, $\phi_\alpha(W \cap U_\alpha \cap U_j) = \phi_\alpha(W \cap U_\alpha) \cap \phi_\alpha(U_\alpha \cap U_j)$, which is open since W is $\mathcal{A}_1$-open. Since the atlases are equivalent, we also know that $\phi_{\alpha j}$ is a homeomorphism. Hence $\phi_j(W \cap U_\alpha \cap U_j)$ is open. Since this holds for all $j \in J$, we get that $W \cap U_\alpha$ is open in the $\mathcal{A}_2$-topology. Finally, we write $W = \cup_{\alpha\in I} W \cap U_\alpha$, i.e., as a union of $\mathcal{A}_2$-open set. This shows that W is open in the $\mathcal{A}_2$-topology for all $\alpha \in I$. □

As a consequence a $\mathcal{C}^0$-manifold has a canonically associated topology in which all charts are homeomorphism. This suggests the following

---

[6] Notice that the proof of ths Lemma does not require the condition that the atlas be open. We only need the conditions that the chart maps be open and that the transition functions be continuous.

Definition 4.3. A topological manifold is a topological space endowed with an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ in which all U_α are open and all ϕ_α are homeomorphisms to their images.

Theorem 4.4. *A topological manifold is the same as a $\mathcal{C}^0$-manifold.*

Proof. We have seen above that a $\mathcal{C}^0$-manifold structure defines a topology in which every atlas in the equivalence class has the properties in the definition of a topological manifold; so a $\mathcal{C}^0$-manifold is a topological manifold. On the other hand, the atlas of a topological manifold is open and all transition maps are homeomorphism since they are now compositions of homeomorphisms. The $\mathcal{C}^0$-equivalence class of this atlas then defines a $\mathcal{C}^0$-manifold. □

Also notice the following

Lemma 4.5. *Let M and N be $\mathcal{C}^0$-manifolds and so, consequently, topological manifolds. A map $F\colon M \to N$ is a $\mathcal{C}^0$-map iff it is continuous. In particular, a $\mathcal{C}^0$-isomorphism is the same as a homeomorphism.*

Proof. Suppose that F is a $\mathcal{C}^0$-map. Let $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ be an atlas on M and $\{(V_\beta, \psi_\beta)\}_{\beta\in J}$ be an atlas on N. For every $W \subset N$, $\forall\alpha \in I$ and $\forall\beta \in J$, we have $\phi_\alpha(F^{-1}(W \cap V_\beta) \cap U_\alpha) = F^{-1}_{\alpha,\beta}(\psi_\beta(W \cap V_\beta))$. If W is open, then $\psi_\beta(W \cap V_\beta)$ is open for all β. Since all $F_{\alpha,\beta}$ are continuos, we conclude that $\phi_\alpha(F^{-1}(W \cap V_\beta) \cap U_\alpha)$ is open for all α and all β. Hence, $F^{-1}(W \cap V_\beta)$ is open for all β, so $F^{-1}(W) = \cup_{\beta\in J} F^{-1}(W \cap V_\beta)$ is open. This shows that F is continuous.

On the other hand, if F is continous, then all its representations are also continuous since all chart maps are homeomorphisms. Thus, F is a $\mathcal{C}^0$-map. □

Remark 4.6. In the following we will no longer distinguish between $\mathcal{C}^0$-manifolds and topological manifolds.[7] Both descriptions are useful. Sometimes we are given a set with charts (like in the example of the manifold of lines in the plane). In other cases, we are given a topological space directly (like in all examples when our manifold arises as a subset of another manifold, e.g., $\mathbb{R}^n$).

Remark 4.7. Notice that a $\mathcal{C}$-manifold may equivalently be defined as a topological manifold where all transition functions are of class $\mathcal{C}$.

In the definition of a manifold, several textbooks assume the topology to be Hausdorff and second countable. These properties have important

[7] What we have proved above is that the category of $\mathcal{C}^0$-manifolds and the category of topological manifolds are isomorphic, if you know what categories are.

consequences (like the existence of a partition of unity which is fundamental in several contexts, e.g., in showing the existence of Riemannian metrics, in defining integrals and in proving Stokes theorem), but are not strictly necessary otherwise, so we will not assume them here unless explicitly stated. Also notice that non-Hausdorff manifolds often arise out of important, natural constructions.

Example 4.8 (The line with two origins)**.** Let $M := \mathbb{R} \cup \{*\}$ where $\{*\}$ is a one-element set (and $* \notin \mathbb{R}$). Let $U_1 = \mathbb{R}$, $\phi_1 = \mathrm{Id}$, and $U_2 = (\mathbb{R}\backslash\{0\})\cup\{*\}$ with $\phi_2\colon U_2 \to \mathbb{R}$ defined by $\phi_2(x) = x$ if $x \in \mathbb{R}\backslash\{0\}$ and $\phi_2(*) = 0$. One can easily see that this is a $\mathcal{C}^0$-atlas (actually a $\mathcal{C}^\infty$-atlas, for the transition functions are just identity maps). On the other hand, the induced topology is not Hausdorff, for 0 and $*$ do not have disjoint open neighborhoods.

Remark 4.9. Every manifold that is defined as a subset of $\mathbb{R}^n$ by the implicit function theorem inherits from $\mathbb{R}^n$ the property of being Hausdorff.

4.1. **Manifolds by local data.** The transition maps $\phi_{\alpha\beta}$ are actually all what is needed to define a manifold (with a specific atlas). Namely, assume that we have an index set I and

(1) for each $\alpha \in I$ a nonempty open subset V_α of $\mathbb{R}^n$, and
(2) for each β different from α an open subset $V_{\alpha\beta}$ of V_α and a $\mathcal{C}$-map $\phi_{\alpha\beta}\colon V_{\alpha\beta} \to V_{\beta\alpha}$,

such that, for all α, β, γ,

(i) $\phi_{\alpha\beta} \circ \phi_{\beta\alpha} = \mathrm{Id}$, and
(ii) $\phi_{\beta\gamma}(\phi_{\alpha\beta}(x)) = \phi_{\alpha\gamma}(x)$ for all $x \in V_{\alpha\beta} \cap V_{\alpha\gamma}$.

On the topological space $\widetilde{M}$, defined as the disjoint union of all the V_αs, we introduce the relation $x \sim y$ to hold if either $x = y$ or, for some α and β, $x \in V_{\alpha\beta}$ and $y = \phi_{\alpha\beta}(x)$. By the conditions above this is an equivalence relation. We then define M as the quotient space $\widetilde{M}/\sim$ with the quotient topology. We denote by $\pi\colon \widetilde{M} \to M$ the canonical projection and set

$$U_\alpha := \pi(V_\alpha).$$

Note that, since $\pi^{-1}(U_\alpha) = V_\alpha \sqcup \bigsqcup_{\beta\neq\alpha} \phi_{\alpha\beta}(V_{\alpha\beta})$ and the $\phi_{\alpha\beta}$s are homeomorphisms, each U_α is open. Also note that for each $q \in U_\alpha$ there is a unique $x_q \in V_\alpha$ with $\pi(x_q) = q$; we use this to define a map

$$\phi_\alpha\colon U_\alpha \to \mathbb{R}^n$$

which sends q to x_q. It is clear that this map is continuous and open and that its image is V_α. Moreover, if $x \in U_\alpha \cap U_\beta$, the unique $x_q \in V_\alpha$

with $\pi(x_q) = q$ and the unique $y_q \in V_\beta$ with $\pi(y_q) = q$ are related by $y_q = \phi_{\alpha\beta}(x_q)$. It then follows that $\phi_{\alpha\beta}(x) = \phi_\beta(\phi_\alpha^{-1}(x))$ for all $x \in V_{\alpha\beta}$. Hence, $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ is a $\mathcal{C}$-atlas on M. We say that the local data $(V_\alpha, V_{\alpha\beta}, \phi_{\alpha\beta})$ define the manifold M by the $\mathcal{C}$-equivalence class of this atlas.

Remark 4.10. This definition of a manifold is equivalent to the previous one. Above we have seen how to define M and assign it an atlas. Conversely, if we start with a manifold M and a $\mathcal{C}$-atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ on it, we define $V_\alpha := \phi_\alpha(U_\alpha)$, $V_{\alpha\beta} := \phi_\alpha(U_\alpha \cap U_\beta)$ and $\phi_{\alpha\beta}$ as the usual transition maps. One can easily see that the two constructions are inverse to each other.

Example 4.11. Let $I = \{1, 2\}$, $V_1 = V_2 = \mathbb{R}$, $V_{12} = V_{21} = \mathbb{R} \setminus \{0\}$ and $\phi_{12} = \phi_{21} = \mathrm{Id}$. Then M is the line with two origins of Example 4.8. This example shows that manifolds constructed by local data may be non Hausdorff.

Example 4.12. Let $I = \{1, 2\}$, $V_1 = V_2 = \mathbb{R}^n$, $V_{12} = V_{21} = \mathbb{R}^n \setminus \{0\}$ and $\phi_{12}(\mathbf{x}) = \phi_{21}(\mathbf{x}) = \frac{\mathbf{x}}{||\mathbf{x}||^2}$. As this actually defines the atlas one gets using the stereographic projections, we see that M is the n-sphere S^n.

5. Bump functions and partitions of unity

A bump function is a nonnegative function that is identically equal to 1 in some neighborhood and zero outside of a larger compact neighborhood.[8] Bump functions are used to extend locally defined objects to global ones. A notion that will be useful is that of support of a function, defined as the closure of the set on which the function does not vanish:

$$\operatorname{supp} f := \overline{\{x \in M \mid f(x) \neq 0\}}, \quad f \in \mathcal{C}(M).$$

An important fact is that bump functions exist. We start with the case of $\mathbb{R}$. Following [5], we first define

$$f(t) := \begin{cases} \mathrm{e}^{-\frac{1}{t}}, & t > 0, \\ 0 & t \leq 0, \end{cases}$$

which is $\mathcal{C}^\infty$ and hence $\mathcal{C}^k$ for every k. Next we set

$$g(t) := \frac{f(t)}{f(t) + f(1-t)}$$

and finally

$$h(t) = g(t+2)g(2-t).$$

[8] Definitions of bump functions vary in the literature.

Notice that h is $\mathcal{C}^\infty$, and hence $\mathcal{C}^k$ for every k, nonnegative, is identically equal to 1 in $[-1,1]$ and has support equal to $[-2,2]$. More generally, for every $y \in \mathbb{R}^n$ and every $R > 0$, we define

$$\psi_{y,R}(x) := h\left(\frac{2||x-y||}{R}\right).$$

This is a $\mathcal{C}^\infty$-function on $\mathbb{R}^n$ which is nonnegative, equal to 1 in the closed ball with center y and radius $R/2$ and with support the closed ball with center y and radius R.

Lemma 5.1. *Let M be a Hausdorff $\mathcal{C}^k$-manifold, $k \geq 0$ Then for every $q \in M$ and for every open U with $U \ni q$, there is a bump function $\psi \in \mathcal{C}^k(M)$ with $\operatorname{supp}\psi \subset U$ which is identically equal to 1 in an open subset V of U that contains q.*

Proof. Pick an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$, and let α be an index such that $U_\alpha \ni q$. Let $U' = \phi_\alpha(U \cap U_\alpha)$ and $y = \phi_\alpha(q)$. Let $R > 0$ and $\epsilon > 0$ be such that the open ball with center y and radius $R + \epsilon$ is contained in U' (this is possible, by definition, since U' is open). Let V_α denote the open ball with center y and radius $R/2$, W_α the open ball with center y and radius R and K_α the closure of W_α, i.e., the closed ball with center y and radius R. Notice that by the Heine–Borel theorem K_α is compact. Then set $V = \phi_\alpha^{-1}(V_\alpha)$, $W = \phi_\alpha^{-1}(W_\alpha)$ and $K = \phi_\alpha^{-1}(K_\alpha)$. Finally, set $\psi(x) = \psi_{y,R}(\phi_\alpha(x))$ for $x \in U_\alpha$ and $\psi(x) = 0$ for $x \in M\backslash U_\alpha$. We claim that ψ has the desired properties.

First, observe that ψ is identically equal to 1 in V and that V is open in U_α and hence in M. Next observe that by Lemma A.8 K is compact in U_α and hence in M. By the Hausdorff condition, Lemma A.11 implies that K is also closed in M. Hence $\operatorname{supp}\psi = K \subset U_\alpha$.

Finally, observe that W is also open in U_α and hence in M; hence $W \cap U_\beta$ is open for all β. We also clearly have that $K \cap U_\beta$ is closed for all β as we have already proved that K is closed. Let $W_\beta := \phi_\beta(W \cap U_\beta)$ and $K_\beta := \phi_\beta(K \cap U_\beta)$. We have that W_β is open, K_β is closed, and K_β is the closure of W_β. Now the representation ψ_β of ψ in $\phi_\beta(U_\beta)$ has support equal to K_β, so it is zero in the complement of K_β and hence smooth for every x in there. For $x \in K_\beta$, we have $\psi_\beta(x) = \psi_{y,R}(\phi_{\beta\alpha}(x))$. Hence ψ_β is of class $\mathcal{C}^k$ in the open subset W_β. Finally, let χ denote ψ_β or one of its derivatives up to order k, and $\chi_{y,R}$ the corresponding derivative of $\psi_{y,R}$. Then $\chi(x) = \chi_{y,R}(\phi_{\beta\alpha}(x))$ for every $x \in K_\beta$. By the continuity of $\chi_{y,R}$ and of $\phi_{\beta\alpha}$, we then have that χ is continuous on the whole of K_β. Hence ψ_β is of class $\mathcal{C}^k$. □

Remark 5.2. In the Example 4.8 of the line with two roigins, we see that a bump function around 0 in $\phi_1(U_1)$ has a support K_1, but the corresponding K is not closed as K_2 is $K_1 \setminus \{0\}$.

As a first application, we can give the following nice characterization of $\mathcal{C}$-maps.

Proposition 5.3. *Let $F\colon M \to N$ be a set theoretic map between $\mathcal{C}$-manifolds with N Hausdorff. Then F is a $\mathcal{C}$-map iff $F^*(\mathcal{C}(N)) \subset \mathcal{C}(M)$.*

Proof. If F is a $\mathcal{C}$-map and f a $\mathcal{C}$-function, we immediately see, choosing representations in charts, that F^*f is also a $\mathcal{C}$-function.

If, on the other hand, $F^*(\mathcal{C}(N)) \subset \mathcal{C}(M)$, we see that F is a $\mathcal{C}$-map by the following consideration. Let F_{WU} be a representation. Pick any point $p \in W$ and let ψ be a bump function as in Lemma 5.1 with $q = F(p)$. Define $f^i(x) := \phi^i_U(x)\,\psi(x)$ for $x \in U$ and 0 otherwise. Then $f^i \in \mathcal{C}(N)$ and hence $F^*f^i \in \mathcal{C}(M)$; i.e., $(F^*f^i)\circ\phi_W^{-1}$ is a $\mathcal{C}$-function on $\phi_W(W)$. Denoting by V the neighborhood of q where ψ is identically equal to 1, for $u \in \phi_W(F^{-1}(V)\cap W)$ we have $(F^*f^i)\circ\phi_W^{-1}(u) = F^i_{WV}(u)$, which shows that the ith component of F_{WV} is a $\mathcal{C}$-map in a neighborhood of $\phi_W(p)$. Since both p and i are arbitrary, F_{WU} is a $\mathcal{C}$-map. □

Remark 5.4. The condition that the target be Hausdorff is essential. Take for example N to be the line with two origins of Example 4.8 and $M = \mathbb{R}$. Consider the map $F\colon M \to N$ defined by

$$F(x) = \begin{cases} 0 & x \leq 0 \\ * & x > 0 \end{cases}$$

This map is not continuous: in fact, the preimage of the open set $\mathbb{R} = U_1 \subset N$ is the interval $(-\infty, 0]$ which is not open in M. On the other hand, the pullback of every continuous function f on N is the constant function on M, which is continuos (even C^∞). To see this, simply observe that if $x_0 := f(0)$ and $x_* := f(*)$ where distinct points in $\mathbb{R}$, then we could find disjoint open neighborhoods U_0 and U_* of them. But then $f^{-1}(U_0)$ and $f^{-1}(U_*)$ would be disjoint open neighborhoods of 0 and $*$, respectively, which is impossible, since N is not Hausdorff.

The next important concept is that of partition of unity, roughly speaking the choice of bump functions that decompose the function 1. This is needed for special constructions (e.g., of integration or of Riemannian metrics) and is not guaranteed unless extra topological assumptions are made. Even with assumptions, one in general needs infinitely many bump functions. To make sense of their sum, one assumes that in a neighborhood of each point only finitey many of them are different from zero. To make this more precise, we say that a collection $\{T_i\}_{i\in I}$ of subsets of a topological space is locally finite if every point in the space possesses an open neighborhood that intersects nontrivially only finitely many T_is.

Definition 5.5. Let M be a $\mathcal{C}$-manifold. A partition of unity on M is a collection $\{\rho_j\}_{j\in J}$ of of $\mathcal{C}$-bump functions on M such that:

(1) $\{\operatorname{supp}\rho_j\}_{j\in J}$ is locally finite, and
(2) $\sum_{j\in J}\rho_J(x) = 1$ for all $x \in M$.

One often starts with a cover $\{U_\alpha\}_{\alpha\in I}$ of M—e.g., by charts—and looks for a partition of unity $\{\rho_j\}_{j\in J}$ such that for every $j \in J$ there is an $\alpha_j \in I$ such that $\operatorname{supp}\rho_j \subset U_{\alpha_j}$. In this case, one says that the partition of unity is subordinate to the given cover.

Theorem 5.6. *Let M be a compact Hausdorff $\mathcal{C}^k$-manifold, $k \geq 0$. Then for every cover by charts there is a finite partition of unity subordinate to it.*

Proof. Let $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ be an atlas. For $x \in U_\alpha$, let $\psi_{x,\alpha}$ be a bump function with support inside U_α and equal to 1 on an open subset $V_{x,\alpha}$ of U_α containing x, see Lemma 5.1. Since $\{V_{x,\alpha}\}_{x\in M,\alpha\in I}$ is clearly a cover of M and M is compact, we have a finite subcover $\{V_{x_j,\alpha_j}\}_{j\in J}$. Since each x is contained in some V_{x_k,α_k}, we have $\psi_{x_k,\alpha_k}(x) = 1$ and hence $\sum_{j\in J}\psi_{x_j,\alpha_j}(x) > 0$. Thus,

$$\rho_j := \frac{\psi_{x_j,\alpha_j}}{\sum_{k\in J}\psi_{x_k,\alpha_k}}, \quad j \in J$$

is a partition of unity subordinate to the cover $\{U_\alpha\}_{\alpha\in I}$. □

A more general theorem, for whose proof we refer to the literature, e.g., [3, 5], is the following:

Theorem 5.7. *Let M be a Hausdorff, second countable $\mathcal{C}^k$-manifold, $k \geq 0$. Then for every open cover there is a partition of unity subordinate to it.*

Recall that a topological space S is second countable if there is a countable collection $\mathcal{B}$ of open sets such that every open set of S can be written as a union of some elements of $\mathcal{B}$. Note that $\mathbb{R}^n$ is second countable with $\mathcal{B}$ given, e.g., by the open balls with rational radius and rational center coordinates. As a subset of a second countable space is automatically second countable in the relative topology, we have that manifolds defined via the implicit function theorem in $\mathbb{R}^n$ are second countable. Hence we have

Remark 5.8. Every manifold that is defined as a subset of $\mathbb{R}^n$ by the implicit function theorem inherits from $\mathbb{R}^n$ the property of being Hausdorff and second countable.

6. Differentiable manifolds

A $\mathcal{C}^k$-manifold with $k \geq 1$ is also called a differentiable manifold. If $k = \infty$, one also speaks of a smooth manifold. The $\mathcal{C}^k$-morphisms are also called differentiable maps, and also smooth maps in case $k = \infty$. Recall the following

Definition 6.1. Let $F \colon U \to V$ be a differentiable map between open subsets of Cartesian powers of $\mathbb{R}$. The map F is called an immersion if $\mathrm{d}_x F$ is injective $\forall x \in U$ and a submersion if $\mathrm{d}_x F$ is surjective $\forall x \in U$.

Then we have the

Definition 6.2. A differentiable map between differentiable manifolds is called an immersion if all its representations are immersions and a submersion if all its representations are submersions. An embedding of differentiable manifolds is an embedding in the topological sense, see Definition A.12, which is also an immersion.

Observe that to check whether a map is an immersion or a submersion one just has to consider all representations for a given choice of atlases.

One can prove that the image of an embedding is a submanifold (and this is one very common way in which submanifolds arise in examples).

Remark 6.3. Some authors call submanifolds the images of (injective) immersions and embedded submanifolds (or regular submanifolds) the images of embeddings. Images of immersions are often called immersed submanifolds. This terminology unfortunately is different in different textbooks. Notice that only the image of an embedding is a submanifold if we stick to Definition 3.14.

Locally, we have the following characterization.

Proposition 6.4. *Let $F \colon N \to M$ be an injective immersion. If M is Hausdorff, then every point p in N has an open neighborhood U such that $F|_U$ is an embedding.*

Proof. Let (V, ψ) be a chart neighborhood of p. Since $\psi(V)$ is open, we can find an open ball, say of radius R, centered at $\psi(p)$ and contained in $\psi(V)$. The closed ball with radius $R/2$ centered at $\psi(p)$ is then also contained in $\psi(V)$ and is compact. Its preimage K under ψ is then also compact, as ψ is a homeomorphism. By Lemma A.13, the restriction of F to K is an embedding in the topological sense. It follows that the restriction of F to an open neighborhood U of p contained in K (e.g., the preimage under ψ of the open ball with radius $R/4$ centered at $\psi(p)$) is also an embedding in the topolological sense, but it is also an injective immersion. $\square$

6.1. **The tangent space.** Recall that to an open subset of $\mathbb{R}^n$ we associate another copy of $\mathbb{R}^n$, called its tangent space. Elements of this space, the tangent vectors, also have the geometric interpretation of velocities of curves passing through a point or of directions along which we can differentiate functions. We will use all these viewpoints to give different caracterizations of tangent vectors to a manifold, even though we relegate the last one, directional derivatives, to Section 7. In the following M is an n-dimensional $\mathcal{C}^k$-manifold, $k \geq 1$.

Let us consider first the case when M is defined in terms of constraints, i.e., as $\Phi^{-1}(c)$ with $\Phi\colon \mathbb{R}^n \to \mathbb{R}^l$ satisfying the condition of the implicit function theorem that $\mathrm{d}_q\Phi$ is surjective for all $q \in M$. We can then naturally define the tangent vectors at $q \in M$ as those vectors in $\mathbb{R}^n$ that do not lead us outside of M, i.e., as the directions along which Φ does not change. More precisely, a vector $v \in \mathbb{R}^n$ is tangent to M at q if $\sum_{j=1}^n v^j \frac{\partial \Phi^i}{\partial x^j}(q) = 0$ for all $i = 1, \dots, l$ (or, equivalently, $\mathrm{d}_q\Phi\, v = 0$). This viewpoint has several problems. The first is that it requires M to be presented in terms of constraints. The second is that it is not immediately obvious that this definition is independent of the choice of constraints. The third is that this definition is not necessarily the most practical way of defining the tangent vectors when one needs to make computations. It is on the other hand useful to remark that tangent vectors at $q \in M$, according to this definition, are also the same as the possible velocities of curves through q in M. Namely, let $\gamma\colon I \to \mathbb{R}^n$ be a differentiable map, with I an open interval, such that $\gamma(I) \subset M$. This means that $\Phi(\gamma(t)) = c\ \forall t \in I$. Let $q = \gamma(u)$ for some $u \in I$. Then, by the chain rule, we get $\sum_{j=1}^n \frac{\mathrm{d}\gamma^j}{\mathrm{d}t} \frac{\partial \Phi^i}{\partial x^j}(q) = 0$ for all $i = 1, \dots, l$, which shows that $\frac{\mathrm{d}\gamma}{\mathrm{d}t}$ is tangent to M at q.

Notice that the last viewpoint, that of tangent vectors as possible velocities of curves, can now be generalized also to manifolds not given in terms of constraints. Namely, let $\gamma\colon I \to M$ be a differentiable map, where I is an open interval with the standard manifold structure. For a fixed u in I, we set $q := \gamma(u)$. We wish to think of the velocity of γ

at u as a tangent vector at q.[9] The problem is that we do not know how to compute derivatives of maps between manifolds. The solution is to pick a chart (U, ϕ_U) on M with $U \ni q$. We now know how yet to differentiate $\phi_U \circ \gamma \colon I \to \mathbb{R}^n$ and define

$$v_U = \frac{\mathrm{d}}{\mathrm{d}t}\phi_U(\gamma(t))|_{t=u}.$$

Notice that v_U is an element of $\mathbb{R}^n$ and we wish to think of it as the tangent vector we were looking for. We now have another problem, however; namely, the value of v_U depends on the choice of chart. On the other hand, we know exactly how to relate values corresponding to different chart. Let in fact (V, ϕ_V) be another chart with $V \ni q$. We define

$$v_V = \frac{\mathrm{d}}{\mathrm{d}t}\phi_V(\gamma(t))|_{t=u}.$$

For t in a neighborhood of u, we have $\phi_V(\gamma(t)) = \phi_{U,V}(\phi_U(\gamma(t)))$; hence, by the chain rule,

$$v_V = \mathrm{d}_{\phi_U(q)}\phi_{U,V}\, v_U.$$

All this motivates the following

Definition 6.5. A coordinatized tangent vector at $q \in M$ is a triple (U, ϕ_U, v) where (U, ϕ_U) is a chart with $U \ni q$ and v is an element of $\mathbb{R}^n$. Two coordinatized tangent vectors (U, ϕ_U, v) and (V, ϕ_V, w) at q are defined to be equivalent if $w = \mathrm{d}_{\phi_U(q)}\phi_{U,V} v$. A tangent vector at $q \in M$ is an equivalence class of coordinatized tangent vectors at q. We denote by T_qM, the tangent space of M at q, the set of tangent vectors at q.

A chart (U, ϕ_U) at q defines a bijection of sets

$$\begin{array}{rccc} \Phi_{q,U}\colon & T_qM & \to & \mathbb{R}^n \\ & [(U, \phi_U, v)] & \mapsto & v \end{array} \tag{6.1}$$

We will also simply write Φ_U when the point q is understood. Using this bijection, we can transfer the vector space structure from $\mathbb{R}^n$ to T_qM making Φ_U into a linear isomorphism. A crucial result is that this linear structure does not depend on the choice of the chart:

Lemma 6.6. *T_qM has a canonical structure of vector space for which $\Phi_{q,U}$ is an isomorphism for every chart (U, ϕ_U) containing q.*

---

[9] For this not to be ambiguous, we should assume that u is the only preimage of q; otherwise, we can think that γ defines a family of tangent vectors at u.

Proof. Given a chart (U, ϕ_U), the bijection Φ_U defines the linear structure

$$\lambda \cdot_U [(U, \phi_U, v)] = [(U, \phi_U, \lambda v)],$$
$$[(U, \phi_U, v)] +_U [(U, \phi_U, v')] = [(U, \phi_U, v + v')],$$

$\forall \lambda \in \mathbb{R}$ and $\forall v, v' \in \mathbb{R}^n$. If (V, ϕ_V) is another chart, we have

$$\begin{aligned}\lambda \cdot_U [(U, \phi_U, v)] &= [(U, \phi_U, \lambda v)] = \\ &= [(V, \phi_V, \mathrm{d}_{\phi_U(q)}\phi_{U,V}\lambda v)] = [(V, \phi_V, \lambda \mathrm{d}_{\phi_U(q)}\phi_{U,V} v)] = \\ &= \lambda \cdot_V [(V, \phi_V, \mathrm{d}_{\phi_U(q)}\phi_{U,V} v)] = \lambda \cdot_V [(U, \phi_U, v)],\end{aligned}$$

so $\cdot_U = \cdot_V$. Similarly,

$$\begin{aligned}&[(U, \phi_U, v)] +_U [(U, \phi_U, v')] = [(U, \phi_U, v + v')] = \\ &= [(V, \phi_V, \mathrm{d}_{\phi_U(q)}\phi_{U,V}(v + v'))] = [(V, \phi_V, \mathrm{d}_{\phi_U(q)}\phi_{U,V} v + \mathrm{d}_{\phi_U(q)}\phi_{U,V} v')] = \\ &= [(V, \phi_V, \mathrm{d}_{\phi_U(q)}\phi_{U,V} v)] +_V [(V, \phi_V, \mathrm{d}_{\phi_U(q)}\phi_{U,V} v')] = \\ &= [(U, \phi_U, v)] +_V [(U, \phi_U, v')],\end{aligned}$$

so $+_U = +_V$. □

From now on we will simply write $\lambda[(U, \phi_U, v)]$ and $[(U, \phi_U, v)] + [(U, \phi_U, v')]$ without the $_U$ label.

Notice that in particular we have

$$\boxed{\dim T_q M = \dim M}$$

where dim denotes on the left-hand-side the dimension of a vector space and on the right-hand-side the dimension of a manifold.

Let now $F\colon M \to N$ be a differentiable map. Given a chart (U, ϕ_U) of M containing q and a chart (V, ψ_V) of N containing $F(q)$, we have the linear map

$$\mathrm{d}_q^{U,V} F := \Phi_{F(q),V}^{-1} \mathrm{d}_{\phi_U(q)} F_{U,V} \Phi_{q,U}\colon T_q M \to T_{F(q)} N.$$

Lemma 6.7. *The linear map $\mathrm{d}_q^{U,V} F$ does not depend on the choice of charts, so we have a canonically defined linear map*

$$\boxed{\mathrm{d}_q F\colon T_q M \to T_{F(q)} N}$$

called the **differential** *of F at q.*

Proof. Let $(U', \phi_{U'})$ be also a chart containing q and $(V', \psi_{V'})$ be also a chart containing $F(q)$. Then

$$\begin{aligned}
\mathrm{d}_q^{U,V} F[(U, \phi_U, v)] &= [(V, \psi_V, \mathrm{d}_{\phi_U(q)} F_{U,V} v)] = \\
&= [(V', \psi_{V'}, \mathrm{d}_{\psi(F(q))} \psi_{V,V'} \, \mathrm{d}_{\phi_U(q)} F_{U,V} v)] = \\
&= [(V', \psi_{V'}, \mathrm{d}_{\phi'_U(q)} F_{U',V'} \, (\mathrm{d}_{\phi_U(q)} \phi_{U,U'})^{-1} v)] = \\
&= \mathrm{d}_q^{U',V'} F[(U', \phi_{U'}, (\mathrm{d}_{\phi_U(q)} \phi_{U,U'})^{-1} v)] = \mathrm{d}_q^{U',V'} F[(U, \phi_U, v)],
\end{aligned}$$

so $\mathrm{d}_q^{U,V} F = \mathrm{d}_q^{U',V'}$. □

We also immediately have the following

Lemma 6.8. *Let $F\colon M \to N$ and $G\colon N \to Z$ be differentiable maps. Then*

$$\mathrm{d}_q(G \circ F) = \mathrm{d}_{F(q)} G \, \mathrm{d}_q F$$

for all $q \in M$.

Remark 6.9. Notice that we can now characterize immersions and submersions, introduced in Definition 6.2, as follows: A differentiable map $F\colon M \to N$ is an immersion iff $\mathrm{d}_q F$ is injective $\forall q \in M$ and is a submersion iff $\mathrm{d}_q F$ is surjective $\forall q \in M$.

We now return to our original motivation:

Remark 6.10 (Tangent space by constraints)**.** Suppose M is a submanifold of $\mathbb{R}^n$ defined by l constraints satisfying the conditions of the implicit function theorem. We may reorganize the constraints as a map $\Phi\colon \mathbb{R}^n \to \mathbb{R}^l$ and obtain $M = \Phi^{-1}(c)$ for some $c \in \mathbb{R}^l$. The conditions of the implicit function theorem are that $\mathrm{d}_q\Phi$ is surjective for all $q \in M$. If we denote by $\iota\colon M \to \mathbb{R}^n$ the inclusion map, we have that $\Phi(\iota(q)) = c \ \forall q \in M$, i.e., $\Phi \circ \iota$ is constant. This implies $\mathrm{d}_q(\Phi \circ \iota) = 0$ and hence, by Lemma 6.8, $\mathrm{d}_{\iota(q)}\Phi \, \mathrm{d}_q\iota = 0$, which in turns implies $\mathrm{d}_q\iota(T_qM) \subset \ker \mathrm{d}_{\iota(q)}\Phi$. Since $\mathrm{d}_{\iota(q)}\Phi$ is surjective, $\mathrm{d}_q\iota$ is injective and $\dim T_qM = \dim M = n - l$, we actually get $\mathrm{d}_q\iota(T_qM) = \ker \mathrm{d}_{\iota(q)}\Phi$, which can be rewritten as

$$T_qM = \ker \mathrm{d}_q\Phi$$

if we abandon the pedantic distinction between q and $\iota(q)$ and regard T_qM as a subspace of $\mathbb{R}^n$. This is a common way of computing the tangent space. To be more explicit, let $\Phi^1, \dots, \Phi^l$ be the components of Φ. Then $T_qM = \{v \in \mathbb{R}^n \mid \sum_{j=1}^n \frac{\partial \Phi^i}{\partial x^j} v^j = 0 \ \forall i = 1, \dots, l\}$. This can also be rephrased as saying that v is tangent to M at q if "$q + \epsilon v$ belongs to M or an infinitesimal ϵ." Another interpretation is that, if M is defined by constraints, then T_qM is defined by the linearization of the constraints at q. One often writes this also using gradients and scalar products on $\mathbb{R}^n$, $T_qM = \{v \in \mathbb{R}^n \mid \nabla\Phi^i \cdot v = 0\ \forall i = 1, \dots, l\}$, and interprets this by saying that v is tangent to M at q if it is orthogonal to the gradients of all contraints. This last viewpoint, however, makes an unnecessary use of the Euclidean structure of $\mathbb{R}^n$.

Example 6.11. The n-dimensional unit sphere S^n is the preimage of 1 of the function $\phi(x) = \sum_{i=1}^{n+1}(x^i)^2$. By differentiating ϕ we then get that the tangent space at $x \in S^n$ is the space of vectors v in $\mathbb{R}^{n+1}$ satisfying $\sum_{i=1}^{n+1} v^i x^i = 0$. Making use of the Euclidean structure, we can also say that the tangent vectors at $x \in S^n$ are the vectors v in $\mathbb{R}^{n+1}$ orthogonal to x.

We finally come back to the other initial viewpoint in this subsection. A differentiable curve in M is a differentiable map $\gamma\colon I \to M$, where I is an open subset of $\mathbb{R}$ with its standard manifold structure. For $t \in I$, we define the velocity of γ at t as

$$\boxed{\dot\gamma(t) := \mathrm{d}_t\gamma 1 \in T_{\gamma(t)}M}$$

where 1 is the vector 1 in $\mathbb{R}$. Notice that for M an open subset of $\mathbb{R}^n$ this coincides with the usual definition of velocity.

For $q \in M$, define $\mathcal{P}_q$ as the space of differentiable curves $\gamma\colon I \to M$ such that $I \ni 0$ and $\gamma(0) = q$. It is easy to verify that the map $\mathcal{P}_q \to T_qM$, $\gamma \mapsto \dot\gamma(0)$ is surjective, so we can think of T_qM as the space of all possible velocities at q.

This observation together with Remark 6.10 yields a practical way of computing the tangent spaces of a submanifold of $\mathbb{R}^n$.

Example 6.12. Consider the group $O(n)$ of orthogonal $n\times n$ matrices. Since a matrix is specified by its entries, we may identify the space of $n \times n$ matrices with $\mathbb{R}^{n^2}$. A matrix A is orthogonal if $A^tA = \mathrm{Id}$. We can then consider the map $\phi(A) = A^tA - \mathrm{Id}$ and regard $O(n)$ as the preimage of the zero matrix. We have however to be careful with the target space: since the image of ϕ consists of symmetric matrices, taking the whole space of $n \times n$ matrices would make some constraints redundant. Instead we consider ϕ as a map from all $n \times n$ matrices to the symmetric ones, hence as a map $\mathbb{R}^{n^2} \to \mathbb{R}^{\frac{n(n+1)}{2}}$. This shows that $\dim O(n) = \frac{n(n-1)}{2}$. Alternatively, we may compute the dimension of $O(n)$ by computing that of its tangent space at some point, e.g., at the identity matrix. Namely, consider a path $A(t)$ with $A(0) = \mathrm{Id}$. Differentianting the defining relation and denoting $\dot A(0)$ by B, we get $B^t + B = 0$. This shows that tangent vectors at the identity matrix are the antisymmetric matrices and hence that $\dim T_{\mathrm{Id}}O(n) = \frac{n(n-1)}{2}$.

More examples of this sort can be analyzed by considering the general version of the implicit function theorem.

Theorem 6.13 (Implicit function theorem)**.** *Let $F\colon Z \to N$ be a $\mathcal{C}^k$-map ($k > 0$) of $\mathcal{C}^k$-manifolds of dimensions $m+n$ and n, respectively. Given $c \in N$, we define $M := F^{-1}(c)$. If for every $q \in M$ the linear map d_qF is surjective, then M has a unique structure of m-dimensional $\mathcal{C}^k$-manifold such that the inclusion map $\iota\colon M \to Z$ is an embedding.*

The proof is similar to the one in Cartesian powers of $\mathbb{R}$ by considering local charts. See, e.g., [5] for details. The considerations of Remark 6.10 generalize to this case. Namely, the tangent space at $q \in M$ can be realized as the kernel of d_qF.

6.2. **The tangent bundle.** We can glue all the tangent spaces of an n-dimensional $\mathcal{C}^k$-manifold M, $k \geq 1$, together:

$$\boxed{TM := \cup_{q\in M} T_qM}$$

An element of TM is usually denoted as a pair (q, v) with $q \in M$ and $v \in T_qM$.[10] We introduce the surjective map $\pi\colon TM \to M$, $(q, v) \mapsto q$. Notice that the fiber T_qM can also be obtained as $\pi^{-1}(q)$.

TM has the following structure of $\mathcal{C}^{k-1}$-manifold. Let $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ be an atlas in the equivalence class defining M. We set $\widehat{U}_\alpha := \pi^{-1}(U_\alpha)$ and

$$\begin{array}{rccc} \widehat{\phi}_\alpha\colon & \widehat{U}_\alpha & \to & \mathbb{R}^n \times \mathbb{R}^n \\ & (q, v) & \mapsto & (\phi_\alpha(q), \Phi_{q,U_\alpha} v) \end{array}$$

where Φ_{q,U_α} is the isomorphism defined in (6.1). Notice that the chart maps are linear in the fibers. The transition maps are then readily computed as

$$\boxed{\widehat{\phi}_{\alpha\beta}(x, w) = (\phi_{\alpha\beta}(x), \mathrm{d}_x\phi_{\alpha\beta} w)} \tag{6.2}$$

Namely, they are the tangent lifts of the transition maps for M and are clearly $\mathcal{C}^{k-1}$.

Definition 6.14. The tangent bundle of the $\mathcal{C}^k$-manifold M, $k \geq 1$, is the $\mathcal{C}^{k-1}$-manifold defined by the equivalence class of the above atlas.

Remark 6.15. Observe that another atlas on M in the same $\mathcal{C}^k$-equivalence class yields an atlas on TM that is $\mathcal{C}^{k-1}$-equivalent to previous one.

[10]Notice that we now denote by v a tangent vector at q, i.e., an equivalence class of coordinatized tangent vectors at q, and no longer an element of $\mathbb{R}^n$.

Remark 6.16. Notice that $\pi\colon TM \to M$ is a $\mathcal{C}^{k-1}$-surjective map and, if $k > 1$, a submersion.

Definition 6.17. If M and N are $\mathcal{C}^k$-manifolds and $F\colon M \to N$ is a $\mathcal{C}^k$-map, then the tangent lift

$$\widehat{F}\colon TM \to TN$$

is the $\mathcal{C}^{k-1}$-map

$$(q, v) \mapsto (F(q), \mathrm{d}_q F v).$$

6.3. **Vector fields.** A vector field is the attachment of a vector to each point; i.e., a vector field X on M is the choice of a vector $X_q \in T_q M$ for all $q \in M$. We also want this attachment to vary in the appropriate differentiability degree. More precisely:

Definition 6.18. A vector field on a $\mathcal{C}^k$-manifold M is a $\mathcal{C}^{k-1}$-map $X\colon M \to TM$ such that $\pi \circ X = \mathrm{Id}_M$.

Remark 6.19. In an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$, M and the corresponding atlas $\{(\widehat{U}_\alpha, \widehat{\phi}_\alpha)\}_{\alpha\in I}$, a vector field X is represented by a collection of $\mathcal{C}^{k-1}$-maps $X_\alpha\colon \phi_\alpha(U_\alpha) \to \mathbb{R}^n$. All these maps are related by

$$\boxed{X_\beta(\phi_{\alpha\beta}(x)) = \mathrm{d}_x\phi_{\alpha\beta}\, X_\alpha(x)} \tag{6.3}$$

for all $\alpha, \beta \in I$ and for all $x \in \phi_\alpha(U_\alpha \cap U_\beta)$. Notice that a collection of maps X_α satisfying all these relations defines a vector field and this is how often vector fields are introduced (cf. equation (3.2) on page 9 for functions).

Remark 6.20. The vector at q defined by the vector field X is usually denoted by X_q as well as by $X(q)$. The latter notation is often avoided as one may apply a vector field X to a function f, see below, and in this case the standard notation is $X(f)$. We also use X_α to denote the representation of X in the chart with index α, but this should not create confusion with the notation X_q for X at the point q.

Note that vector fields may be added and multiplied by scalars and by functions: if X and Y are vector fields, λ a real number and f a function, we set

$$(X+Y)_q := X_q + Y_q,$$
$$(\lambda X)_q := \lambda X_q,$$
$$(fX)_q := f(q) X_q.$$

This way the set $\mathfrak{X}^{k-1}(M)$ of vector fields on M acquires the structure of vector space over $\mathbb{R}$ and of module over $\mathcal{C}^{k-1}(M)$.

The explicit representation of a vector field over an open subset U of $\mathbb{R}^n$ depends on a choice of coordinates. If we change coordinates by a diffeomorphism ϕ, the expression of a vector field changes by the differential of ϕ. We have already made use of this in equation (6.3). We now want to generalize this to manifolds.

Remark 6.21 (The push-forward of vector fields)**.** Let $F\colon M \to N$ be a $\mathcal{C}^k$-map of $\mathcal{C}^k$-manifolds. If X is a vector field on M, then $\mathrm{d}_q F\, X_q$ is a vector in $T_{F(q)}N$ for each $q \in M$. If F is a $\mathcal{C}^k$-diffeomorphism, we can perform this construction for each $y \in N$, by setting $q = F^{-1}(y)$, and define a vector field, denoted by F_*X, on N:

$$(F_*X)_{F(q)} := \mathrm{d}_q F\, X_q, \quad \forall q \in M, \tag{6.4}$$

or, equivalently,

$$(F_*X)_y = \mathrm{d}_{F^{-1}(y)} F\, X_{F^{-1}(y)}, \quad \forall y \in N.$$

The $\mathbb{R}$-linear map $F_*\colon \mathfrak{X}^{k-1}(M) \to \mathfrak{X}^{k-1}(N)$ is called the push-forward of vector fields. Note the if $G\colon N \to Z$ is also a diffeomorphism we immediately have

$$(G \circ F)_* = G_* F_*.$$

We also obviously have $(F_*)^{-1} = (F^{-1})_*$.

In case of a change of coordinates ϕ on an open subset of $\mathbb{R}^n$, the change of representation of a vector field is precisely described by the push-forward by ϕ. In particular, we have $X_\alpha = (\phi_\alpha)_* X$ for the chart labeled by α,[11] and equation (6.3) can be written in the more transparent form

$$\boxed{X_\beta = (\phi_{\alpha\beta})_* X_\alpha} \tag{6.5}$$

Remark 6.22. The push-forward is also natural from the point of view of our motivation of vectors as possible velocities of curves. If γ is a curve in M tangent to X (i.e., $\frac{\mathrm{d}}{\mathrm{d}t}\gamma(t) = X_{\gamma(t)}$ for all t), then $F \circ \gamma$ is tangent to F_*X (i.e., $\frac{\mathrm{d}}{\mathrm{d}t}F(\gamma(t)) = (F_*X)_{F(\gamma(t))}$ for all t), as is easily verified.

Remark 6.23. The push-forward of vector fields is compatible with the push-forward of functions defined in Remark 3.13. Namely, a simple

[11]We resort here to a very common and very convenient abuse of notation. The precise, but pedantic expression should be $X_\alpha = (\phi_\alpha)_* X_{|_{U_\alpha}}$ as ϕ_α is a diffeomorphism from U_α to $\phi_\alpha(U_\alpha)$. Similarly, (6.5) pedantically reads

$$(X_\beta)_{|_{\phi_\beta(U_\alpha \cap U_\beta)}} = (\phi_{\alpha\beta})_* (X_\alpha)_{|_{\phi_\alpha(U_\alpha \cap U_\beta)}}.$$

calculation shows that, if X and f are a vector field and a function on M and $F\colon M\to N$ is a diffeomorphism, then

$$F_*(fX) = F_*f\, F_*X.$$

Remark 6.24. If M and N are open subsets of $\mathbb{R}^n$ and we write $\bar{X} := F_*X$, then, regarding X and $\bar{X}$ as maps from M or N to $\mathbb{R}^n$, (6.4) explicitly reads

$$\bar{X}^{\bar{\jmath}}(\bar{x}) = \sum_{j=1}^{n} \frac{\partial F^{\bar{\jmath}}}{\partial x^j}(x) X^j(x), \quad \forall x \in M, \tag{6.6}$$

where $\bar{x} := F(x)$.

We finally come to a last interpretation of vector fields. If U is an open subset of $\mathbb{R}^n$, X a $\mathcal{C}^{k-1}$-vector field and f a $\mathcal{C}^k$-function ($k>0$), then we can define

$$X(f) = \sum_{i=1}^{n} X^i \frac{\partial f}{\partial x^i},$$

where on the right hand side we regard X as a map $U\to\mathbb{R}^n$. Notice that the map $\mathcal{C}^k(U)\to\mathcal{C}^{k-1}(U)$, $f\mapsto X(f)$, is $\mathbb{R}$-linear and satisfies the Leibniz rule

$$X(fg) = X(f)g + fX(g).$$

This is a derivation in the terminology of subsection 7.1. If we now have a $\mathcal{C}^k$-manifold M and a vector field X on it, we can still define a derivation $\mathcal{C}^k(M)\to\mathcal{C}^{k-1}(M)$ as follows. First we pick an atlas $\{(U_\alpha,\phi_\alpha)\}_{\alpha\in I}$. We then have the representation X_α of X in the chart (U_α,ϕ_α) as in Remark 6.19. If f is a function on M, we assign to it its representation f_α as in Remark 3.4. We can then compute $g_\alpha := X_\alpha(f_\alpha)\in\mathcal{C}^{k-1}(\phi_\alpha(U_\alpha))$ for all $\alpha\in I$. From (3.4) and (6.5), we get, using Remark 6.23, that

$$g_\beta = X_\beta(f_\beta) = ((\phi_{\alpha\beta})_*X_\alpha)((\phi_{\alpha\beta})_*f_\alpha) = (\phi_{\alpha\beta})_*(X_\alpha(f_\alpha)) = (\phi_{\alpha\beta})_*g_\alpha,$$

which shows, again by (3.4), that the g_αs are the representation of a $\mathcal{C}^{k-1}$-function g. We then set $X(f) := g$. In the $k=\infty$ case, one can define vector fields as in Section 7.1. In this case, the interpretation of the derivation $f\mapsto X(f)$ is immediate.

6.4. **Integral curves.** To a vector field X we associate the ODE

$$\dot{q} = X(q).$$

A solution, a.k.a. an integral curve, is a path $q\colon I\to M$ such that $\dot{q}(t) = X(q(t))\in T_{q(t)}M$ for all $t\in I$.

Note that, by Remark 6.22, a diffeomorphism F sends a solution γ of the ODE associated to X to a solution $\gamma \circ F$ of the ODE associated to F_*X.

Assume $k > 1$, so the vector field is continuously differentiable. The local existence and uniqueness theorem as well as the theorem on dependence on the initial values extend immediately to the case of Hausdorff $\mathcal{C}^k$-manifolds, as it enough to have them in charts. The solution is computed by solving the equation in a chart and, when we are about to leave the chart, by taking the end point of the solution as a new initial condition.

More precisely, if we want to solve the equation with initial value at some point $q \in M$, we pick a chart (U_α, ϕ_α) around q and solve the ODE for X_α in $\mathbb{R}^n$ with initial condition at $\phi_\alpha(q)$. Composing with ϕ_α^{-1} then yields a solution in U_α that we denote by γ_α. If (U_β, ϕ_β) is another chart around q, we get in principle another solution γ_β. However, by Remark 6.22, we immediately see that $\gamma_\alpha = \gamma_\beta$ in $U_\alpha \cap U_\beta$. When the solutions leave the intersection, by uniqueness of limits on a Hausdorff space, we get a unique value that shows that the solutions keep staying equal.[12] The resulting solution is simply denoted by γ with no reference to the charts.

Remark 6.25. On a non-Hausdorff manifold the above construction fails. Take the example of the line with two origins of Remark 4.8. Let X be the vector field which in each of the two charts is the constant vector 1. If we start with initial value $q \notin \{0, *\}$, then we may construct two distinct solutions: one passing through 0 but not through $*$ and another passing through $*$ but not through 0.

If the vector field vanishes at a point, then the integral curve passing through that point is constant. If the vector field does not vanish at a point, then it does not vanish on a whole neighborhood, so that through each point in that neighborhood we have a true (i.e., nonconstant) curve. The neighborhood can then be described as the collection of all these curves. By a diffeomorphism one can actually stretch these curves to straight lines, so that the neighborhood looks like an open subset of $\mathbb{R}^n$ with the integral curves being parallel to the first axis. More precisely, we have the

Proposition 6.26. *Let X be a vector field on a Hausdorff manifold M. Let $m \in M$ be a point such that $X_m \neq 0$. Then there is a chart (U, ϕ_U)*

[12] Set $T := \sup\{t : \gamma_\alpha(t) \in U_\alpha \cap U_\beta\} = \sup\{t : \gamma_\beta(t) \in U_\alpha \cap U_\beta\}$. By uniqueness of limits we have $q_1 := \lim_{t\to T} \gamma_\alpha(t) = \lim_{t \to T} \gamma_\beta(t)$. We now start again solving the equation with initial condition at q_1.

with $U \ni m$ such that $(\phi_U)_* X|_U$ is the constant vector field $(1,0,\dots,0)$. As a consequence, if γ is an integral curve of X passing through U, then $\phi_U \circ \gamma$ is of the form $\{x \in \phi_U(U) \mid x^1(t) = x_0^1 + t; x^j(t) = x_0^j, j > 1\}$ where the x_0^is are constants.

Proof. Let (V, ϕ_V) be a chart with $V \ni m$. We can assume that $\phi_V(m) = 0$ (otherwise we compose ϕ_V with the diffeomorphism of $\mathbb{R}^n$, $n = \dim M$, $x \mapsto x - \phi_V(m)$). Let $X_V := (\phi_V)_* X|_V$. We have $X_V(0) \neq 0$, so we can find a linear isomorphism A of $\mathbb{R}^n$ such that

$$AX_V(0) = (1,0,\dots,0).$$

Define $\phi'_V := A \circ \phi_V$ and $X'_V := (\phi'_V)_* X|_V$. Let $\widetilde{V}$ be an open subset of $\phi'_V(V)$, W an open subset of the intersection of $\phi'_V(V)$ with $x^1 = 0$, and $\epsilon > 0$, such that the map

$$\begin{array}{rccc} \sigma\colon & (-\epsilon,\epsilon) \times W & \to & \widetilde{V} \\ & (t, a_2, \dots, a_n) & \mapsto & \Phi_t^{X'_V}(0, a_2, \dots, a_n) \end{array}$$

is defined. The differential of σ at 0 is readily computed to be the identity map. In particular, it is invertible; hence, by the inverse function theorem, Theorem 2.24, we can find open neighborhoods $\widehat{W}$ of $(-\epsilon,\epsilon) \times W$ and $\widehat{V}$ of $\widetilde{V}$ such that the restriction of $\sigma\colon \widehat{W} \to \widehat{V}$ is a diffeomorphism. We then define $\widehat{\phi}_{V'} := \sigma^{-1} \circ \phi'_V$ and $\widehat{X} := (\widehat{\phi}_{V'})_* X = \sigma_*^{-1} X'_V$. We claim that $\widehat{X} = (1,0,\dots,0)$. In fact, using (6.6),

$$(\sigma_*(1,0,\dots,0))^i = \frac{\partial \sigma^i}{\partial t} = \frac{\partial (\Phi_t^{X'_V})^i}{\partial t} = (X'_V)^i.$$

□

6.5. **Flows.** An integral curve is called maximal if it cannot be further extended (i.e., it is not the restriction of a solution to a proper subset of its domain). On a Hausdorff manifold, through every point passes a unique maximal integral curve and to a vector field X we may then associate its flow Φ_t^X (see [5, paragraph 1.48] for more details): For $x \in M$ and t in a neighborhood of 0, $\Phi_t^X(x)$ is the unique solution at time t to the Cauchy problem with initial condition at x. Explicitly,

$$\frac{\partial}{\partial t}\Phi_t^X(x) = X(\Phi_t^X(x))$$

and $\Phi_0^X(x) = x$. We can rewrite this last condition more compactly as

$$\Phi_0^X = \mathrm{Id}_M$$

and use the existence and uniqueness theorem to show that

$$\Phi_{t+s}^X(x) = \Phi_t^X(\Phi_s^X(x)) \tag{6.7}$$

for all x and for all s and t such that the flow is defined.

By the existence and uniqueness theorem and by the theorem on dependency on the initial conditions, for each point $x \in M$ there is an open neighborhood $U \ni x$ and an $\epsilon > 0$ such that for all $t \in (-\epsilon, \epsilon)$ the map $\Phi_t^X : U \to \Phi_t^X(U)$ is defined and is a diffeomorphism.

A vector field X with the property that all its integral curves exist for all $t \in \mathbb{R}$ is called complete. If X is a complete vector field, then its flow is a diffeomorphism

$$\Phi_t^X : M \to M$$

for all $t \in \mathbb{R}$. It is often called a global flow. Equation (6.7) can then be rewritten more compactly as

$$\Phi_{t+s}^X = \Phi_t^X \circ \Phi_s^X .$$

To see whether a vector field is complete, it is enough to check that all its integral curves exist for some global time interval. In fact, we have the

Lemma 6.27. *If there is an $\epsilon > 0$ such that all the integral curves of a vector field X exist for all $t \in (-\epsilon, \epsilon)$, then they exist for all $t \in \mathbb{R}$ and hence X is complete.*

Proof. Fix $t > 0$ (we leave the analogous proof for $t < 0$ to the reader). Then there is an integer n such $t/n < \epsilon$. For each initial condition x, we can then compute the integral curve up to time t/n and call x_1 its end point. Next we can compute the integral curve with initial condition x_1 up to time t/n and call x_2 its end point, and so on. The concatenation of all these integral curves is then an integral curve extending up to time t. □

We then have the fundamental

Theorem 6.28. *Every compactly supported vector field is complete. In particular, on a compact manifold every vector field is complete.*

The support of a vector field is defined, like in the case of functions, as the closure of the set on which it does not vanish:

$$\operatorname{supp} X := \overline{\{q \in M \mid X_q \neq 0\}}.$$

A vector field X is called compactly supported if $\operatorname{supp} X$ is compact.

Proof. For every $q \in \operatorname{supp} X$ there is a neighborhood $U_q \ni q$ and an $\epsilon_q > 0$ such that all integral curves with initial condition in U_q exist for all $t \in (-\epsilon_q, \epsilon_q)$. Since $\{U_q\}_{q \in \operatorname{supp} X}$ is a covering of $\operatorname{supp} X$, and $\operatorname{supp} X$ is compact, we may find a finite collection of points $q_1, \dots, q_n$ in $\operatorname{supp} X$ such that $\{U_{q_1}, \dots, U_{q_n}\}$ is also a covering. Hence all integral curves with initial condition in $\operatorname{supp} X$ exist for all $t \in (-\epsilon, \epsilon)$ with $\epsilon = \min\{\epsilon_{q_1}, \dots, \epsilon_{q_n}\}$.

Outside of $\operatorname{supp} X$, the vector field vanishes, so the integral curves are constant and exist for all t in $\mathbb{R}$. As a consequence, all integral curves on the whole manifold exists for all $t \in (-\epsilon, \epsilon)$. We finally apply Lemma 6.27 □

Remark 6.29. For several local construction (e.g., the Lie derivative), we will pretend that the flow of a given vector field X is complete. The reason is that in these local constructions, we will always consider the neighborhood of some point q and we will tacitly replace X by ψX, where ψ is a bump function supported in a compact neighborhood of q.

7. Derivations

In this section we discuss the interpretation of tangent vectors as directions along which one can differentiate functions. To be more explicit, let $\gamma\colon I \to M$ be a differentiable curve and let f be a differentiable function on M. Then $f \circ \gamma$ is a differentiable function on I which we can differentiate. If $u \in I$ and (U, ϕ_U) is a chart with $\gamma(u) \in U$, we have

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\gamma(t))|_{t=u} = \frac{\mathrm{d}}{\mathrm{d}t} f_U(\phi_U(\gamma(t)))|_{t=u} = \sum_i v_U^i \frac{\partial f_U}{\partial x^i},$$

where f_U and v_U are the representations of f and of the tangent vector in the chart (U, ϕ_U), respectively. Notice that in this formula it is enough for f to be defined in a neighborhood of $\gamma(u)$.

This idea leads, in the case of smooth manifolds, to a definition of the tangent space where the linear structure is intrinsic and does not require choosing charts (not even at an intermediate stage). The construction is also more algebraic in nature.

The characterizing algebraic property of a derivative is the Leibniz rule for differentiating products. From the topological viewpoint, derivatives are characterized by the fact that, being defined as limits, they only see an arbitrarily small neighborhood of the point where we differentiate. The latter remark then suggests considering functions "up to a change of the definition domain," a viewpoint that turns out to be quite useful.

Let M be a $\mathcal{C}^k$-manifold, $k \geq 0$. For $q \in M$ we denote by $\mathcal{C}^k_q(M)$ the set of $\mathcal{C}^k$-functions defined in a neighborhood of q in M. Notice that by pointwise addition and multiplication of functions (on the intersection of their definition domains), $\mathcal{C}^k_q(M)$ is a commutative algebra.

Definition 7.1. We define two functions in $\mathcal{C}^k_q(M)$ to be equivalent if they coincide in a neighborhood of q.[13] An equivalence class is called a germ of $\mathcal{C}^k$-functions at q. We denote by $\mathcal{C}^k_q M$ the set of germs at q with the inherited algebra structure.

Notice that two equivalent functions have the same value at q. This defines an algebra morphism, called the evaluation at q:

$$\begin{array}{rccc} \mathrm{ev}_q\colon & \mathcal{C}^k_q M & \to & \mathbb{R} \\ & [f] & \mapsto & f(q) \end{array}$$

where on the right hand side f denotes a locally defined function in the class of $[f]$. We are now ready for the

Definition 7.2. A derivation at q in M is a linear map $D\colon \mathcal{C}^k_q M \to \mathbb{R}$ satisfying the Leibniz rule

$$D(fg) = Df\, \mathrm{ev}_q\, g + \mathrm{ev}_q f\, Dg,$$

for all $f, g \in \mathcal{C}^k_q M$. Notice that a linear combination of derivations at q is also a derivation at q. We denote by $\mathrm{Der}^k_q M$ the vector space of derivations at q in M. We wish to consider this vector space, which we have defined without using charts, as the intrinsic definition of the tangent space: we will see in Theorem 7.8 that this interpretation agrees with our previous definition but only in the case of smooth manifolds.

Remark 7.3. Notice that if U is an open neighborhood of q, regarded as a $\mathcal{C}^k$-manifold, a germ at $q \in U$ is the same as a germ at $q \in M$. So we have $\mathcal{C}^k_q U = \mathcal{C}^k_q M$. As a consequence we have

$$\boxed{\mathrm{Der}^k_q U = \mathrm{Der}^k_q M}$$

for every open neighborhood U of q in M.

The first algebraic remark is the following

Lemma 7.4. *A derivation vanishes on germs of constant functions (the germ of a constant function at q is an equivalence class containing a function that is constant in a neighborhood of q).*

Proof. Let D be a derivation at q. First consider the germ 1 (the equivalence class containing a function that is equal to 1 in a neighborhood of q). From $1 \cdot 1 = 1$, it follows that

$$D1 = D1\,1 + 1D1 = 2D1,$$

[13] More pedantically, $f \sim q$ if there is a neighborhood U of q in M contained in the definition domains of f and g such that $f_{|U} = g_{|U}$.

so $D1 = 0$. Then observe that, if f is the germ of a constant function, then $f = k1$, where k is the evaluation of f at q. Hence, by linearity, we have $Df = k\,D1 = 0$. $\square$

Remark 7.5. Notice that all the above extends to a more geneal context: one may define derivations an any algebra with a character (an algebra morphism to the ground field). The above Lemma holds in the case of algebras with one.

Let now $F\colon M \to N$ be a $\mathcal{C}^k$-morphism. Then we have an algebra morphism $F^*\colon \mathcal{C}^k_{F(q)}(N) \to \mathcal{C}^k_q(M)$, $f \mapsto f \circ F_{|_{F^{-1}(V)}}$, where V is the definition domain of f. This clearly descends to germs, so we have an algebra morphism

$$F^*\colon \mathcal{C}^k_{F(q)}N \to \mathcal{C}^k_q M,$$

which in turn induces a linear map of derivations

$$\begin{array}{rccc} \mathrm{der}^k_q F\colon & \mathrm{Der}^k_q M & \to & \mathrm{Der}^k_{F(q)} N \\ & D & \mapsto & D \circ F^* \end{array}$$

It then follows immediately that, if $G\colon N \to Z$ is also a $\mathcal{C}^k$-morphism, then

$$\mathrm{der}^k_q(G \circ F) = \mathrm{der}^k_{F(q)} G\, \mathrm{der}^k_q F.$$

This in particular implies that, if F is a $\mathcal{C}^k$-isomorphism, then $\mathrm{der}^k_q F$ is a linear isomorphism.

Let (U, ϕ_U) be a chart containing q. We then have an isomorphism $\mathrm{der}^k_q \phi_U\colon \mathrm{Der}^k_q U \to \mathrm{Der}^k_{\phi_U(q)} \phi_U(U)$. As in Remark 7.3, we have $\mathrm{Der}^k_q U = \mathrm{Der}^k_q M$ and $\mathrm{Der}^k_{\phi_U(q)} \phi_U(U) = \mathrm{Der}^k_{\phi_U(q)} \mathbb{R}^n$.[14] Hence we have an isomorphism

$$\mathrm{der}^k_q \phi_U\colon \mathrm{Der}^k_q M \xrightarrow{\sim} \mathrm{Der}^k_{\phi_U(q)} \mathbb{R}^n$$

for each chart (U, ϕ_U) containing q. It remains for us to understand derivations at a point of $\mathbb{R}^n$:

Lemma 7.6. *For every $y \in \mathbb{R}^n$, the linear map*

$$\begin{array}{rccc} A_y\colon & \mathrm{Der}^k_y \mathbb{R}^n & \to & \mathbb{R}^n \\ & D & \mapsto & \begin{pmatrix} Dx^1 \\ \vdots \\ Dx^n \end{pmatrix} \end{array}$$

is surjective for $k \geq 1$ and an isomorphism for $k = \infty$ (here $x^1, \dots, x^n$ denote the germs of the coordinate functions on $\mathbb{R}^n$).

[14] To be more precise, we regard U as a submanifold of M and $\phi_U\colon U \to \phi_U(U)$ as a diffeomorphism.

Proof. For $k \geq 1$ we may also define the linear map

$$\begin{array}{rccc} B_y\colon & \mathbb{R}^n & \to & \mathrm{Der}^k_y\mathbb{R}^n \\ & \mathbf{v} = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} & \mapsto & D_{\mathbf{v}} \end{array}$$

with

$$D_{\mathbf{v}}[f] = \sum_{i=1}^n v^i \frac{\partial f}{\partial x^i}(y),$$

where f is a representative of $[f]$. Notice that $A_y B_y = \mathrm{Id}$, which implies that A_y is surjective.

It remains to show that, for $k = \infty$, we also have $B_y A_y = \mathrm{Id}$. Let f be a representative of $[f] \in \mathcal{C}^\infty_y \mathbb{R}^n$. As a function of x, f may be Taylor-expanded around y as

$$f(x) = f(y) + \sum_{i=1}^n (x^i - y^i) \frac{\partial f}{\partial x^i}(y) + R_2(x),$$

where the rest can be written as

$$R_2(x) = \sum_{i,j=1}^n (x^i - y^i)(x^j - y^j) \int_0^1 (1-t) \frac{\partial^2 f}{\partial x^i \partial x^j}(y + t(x-y))\, dt.$$

(To prove this formula just integrate by parts.[15]) Define

$$\sigma_i(x) := \frac{\partial f}{\partial x^i}(y) + \sum_{j=1}^n (x^j - y^j) \int_0^1 (1-t) \frac{\partial^2 f}{\partial x^i \partial x^j}(y + t(x-y))\, dt,$$

so we can write

$$f(x) = f(y) + \sum_{i=1}^n (x^i - y^i)\sigma_i(x).$$

Observe that, for all i, both $x^i - y^i$ and σ_i are $\mathcal{C}^\infty$-functions;[16] the first vanishes at $x = y$, whereas for the second we have

$$\sigma_i(y) = \frac{\partial f}{\partial x^i}(y).$$

[15] Observe that we may write

$$R_2(x) = \int_0^1 (1-t) \frac{\partial^2}{\partial t^2} f(y + t(x-y))\, dt.$$

[16] Here it is crucial to work with $k = \infty$. For $k \geq 2$ finite, in general σ_i is only $\mathcal{C}^{k-2}$, and for $k = 1$ it is not even defined.

For a derivation $D \in \mathrm{Der}_y^\infty \mathbb{R}^n$, we then have, also using Lemma 7.4,

$$D[f] = \sum_{i=1}^n Dx^i \frac{\partial f}{\partial x^i}(y) = B_y A_y(D)[f],$$

which completes the proof. $\square$

From now on, we simply write Der_q and der_q instead of Der_q^∞ and der_q^∞.

Corollary 7.7. *For every q in a smooth manifold, we have*

$$\boxed{\dim \mathrm{Der}_q M = \dim M}$$

We finally want to compare the construction in terms of derivations with the one in terms of equivalence classes of coordinatized tangent vectors.

Theorem 7.8. *Let M be a smooth manifold, $q \in M$, and (U, ϕ_U) a chart containing q. Then the isomorphism*

$$\tau_{q,U} := (\mathrm{der}_q \phi_U)^{-1} A_{\phi(q)}^{-1} \Phi_{q,U} \colon T_q M \xrightarrow{\sim} \mathrm{Der}_q M$$

does not depend on the choice of chart. We will denote this canonical isomorphism simply by τ_q.

If $F\colon M \to N$ is a smooth map, we have $\mathrm{d}_q F = \tau_{F(q)}^{-1} \,\mathrm{der}_q F\, \tau_q$.

Proof. Explicitly we have,

$$(\tau_{q,U}[(U, \phi_U, v)])[f] = \sum_{i=1}^n v^i \frac{\partial (f \circ \phi_U^{-1})}{\partial x^i}(\phi_U(q)),$$

for every representative f of $[f] \in \mathcal{C}_q^\infty M$. We then have, by the chain rule,

$$\begin{aligned}(\tau_{q,V}[(U, \phi_U, v)])[f] &= (\tau_{q,V}[(V, \phi_V, \mathrm{d}_{\phi_U(q)}\phi_{U,V} v)])[f] = \\ &= \sum_{i,j=1}^n \frac{\partial \phi_{U,V}^i}{\partial x^j}(\phi_U(q))\, v^j\, \frac{\partial (f \circ \phi_V^{-1})}{\partial x^i}(\phi_V(q)) = \\ &= \sum_{i=1}^n v^i \frac{\partial (f \circ \phi_U^{-1})}{\partial x^i}(\phi_U(q)) = (\tau_{q,U}[(U, \phi_U, v)])[f].\end{aligned}$$

The last statement of the Theorem also easily follows from the chain rule in differentiating $f \circ F$, $f \in [f] \in \mathcal{C}_{F(q)}^\infty N$. $\square$

7.1. **Vector fields as derivations.** We now want to show that vector fields on a smooth Hausdorff manifold are the same as derivations on its algebra of functions.

Definition 7.9. A derivation on the algebra of functions $\mathcal{C}^\infty(M)$ of a smooth manifold M is an $\mathbb{R}$-linear map $D\colon \mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M)$ that satisfies the Leibniz rule

$$D(fg) = Df\, g + f\, Dg.$$

Notice that a linear combination of derivations is also a derivation. We denote by $\mathrm{Der}(M)$ the $\mathcal{C}^\infty(M)$-module of derivations on $\mathcal{C}^\infty(M)$.

Remark 7.10. This construction can be generalized to any algebra A. By $\mathrm{Der}(A)$ one then denotes the algebra of derivations on A. In the case $A = \mathcal{C}^\infty(M)$, $\mathrm{Der}(M)$ may be used as a shorthand notation for $\mathrm{Der}(\mathcal{C}^\infty(M))$.

Remark 7.11. On a $\mathcal{C}^k$-manifold M, $k \geq 1$, one can define derivations as linear maps $\mathcal{C}^k(M) \to \mathcal{C}^{k-1}(M)$ that satisfy the Leibniz rule.

The first remark is that derivations, like derivatives, are insensitive to changing functions outside of a neighborhood:

Lemma 7.12. *Let M be a Hausdorff $\mathcal{C}^k$-manifold, $k \geq 1$. Let D be a derivation and f a function that vanishes on some open subset U. Then $Df(q) = 0$ for all $q \in U$.*

Proof. Let ψ be a bump function as in Lemma 5.1. Then $f = (1-\psi)f$. In fact, ψ vanishes outside of U, whereas f vanishes inside U. We then have $Df = D(1-\psi)\, f + (1-\psi)Df$. Since $f(q) = 0 = 1 - \psi(q)$, we get $Df(q) = 0$. □

We then want to connect derivations with derivations at a point q. Notice that, for every $\mathcal{C}^k$-manifold, $k \geq 0$, we have a linear map

$$\gamma_q\colon \mathcal{C}^k(M) \to \mathcal{C}^k_q M$$

that associates to a function its germ at q.

Lemma 7.13. *Let M be a Hausdorff $\mathcal{C}^k$-manifold. Then, for every $q \in M$, γ_q is surjective.*

Proof. Let $[f] \in \mathcal{C}^k_q M$. Let $g \in \mathcal{C}^k(W)$ be a representative of $[f]$ in some open neighborhood W of q. Then pick an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$, and let α be an index such that $U_\alpha \ni q$. Let U be an open neighborhood of q strictly contained in $W \cap U_\alpha$ (simply take the preimage by ϕ_α of an open ball centered at $\phi_\alpha(q)$ strictly contained in $\phi_\alpha(W \cap U_\alpha)$) and let ψ be a bump function as in Lemma 5.1. Let $h := g\psi \in \mathcal{C}^k(U)$. Then

$[h] = [f]$. Since g is identically equal to zero in the complement of U inside U_α, we can extend it by zero to get a $\mathcal{C}^k$-function on the whole of M. □

Theorem 7.14. *If M is a Hausdorff smooth manifold, we have a canonical $\mathcal{C}^\infty(M)$-linear isomorphism*

$$\tau\colon \mathfrak{X}(M) \to \mathrm{Der}(M),$$

where $\mathfrak{X}(M)$ is the $\mathcal{C}^\infty(M)$-module of vector fields on M.

Proof. If X is a vector field and f is a function, we define $((\tau(X))f)(q) := (\tau_q X(q))\gamma_q f$. It is readily verified that $\tau(X)$ is a derivation. It is also clear that τ is $\mathcal{C}^\infty(M)$-linear and injective. We only have to show that it is surjective.

If D is a derivation and $[f] \in \mathcal{C}^\infty_q$, we define $D_q[f] := (Df)(q)$ for any $f \in \gamma_q^{-1}([f])$. (By Lemma 7.13 we know that γ_q is surjective.) By Lemma 7.12 this is readily seen to be independent of the choice of f and to be a derivation at q. We then define $X_q := \tau_q^{-1}(D_q)$, which is readily seen to depend smoothly on q. Hence we have found an inverse map to τ. □

Remark 7.15. Because of the canonical identification proved above, from now on we will use interchangeably T_qM and $\mathrm{Der}_q M$, $\mathrm{d}_q F$ and $\mathrm{der}_q F$, $\mathfrak{X}(M)$ and $\mathrm{Der}(C^\infty(M))$. (We will also always assume M to be Hausdorff.)

Let us now concentrate on the case where $M = U$ is an open subset of $\mathbb{R}^n$ (this is also the case of the representation in a chart). A vector field X on U may be regarded as a map $q \mapsto (X^1(q), \dots, X^n(q))$ or as a derivation that we write as

$$X = \sum_{i=1}^n X^i \frac{\partial}{\partial x^i}.$$

This useful notation also has a deeper meaning: $\left(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}\right)$ is a $C^\infty(U)$-linearly independent system of generators of $\mathfrak{X}(U) = \mathrm{Der}(U)$ as a module over $C^\infty(U)$.

Remark 7.16. A useful, quite common notation consists in defining

$$\partial_i := \frac{\partial}{\partial x^i}.$$

With this notation, a vector field on U reads $X = \sum_{i=1}^n X^i \partial_i$. This notation is neater and creates no ambiguity when a single set of coordinates is used. Notice that, if f is a function, we may also write $\partial_i f$ instead of the more cumbersome $\frac{\partial f}{\partial x^i}$.

7.2. **The Lie bracket.** Derivations are in particular endomorphisms and endomorphisms may be composed. However, in general, the composition of two derivations is not a derivation. In fact,

$$XY(fg) = X(Y(f)g + fY(g)) = \\ = XY(f)g + Y(f)X(g) + X(f)Y(g) + fXY(g).$$

On the other hand, we can get rid of the unwanted terms $Y(f)X(g)$ and $X(f)Y(g)$ by skew-symmetrizing. This shows that

$$[X, Y] := XY - YX \tag{7.1}$$

is again a derivation. The operation $[\ ,\ ]$ is called the Lie bracket. Note that

$$[X, [Y, Z]] = XYZ - XZY - YZX + ZYX.$$

This shows that

$$[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0$$

for all vector fields X, Y, Z.

This is just an example of a more general setting:

Definition 7.17. A Lie algebra is a vector space V endowed with a bilinear map $[\ ,\ ]\colon V \times V \to V$, which is skew symmetric, i.e.,

$$[a, b] = -[b, a]\ \forall a, b \in V,$$

and satifies the Jacobi identity

$$[a, [b, c]] = [[a, b], c] + [b, [a, c]],\ \forall a, b, c \in V.$$

The operation is usually called a Lie bracket.

Remark 7.18. Using skew-symmetry, the Jacobi identity may equivalently be written

$$[a, [b, c]] + [c, [a, b]] + [b, [c, a]] = 0,\ \forall a, b, c \in V.$$

Example 7.19. $V := \mathrm{Mat}(n \times n, \mathbb{R})$ with $[A, B] := AB - BA$ is a Lie algebra, where AB denotes matrix multiplication. More generally, $V := \mathrm{End}(W)$, W some vector space, $[A, B] := AB - BA$ is a Lie algebra, where AB denotes composition of endomorphism.

Definition 7.20. A subspace W of a Lie algebra $(V, [\ ,\ ])$ is called a Lie subalgebra if $[a, b] \in W\ \forall a, b \in W$. Notice that W is a Lie algebra itself with the restriction of the Lie bracket of V.

Example 7.21. $\mathrm{Der}(M)$ is a Lie subalgebra of $\mathrm{End}_{\mathbb{R}}(C^\infty(M))$.[17]

[17] More generally, if A is an algebra, i.e., a vector space with a bilinear operation, we may still define derivations and $\mathrm{Der}(A)$ is a Lie subalgebra of $\mathrm{End}(A)$.

Remark 7.22. Notice that $\mathfrak{X}(M)$ is also a module over $\mathcal{C}^\infty(M)$; however, the Lie bracket is not $\mathcal{C}^\infty(M)$-bilinear. Instead, as follows immediately from (7.1), if f is a function and X, Y are vector fields, one has

$$[X, fY] = f[X, Y] + X(f)Y, \qquad [fX, Y] = f[X, Y] - Y(f)X. \tag{7.2}$$

If we work locally, i.e., for $M = U$ an open subset of $\mathbb{R}^n$, we can write the Lie bracket of vector fields explicitly as follows (we use the notation of Remark 7.16): let $X = \sum_{i=1}^n X^i \partial_i$ and $Y = \sum_{i=1}^n Y^i \partial_i$. Then $[X, Y] = \sum_{i=1}^n [X, Y]^i \partial_i$ with

$$[X, Y]^i = \sum_{j=1}^n \left(X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j} \right). \tag{7.3}$$

If X and Y are $\mathcal{C}^k$-vector fields with $1 \le k < \infty$ we can still define their Lie bracket by this formula, but the results will be a $\mathcal{C}^{k-1}$-vector field.

Remark 7.23. If X and Y are vector fields on a smooth manifold M, their representations X_α and Y_α are vector fields on the open subset $\phi_\alpha(U_\alpha)$ of $\mathbb{R}^n$. The representation $[X, Y]_\alpha$ of $[X, Y]$ is then given by

$$[X, Y]^i_\alpha = \sum_{j=1}^n \left(X^j_\alpha \frac{\partial Y^i_\alpha}{\partial x^j} - Y^j_\alpha \frac{\partial X^i_\alpha}{\partial x^j} \right).$$

This is in particular shows that the $[X, Y]_\alpha$s transform according to (6.2). Notice that $\sum_{i,j=1}^n X^j_\alpha \frac{\partial Y^i_\alpha}{\partial x^j} \frac{\partial}{\partial x^i}$ is also a vector field on $\phi_\alpha(U_\alpha)$ for each α, but in general these vector fields do not transform according to (6.2), so they do not define a vector field on M.

Remark 7.24. On a $\mathcal{C}^k$-manifold, $1 \le k < \infty$, we can define the Lie bracket of vector fields by the local formula. The result will be a globally defined $\mathcal{C}^{k-1}$-vector field on M. This can be checked by an explicit computation.

The Lie bracket of vector fields has several important applications. It also has a geometric meaning, to which we will return in Section 7.4.

7.3. **The push-forward of derivations.** We will now define the push-forward F_* of derivations under a diffeomorphism F. We use the same notation as for the push-forward of vector fields introduced in Remark 6.21, as we will show that the two notions coincide.

Let M and N be smooth manifolds and $F\colon M \to N$ a diffeomorphism. Recall from subsection 3.1 that by $F^*\colon \mathcal{C}^\infty(N) \to \mathcal{C}^\infty(M)$,

$g \mapsto F^*g := g \circ F$ we denote the pullback of functions.[18] Also recall that F^* is an $\mathbb{R}$-linear map and that $F^*(fg) = F^*f\, F^*g$, $\forall f, g \in \mathcal{C}^\infty(N)$. If X is a vector field on M, regarded as a derivation, we define its push-forward F_*X as a composition of endomorphisms of $C^\infty(M)$:

$$F_*X := (F^*)^{-1} X F^*.$$

Namely, if g is a function we have

$$F_*X(g) := (F^*)^{-1}(X(F^*g)),$$

If $G\colon N \to Z$ is also a diffeomorphism, then we clearly have $(G \circ F)_* = G_*F_*$.

Lemma 7.25. *The push-forward maps vector fields to vector fields.*

Proof. We just compute

$$\begin{aligned} F_*X(fg) = (F^*)^{-1}X(F^*(fg)) = (F^*)^{-1}X(F^*f F^*g) = \\ = (F^*)^{-1}(X(F^*f)F^*g + F^*fX(F^*g)) = \\ = (F^*)^{-1}(X(F^*f))\, g + f\, (F^*)^{-1}(X(F^*g)) = F_*X(f)g + fF_*X(g). \end{aligned}$$

□

It is also clear that

$$F_*\colon \mathfrak{X}(M) \to \mathfrak{X}(N)$$

is an $\mathbb{R}$-linear map. By (7.1) we also see that

$$F_*[X, Y] = [F_*X, F_*Y]$$

for all $X, Y \in \mathfrak{X}(M)$; one says that F_* is a morphism of Lie algebras. Moreover, for $f \in \mathcal{C}^\infty(M)$, we have

$$F_*(fX)(g) = (F^*)^{-1}(fX(F^*g)) = (F^*)^{-1}f\, F_*X(g).$$

Using the push-forward F_* of functions, defined in Remark 3.13 as $(F^*)^{-1}$, we then have the nicer looking formula

$$F_*(fX) = F_*f\, F_*X.$$

We can summarize:

[18]The pullback is defined for any smooth map F, but for the following consderations we need a diffeomorphism.

Theorem 7.26. *Let $F\colon M \to N$ be a diffeomorphism. Then the push-forward F_* is an $\mathbb{R}$-linear map from $\mathcal{C}^\infty(M)$ to $\mathcal{C}^\infty(N)$ and from $\mathfrak{X}(M)$ to $\mathfrak{X}(N)$ such that*

$$\begin{aligned} F_*(fg) &= F_*(f)F_*(g),\\ F_*(fX) &= F_*(f)F_*(X),\\ F_*[X,Y] &= [F_*X, F_*Y], \end{aligned}$$

$\forall f, g \in \mathcal{C}^\infty(M)$ and $\forall X, Y \in \mathfrak{X}(M)$. If $G\colon N \to Z$ is also a diffeomorphism, then

$$(G \circ F)_* = G_* F_*.$$

The push-forward of vector fields regarded as derivations agrees with the definition we gave in Remark 6.21:

Proposition 7.27. *Let $F\colon M \to N$ be a diffeomorphism and X a vector field on M. Then*

$$(F_*X)_y = \mathrm{d}_{F^{-1}(y)}F\, X_{F^{-1}(y)}, \quad \forall y \in N.$$

Equivalently,

$$(F_*X)_{F(q)} = \mathrm{d}_qF\, X_q, \quad \forall q \in M. \tag{7.4}$$

Proof. We use the notations of subsection 7.1. Let $[f] \in \mathcal{C}_{F(q)}N$ and $f \in \gamma_{F(q)}^{-1}[f] \subset \mathcal{C}^\infty(N)$. Since $\gamma_q(f \circ F) = F^*[f]$, we get

$$\begin{aligned} (F_*X)_{F(q)}[f] = (F_*X(f))(F(q)) = (X(f \circ F))(q) =\\ = X_q(F^*[f]) = (\mathrm{der}_qF X_q)[f]. \end{aligned}$$

Finally, we recall from Theorem 7.8 on page 36 that, up to the isomorphism $\tau_q\colon T_qM \xrightarrow{\sim} \mathrm{Der}_qM$, der_qF and d_qF are the same thing. $\square$

We finish with the following important

Proposition 7.28. *Let Φ_t^X denote the flow of a vector field X. Then, for every diffeomorphism $F\colon M \to N$ and for every vector field X on M, we have*

$$F \circ \Phi_t^X \circ F^{-1} = \Phi_t^{F_*X}.$$

Proof. Let $\Psi_t := F \circ \Phi_t^X \circ F^{-1}$. For every $y \in N$ we have $\Psi_0(y) = y$ and

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\Psi_t(y) = \mathrm{d}_{\Phi_t^X(F^{-1}(y))}F \frac{\mathrm{d}}{\mathrm{d}t}\Phi_t^X(F^{-1}(y)) = \mathrm{d}_{\Phi_t^X(F^{-1}(y))}F X_{\Phi_t^X(F^{-1}(y))} =\\ = (F_*X)_{F(\Phi_t^X(F^{-1}(y)))} = (F_*X)_{\Psi_t(y)}, \end{aligned}$$

which shows that Ψ_t is the flow of F_*X. $\square$

7.4. **The Lie derivative.** Vector fields may be regarded as directions along which we can differentiate functions. The expression $X(f)$ corresponds precisely to this. If we regard X_q as an element of T_qM, then we also have

$$X(f)(q) = \mathrm{d}_q f X_q.$$

More geometrically, we can define this differentiation as the change of the function along the flow of the vector field. This viewpoint allows extending differentiation to vector fields as well (and also more generally to other objects, like densities or tensor fields, that we will consider in the following).

Let X be a vector field on M and let Φ_t^X denote its flow (we assume here that M is Hausdorff). The Lie derivative of a function f along the vector field X is defined as

$$\mathsf{L}_X f := \frac{\partial}{\partial t}\Big|_{t=0} f \circ \Phi_t^X$$

or, equivalently,

$$\mathsf{L}_X f = \frac{\partial}{\partial t}\Big|_{t=0} (\Phi_{-t}^X)_* f = \lim_{h\to 0} \frac{(\Phi_{-h}^X)_* f - f}{h} = \lim_{h\to 0} \frac{f - (\Phi_h^X)_* f}{h}.$$

A simple computation shows that

$$\mathsf{L}_X f = X(f). \tag{7.5}$$

Namely,

$$\mathsf{L}_X f(q) = (\frac{\partial}{\partial t}\Big|_{t=0} (\Phi_{-t}^X)_* f)(q) = \frac{\partial}{\partial t}\Big|_{t=0} f(\Phi_t^X(q)) = \mathrm{d}_q f X_q.$$

We now extend this definition to vector fields. Namely, the Lie derivative of a vector field Y along the vector field X is defined as

$$\mathsf{L}_X Y := \frac{\partial}{\partial t}\Big|_{t=0} (\Phi_{-t}^X)_* Y = \lim_{h\to 0} \frac{(\Phi_{-h}^X)_* Y - Y}{h} = \lim_{h\to 0} \frac{Y - (\Phi_h^X)_* Y}{h}.$$

Lemma 7.29. *For all $X, Y \in \mathfrak{X}(M)$ we have*

$$\mathsf{L}_X Y = [X, Y]. \tag{7.6}$$

This yields a geometric interpretation of the Lie bracket.

Proof. By definition of the push-forward of a vector field, using push-forward of functions instead of pullback, we have

$$((\Phi_{-t}^X)_* Y)(g) = (\Phi_{-t}^X)_*(Y((\Phi_{-t}^X)_*^{-1} g)).$$

Differentiation at $t = 0$ then yields

$$(\mathsf{L}_X Y)(g) = L_X(Y(g)) - Y(L_X g) = X(Y(g)) - Y(X(g)),$$

where we have also used (7.5). As this holds for all $g \in \mathcal{C}^\infty(M)$, the Lemma is proved. $\square$

Notice that by its definition, or by the explicit formulae (7.5) and (7.6), the Lie derivative L_X is $\mathbb{R}$-linear. Morever, if f is a function and Ξ is a function or a vector field, we have

$$\mathsf{L}_X(f\Xi) = \mathsf{L}_X f\,\Xi + f\,\mathsf{L}_X\Xi.$$

Remark 7.30. The Lie derivative is a beautiful construction that gives a geometrical (or dynamical) meaning to algebraic concepts like the application of a vector field to a function or the Lie bracket of two vector fields. To make full sense of it, it seems however that we should have assumed the flow of the vector field X to be global. The reason why we do not need it is that in order to compute $\mathsf{L}_X\Xi$ at some point $q \in M$ we can replace X by the compactly supported vector field ψX, where ψ is a bump function supported around q. Note that a different bump function $\widetilde{\psi}$ will yield the same result. In fact,

$$\mathsf{L}_{\widetilde{\psi}X}\Xi - \mathsf{L}_{\psi X}\Xi = \lim_{h\to 0} \frac{(\Phi^{\widetilde{\psi}X}_{-h})_* Y - (\Phi^{\psi X}_{-h})_* Y}{h}$$

which vanishes on the neighborhood V of q where ψ and $\widetilde{\psi}$ coincide. In the rest of this section, we will always assume that the flow of X is global or it has been made global by this construction.

The composition property of flows ($\Phi^X_{t+s} = \Phi^X_t \circ \Phi^X_s$) implies the following useful

Lemma 7.31. *Let Ξ be a function or a vector field. Then*

$$\left.\frac{\partial}{\partial t}\right|_{t=s} (\Phi^X_{-t})_*\Xi = (\Phi^X_{-s})_*\mathsf{L}_X\Xi = \mathsf{L}_X(\Phi^X_{-s})_*\Xi$$

for all s for which the flow is defined.

Proof.

$$\left.\frac{\partial}{\partial t}\right|_{t=s} (\Phi^X_{-t})_*\Xi = \lim_{h\to 0} \frac{(\Phi^X_{-s-h})_*\Xi - (\Phi^X_{-s})_*\Xi}{h} = (\Phi^X_{-s})_* \lim_{h\to 0} \frac{(\Phi^X_{-h})_*\Xi - \Xi}{h}.$$

The second equality follows similarly from

$$\lim_{h\to 0} \frac{(\Phi^X_{-s-h})_*\Xi - (\Phi^X_{-s})_*\Xi}{h} = \lim_{h\to 0} \frac{(\Phi^X_{-h})_*(\Phi^X_{-s})_*\Xi - (\Phi^X_{-s})_*\Xi}{h}.$$

$\square$

Corollary 7.32. *Let X a be vector field. A function f is preserved by the flow of X, i.e., $(\Phi^X_t)_* f = f$ $\forall t$, if and only if $X(f) = 0$. A vector field Y is preserved by the flow of X, i.e., $(\Phi^X_t)_* Y = Y$ $\forall t$, if* *and only if* $[X, Y] = 0$. *In particular,* X *is preserved by its own flow:* $(\Phi_t^X)_* X = X \ \forall t$.

Notice that as the Lie bracket is skew-symmetric, the condition for vector fields is symmetric: namely, Y is preserved by the flow of X if and only if X is preserved by the flow of Y. An even more symmetric statement is provided by the next

Proposition 7.33. *The flows of two vector fields commute if and only if the two vector fields Lie commute. In formulae:*

$$\Phi_s^Y \circ \Phi_t^X = \Phi_t^X \circ \Phi_s^Y \ \forall s, t \text{ for which the flows are defined} \Leftrightarrow [X, Y] = 0.$$

Proof. By Proposition 7.28, we have

$$(\Phi_s^Y)^{-1} \circ \Phi_t^X \circ \Phi_s^Y = \Phi_t^{(\Phi_{-s}^Y)_* X}$$

If $[X, Y] = 0$, then $(\Phi_{-s}^Y)_* X = X$ and the flows commute. If on the other hand the flows commute, then we get that $\Phi_t^{(\Phi_{-s}^Y)_* X} = \Phi_t^X \ \forall t, s$; by deriving at $t = 0$, we then get $(\Phi_{-s}^Y)_* X = X \ \forall s$, which implies $[X, Y] = 0$. □

There is one more way of characterizing and computing the Lie bracket:

Lemma 7.34. *On an open subset of* $\mathbb{R}^n$ *we have*

$$\left.\frac{\partial^2}{\partial s \partial t}\right|_{s=t=0} (\Phi_t^X)^{-1} \circ (\Phi_s^Y)^{-1} \circ \Phi_t^X \circ \Phi_s^Y = [Y, X].$$

Proof. Let $\Phi_{s,t} := (\Phi_t^X)^{-1} \circ (\Phi_s^Y)^{-1} \circ \Phi_t^X \circ \Phi_s^Y$. By Proposition 7.28, we have

$$\Phi_{s,t} = (\Phi_t^X)^{-1} \circ \Phi_t^{(\Phi_{-s}^Y)_* X}.$$

Hence

$$\left.\frac{\partial}{\partial t}\right|_{t=0} \Phi_{s,t} = -X + (\Phi_{-s}^Y)_* X.$$

Thus,

$$\left.\frac{\partial^2}{\partial s \partial t}\right|_{s=t=0} \Phi_{s,t} = \mathsf{L}_Y X.$$

□

Finally, we have the following important property:

Lemma 7.35. *Let* Ξ *be a function or a vector field and let* X *and* Y *be vector fields. Then*

$$\mathsf{L}_{X+Y} \Xi = \mathsf{L}_X \Xi + \mathsf{L}_Y \Xi,$$
$$\mathsf{L}_X \mathsf{L}_Y \Xi - \mathsf{L}_Y \mathsf{L}_X \Xi = \mathsf{L}_{[X,Y]} \Xi.$$

These formulae say that the map $X \mapsto \mathsf{L}_X$ from vector fields to linear operators (on the vector space of functions or on that of vector fields) is $\mathbb{R}$-linear and that

$$[\mathsf{L}_X, \mathsf{L}_y] = \mathsf{L}_{[X,Y]},$$

where [,] on the left hand side denotes the commutator of linear operators.

One easy way to prove them is by the the explicit formulae (7.5) and (7.6). The first identity is then obvious, whereas the second is just the definition of the Lie bracket if Ξ is a function and the Jacobi identity if Ξ is a vector field. If we are in the image of a chart (i.e., if we work on an open subset of $\mathbb{R}^n$), we also have a proof based directly on the definition of Lie derivative; this is interesting, for this proof will apply to other cases as well.

Proof. Let $\Psi_t := \Phi_t^{X+Y} \circ \Phi_{-t}^Y \circ \Phi_{-t}^X$. We have $\Psi_0 = \mathrm{Id}$ and $\frac{\partial \Psi_t}{\partial t}|_{t=0} = 0$. Hence

$$0 = \frac{\partial}{\partial t}\Big|_{t=0} (\Psi_t)_*\Xi = -\mathsf{L}_{X+Y}\Xi + \mathsf{L}_X\Xi + \mathsf{L}_Y\Xi.$$

For the second identity consider again $\Phi_{s,t} := (\Phi_t^X)^{-1} \circ (\Phi_s^Y)^{-1} \circ \Phi_t^X \circ \Phi_s^Y$. We have

$$\frac{\partial}{\partial t}\Big|_{t=0} (\Phi_{s,t})_*\Xi = \mathsf{L}_X\Xi - (\Phi_{-s}^Y)_*\mathsf{L}_X(\Phi_s^Y)_*\Xi,$$

so

$$\frac{\partial^2}{\partial s \partial t}\Big|_{s=t=0} (\Phi_{s,t})_*\Xi = -\mathsf{L}_Y\mathsf{L}_X\Xi + \mathsf{L}_X\mathsf{L}_Y\Xi.$$

Again, by Proposition 7.28, we have $\Phi_{s,t} = (\Phi_t^X)^{-1} \circ \Phi_t^{(\Phi_{-s}^Y)_*X}$, so

$$\frac{\partial}{\partial t}\Big|_{t=0} (\Phi_{s,t})_*\Xi = \mathsf{L}_X\Xi - \mathsf{L}_{(\Phi_{-s}^Y)_*X}\Xi.$$

By definition of Lie derivative, $(\Phi_{-s}^Y)_*X = X + s\,\mathsf{L}_YX + O(s^2)$. By the just proved first identity, we then have

$$\mathsf{L}_{(\Phi_{-s}^Y)_*X}\Xi = \mathsf{L}_X\Xi + s\,\mathsf{L}_{\mathsf{L}_YX}\Xi + O(s^2).$$

Finally,

$$\frac{\partial^2}{\partial s \partial t}\Big|_{s=t=0} (\Phi_{s,t})_*\Xi = -\mathsf{L}_{\mathsf{L}_YX}\Xi = \mathsf{L}_{[X,Y]}\Xi.$$

□

7.5. **Plane distributions.** In this section we want to generalize the results of Section 6.4 to the case when we want to integrate several ODEs simultaneously. For simplicity we focus on the smooth case only. The main goal will be to prove the Frobenius theorem, which has several applications. (We will see some in Sections 9.8.2, 10.3 and 10.4.)

Definition 7.36. A k-plane distribution D, or simply a k-distribution or just a distribution,[19] on a smooth n-dimensional manifold M is a collection $\{D_q\}_{q\in M}$ of linear k-dimensional subspaces $D_q \in T_qM$ for all $q \in M$. (Of course we assume $k \leq n$.) The number k is also called the rank of the distribution.

We are interested in distributions that vary smoothly over M. We present here a preliminary definition that is enough for the applications in this section; we will return to a nicer, equivalent characterization later (see Corollary 8.15).

Definition 7.37. A k-distribution D on M is called smooth if every q in M possesses an open neighborhood U and smooth vector fields $X_1, \dots, X_k$ defined on U such that

$$D_x = \operatorname{span}\{(X_1)_x, \dots, (X_k)_x\}$$

for all $x \in U$. The vector fields $X_1, \dots, X_k$ are also called (local) generators for D on U.

Remark 7.38. If one can take U to be the whole of M, one speaks of global generators. In the definition we require the existence of local generators only, as several interesting distributions do not admit global generators. See the examples below.

A vector field X on M is said to be tangent to a distribution D if $X_q \in D_q$ for all $q \in M$. Linear combinations of vector fields tangent to a given distribution are also tangent to it. We denote by $\Gamma(D)$ the $\mathbb{R}$-vector space and $C^\infty(M)$-module of vector fields tangent to D. Note that $\Gamma(D)$ is a subspace of $\mathfrak{X}(M)$.

Definition 7.39. A smooth distribution D on M is called involutive if $\Gamma(D)$ is a Lie subalgebra of $\mathfrak{X}(M)$; i.e., when $[X, Y] \in \Gamma(D)$ for all X and Y in $\Gamma(D)$.

Remark 7.40. Note that a distribution generated by the vector fields $X_1, \dots, X_k$ is involutive if and only if, for all i and j, $[X_i, X_j]$ is a linear

[19] Distributions in this sense have nothing to do with distributions introduced in analysis as continuous linear functionals on spaces of test functions.

combination, over the ring of functions, of the generators. The only-if side follows directly from the the definition. For the if-implication, observe that a vector field tangent to the distribution is necessarily a linear combination of the generators. Moreover, by (7.2), we have

$$\left[\sum_i f_i X_i, \sum_j g_j X_j\right] = \sum_{ij} \left((f_i X_i(g_j) - g_i X_i(f_j))X_j + f_i g_j [X_i, X_j]\right).$$

The first term in the sum is explicitly tangent to the distribution; the second is so by the assumption.

Remark 7.41. The previous remark implies that a smooth distribution of rank 1 is always involutive. In fact, locally it is generated by a single vector field, say X, and by skew-symmetry of the Lie bracket we have $[X, X] = 0$.

Example 7.42. Let X be a nowhere vanishing vector field on M. Let $D_q := \operatorname{span} X_q = \{\lambda X_q,\ \lambda \in \mathbb{R}\}$. Then D is a smooth 1-distribution. It is also involutive by Remark 7.41. Note that different vector fields may generate the same a 1-distribution.

Example 7.43. $D = 0$ (i.e., $D_q = \{0\}$ for all q) is an involutive 0-distribution, and $D = TM$ (i.e., $D_q = T_qM$ for all q) is an involutive n-distribution, $n = \dim M$. Note that $D = TM$ may not admit global generators. In the terminology to be introduced in Section 8.1, the distribution TM has global generators if and only if M is parallelizable. In Lemma 8.32 we will see that, e.g., $M = S^2$ is not parallelizable.

Example 7.44. Let $M = \mathbb{R}^3 \setminus \{0\}$ and D the distribution of planes othogonal to the radial direction; i.e., $D_{\mathbf{x}} = \{\mathbf{v} \in \mathbb{R}^3 \mid \mathbf{v} \cdot \mathbf{x} = 0\}$. In other words, $D_{\mathbf{x}}$ consists of the vectors tangent at $\mathbf{x}$ to the sphere of radius $||\mathbf{x}||$. This shows that this distribution has rank 2, is smooth and is involutive. (Another way to see that it is involutive consists in observing that X is tangent to D if and only if $X(r) = 0$, where r is the function $||\mathbf{x}||$. If X and Y are both tangent, then $[X, Y](r) = X(Y(r)) - Y(X(r)) = 0$ as well.) This distribution has no global generators, as TS^2 does not have them either.

Example 7.45. On $U = \{(x, y, z) \in \mathbb{R}^3 \mid z \neq 0\}$ consider

$$X = y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y},$$
$$Y = z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}.$$

Let $D_p := \text{span}\{X_p, Y_p\}$. Then D is a smooth 2-distribution. It is involutive by Remark 7.40. In fact,

$$Z := [Y, X] = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}$$

can be written as $Z = -\frac{x}{z}X - \frac{y}{z}Y$. (Geometrically observe that the flows of X, Y, and Z are rotations around the x-, y- and z-axes, respectively.)

Example 7.46. On $\mathbb{R}^3$ consider

$$X = \frac{\partial}{\partial x}, \qquad Y = \frac{\partial}{\partial y} + x\frac{\partial}{\partial z}.$$

The distribution defined by $\text{span}\{X, Y\}$ is smooth and of rank 2. It is however not involutive since $[X, Y] = \frac{\partial}{\partial z}$ cannot be written as a $C^\infty(\mathbb{R}^3)$-linear combination of X and Y.

We can generalize the notion of push-forward to distributions mimicking the definition in Remark 6.21.

Definition 7.47. Let D be a distribution on M and let $F\colon M \to N$ be a diffeomorphism. We define the push-forward F_*D of D by

$$(F_*D)_y := \mathrm{d}_{F^{-1}(y)} D_{F^{-1}(y)}$$

for all $y \in N$.

Note that the push-forward of a smooth distribution is also smooth, and the push-forward of an involutive distribution is also involutive. We now come to the generalization of the notion of integral curve.

Definition 7.48. An immersion $\psi\colon N \to M$ with N connected is called an integral manifold for a distribution D on M if

$$\mathrm{d}_n\psi(T_nN) = D_{\psi(n)} \qquad \text{for all } n \in N.$$

An integral manifold that is not a proper restriction of an integral manifold is called maximal.

If ψ is an embedding (which is not much of a restriction in view of Proposition 6.4), restricting $\psi\colon N \to \psi(N)$ to its image allows rewriting the above condition as

$$\psi_*(TN) = D|_{\psi(N)},$$

where the fact that D may be restricted to $\psi(N)$ (i.e., $D_q \in T_q\psi(N)$ for all $q \in \psi(N)$) is part of the condition.

Definition 7.49. A smooth distribution D on M is called integrable if for every q in M there is an integral manifold for D passing through q.

Example 7.50. Example 7.43 yields an integrable distribution with (M, Id) as maximal integral manifold. Another example of integrable distribution is given by Example 7.42, where the (maximal) integral manifolds are the (maximal) integral curves of the vector field. Example 7.45 also yields an integrable distribution where (the images of) the maximal integral manifolds are the connected components of the intersections of U with the spheres centered at the origin.

The fact that these examples of integrable distributions are also involutive is not by chance. In fact, we have the

Lemma 7.51. *If D is integrable, then D is involutive.*

Proof. For each $q \in M$, we can find an integral manifold $\psi\colon N \to M$ with $q \in \psi(N)$. If X and Y are tangent to D, in a neighborhood of q in $\psi(N)$ we can write them as push-forwards of vector fields $\widetilde{X}$ and $\widetilde{Y}$ on N. Since the push-forward preserves Lie bracket and TN is involutive, we see that in this neighborhood $Z := [X, Y]$ is the push-forward of $[\widetilde{X}, \widetilde{Y}]$ and hence tangent to D (note that this is indeed the Lie bracket of X and Y, as they do not have components transverse to $\psi(N)$ by definition). We can compute Z by this procedure at each point of M, which shows that D is involutive. $\square$

We now come to the, far less trivial, converse of the Lemma:

Theorem 7.52 (Frobenius' Theorem)**.** *Let D be an involutive k-distribution on a smooth, Hausdorff n-dimensional manifold M. Then each point $q \in M$ has a chart neighborhood (U, ϕ) such that $\phi_*D = \mathrm{span}\{\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^k}\}$, where $x^1, \dots, x^n$ are coordinates on $\phi(U)$.*

Note that, as a consequence, through each $q \in M$ passes an integral manifold $\psi\colon N \to M$ with $N = \{x \in \phi(U)\colon x^j = \phi^j(q), j > k\}$ and $\psi = \phi^{-1}|_N$. This yields the immediate

Corollary 7.53. *On a smooth, Hausdorff manifold a smooth distribution is involutive if and only if it is integrable.*

Another consequence is that, in a chart like the above, the images of integral manifolds are parallel k-planes. A collection of maximal submanifolds with this property is also called a foliation and the submanifolds are called the leaves of the foliation. One can introduce an equivalence relation $\sim$ on the manifold by saying that two points are equivalent if they belong to the same leaf. The quotient by this relation, denoted by $M/{\sim}$ or better by M/D, is called the leaf space of the distribution D. Typically it is not a manifold and not even a Hausdorff space.

Example 7.54. In Example 7.42, we already have all possibilities for the leaf space: e.g., take M to be a torus realized as the square $[0,1] \times [0,1]$ in $\mathbb{R}^2$ with opposite sides identified and $X = \frac{\partial}{\partial x} + \alpha \frac{\partial}{\partial y}$; one can show that for α rational the maximal curves are embeddings of S^1 and the leaf space is diffeomorphic to S^1, whereas for α irrational (the Kronecker foliation) the maximal curves are dense immersions of $\mathbb{R}$ and the leaf space is not Hausdorff. In Example 7.43, $D = 0$ induces the trivial equivalence relation, so $M/D = M$; on the other hand, in the case $D = TM$ all points are equivalent, so M/D is a point. In Example 7.44, every leaf is a sphere, identified by its radius, so the leaf space if $\mathbb{R}_{>0}$. In Example 7.45 in each half space $z > 0$ and $z < 0$ the leaves are characterized by the radius of the sphere, so the leaf space is the disjoint union of two copies of $\mathbb{R}_{>0}$.

Proof of Frobenius' Theorem. We prove the theorem by induction on the rank k of the distribution. For $k = 1$, this is Proposition 6.26.

We then assume we have proved the theorem for rank $k-1$. Let $(X_1, \dots, X_k)$ be generators of the distribution in a neighborhood of q. Note that in particular they are all not vanishing at q. By Proposition 6.26 we can then find a chart neighborhood (V, χ) of q with $\chi(q) = 0$ and $\chi_* X_1 = \frac{\partial}{\partial y^1}$, where $y^1, \dots, y^n$ are coordinates on $\chi(V)$.

We define new generators of $\chi_* D$ by

$$Y_1 := \chi_* X_1 = \frac{\partial}{\partial y^1}$$

and, for $i > 1$,

$$Y_i := \chi_* X_i - (\chi_* X_i(y^1))\, \chi_* X_1.$$

For $i > 1$ we then have $Y_i(y^1) = 0$ and hence, for $i, j > 1$, we have $[Y_i, Y_j](y^1) = 0$; this means that the expansion of $[Y_i, Y_j]$ in the Y_ls does not contain Y_1.

As a consequence the distribution D' defined on $S := \{y \in \chi(V) \mid y^1 = 0\}$ as the span of $Y_2, \dots, Y_k$ is involutive. By the induction assumption, we can find a neighborhood U of 0 in S and a diffeomorphism τ such that $\tau_* Y_i = \frac{\partial}{\partial w^i}$, $i = 2, \dots, k$, where $w^2, \dots, w^n$ are coordinates on $\tau(U)$.

Let $\widetilde{U}$ be $U \times (-\epsilon, \epsilon)$ for an $\epsilon > 0$ such that $\widetilde{U} \subset \chi(V)$. We then have the projection map $\pi \colon \widetilde{U} \to U$. We finally consider the diffeomorphism

$$\begin{array}{cccc} \widetilde{\tau}\colon & \widetilde{U} & \to & \tau(U) \times (-\epsilon, \epsilon) \\ & (u, y^1) & \mapsto & (\tau(u), y^1) \end{array}$$

and write $x^1 = y^1$, $x^i = \tau^i(y^2, \dots, y^n) = w^i$ for $i > 1$. We denote by $Z_i := \widetilde{\tau}_* Y_i$, $i = 1, \dots, k$ the generators of the distribution $\widetilde{D} := \widetilde{\tau}_* \chi_* D$.

Since $\frac{\partial x^i}{\partial y^1}$ is equal to 1 for $i = 1$ and to 0 otherwise, by (6.6) we get $Z_1 = \frac{\partial}{\partial x^1}$. Finally, we have, for $i = 2, \dots, k$ and $j > 1$,

$$\frac{\partial}{\partial x^1}(Z_i(x^j)) = Z_1(Z_i(x^j)) = [Z_1, Z_i](x^j) = \sum_{l=2}^{k} c_i^l Z_l(x^j),$$

where the c_i^l are functions that are guaranteed to exist by the involutivity of the distribution (note that in the sum we do not have the term for $l = 1$ since $Z_1(x^j) = 0$). For fixed j and fixed $x^2, \dots x^n$, we regard these identities as ODEs in the variable x^1. Note that, for $i = 2, \dots, k$ and $j > k$, we have $Z_i(x^j) = 0$ at $x^1 = 0$ (since at $x^1 = 0$ we have $Z_i = Y_i$). This means that $Z_i(x^j) = 0$, $i = 2, \dots, k$ and $j > k$, is the unique solution with this initial condition. These identities mean that $\widetilde{D}$ is the distribution spanned by vector fields $\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^k}$. □

7.5.1. *Quotients.* Let $\pi\colon M \to N$ be a surjective submersion. Then $\{\ker \mathrm{d}_q\pi\}_{q\in M}$ is an integrable distribution on M of rank $m - n$, with $m = \dim M$ and $n = \dim N$ (we assume M to be connected with connected fibers), with leaves the fibers $\pi^{-1}(z)$, $z \in N$.

If the leaf space N of an involutive distribution D on M can be given a manifold structure such that the canonical projection π is smooth (and hence a submersion by Frobenius theorem), then we have $\ker \mathrm{d}\pi = D$. This also implies that this manifold structure is unique up to diffeomorphism. In fact, by Frobenius theorem, locally the projection is like $U \times V \to V$ for $V \subset \mathbb{R}^{\dim N}$ and $U \subset \mathbb{R}^{\dim M - \dim N}$.

The vector fields tangent to $\ker \mathrm{d}\pi$ are called vertical (we imagine M projecting down to N). We denote by $\mathfrak{V}(M) := \Gamma(\ker \mathrm{d}\pi)$ the Lie algebra of vertical vector fields. We next consider its Lie idealizer

$$N(\mathfrak{V}(M)) = \{X \in \mathfrak{X}(M) \mid [X, Y] \in \mathfrak{V}(M) \text{ for all } Y \in \mathfrak{V}(M)\},$$

i.e., the largest Lie subalgebra of $\mathfrak{X}(M)$ in which $\mathfrak{V}(M)$ sits as an ideal. Elements of $N(\mathfrak{V}(M))$ are called projectable vector fields. The reason is that there is a well defined map

$$\phi\colon N(\mathfrak{V}(M)) \to \mathfrak{X}(N)$$

defined by the assignment

$$X_p \mapsto \mathrm{d}_p\pi X_p.$$

Note that by definition the flows of vertical vector fields change X only vertically, so its projection does not depend on the point in the fiber over $\pi(p)$.[20] Also note that ϕ is surjective and that its kernel is exactly $\mathfrak{V}(M)$. We leave to the reader to verify that ϕ is also a Lie algebra morphism (this is easy to see in local coordinates), so we have that $\mathfrak{X}(N)$ and $N(\mathfrak{V}(M))/\mathfrak{V}(M)$ are isomorphic Lie algebras.

It is often convenient to work with generators.

Definition 7.55. A family $\mathcal{Y}$ of vertical vector fields generates $\mathfrak{V}(M)$ if every $Y \in \mathfrak{V}(M)$ can be written as a finite linear combination of elements of $\mathcal{Y}$, which are hence called generators (i.e., $\mathcal{Y}$ generates $\mathfrak{V}(M)$ as a $C^\infty(M)$-module).

One can then easily see that X is projectable if and only if $[X, Y]$ is vertical for every $Y \in \mathcal{Y}$.

8. Vector bundles

The tangent bundle introduced in the previous sections is an example of a more general structure known as a vector bundle. Many important objects defined on manifolds (e.g., densities, tensor fields, differential forms) are sections of vector bundles.

8.1. General definitions.

Definition 8.1. A $\mathcal{C}^k$-vector bundle of rank r over a $\mathcal{C}^k$-manifold of dimension n is a $\mathcal{C}^k$-manifold E together with a surjection $\pi\colon E \to M$ such that:

(1) $E_q := \pi^{-1}(q)$ is an r-dimensional vector space for all $q \in M$.
(2) E possesses a $\mathcal{C}^k$-atlas of the form $\{(\widetilde{U}_\alpha, \widetilde{\phi}_\alpha)\}_{\alpha\in I}$ with $\widetilde{U}_\alpha = \pi^{-1}(U_\alpha)$ for a $\mathcal{C}^k$-atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ of M and

$$\begin{array}{cccc} \widetilde{\phi}_\alpha\colon & \widetilde{U}_\alpha & \to & \mathbb{R}^n \times \mathbb{R}^r \\ & (q, v \in E_q) & \mapsto & (\phi_\alpha(q), A_\alpha(q)v) \end{array}$$

where $A_\alpha(q)$ is a linear isomorphism for all $q \in U_\alpha$.

An atlas for E like $\{(\widetilde{U}_\alpha, \widetilde{\phi}_\alpha)\}_{\alpha\in I}$ in this definition is called an adapted atlas for the vector bundle. The corresponding atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ for M is called a trivializing atlas.

---

[20]If $\mathsf{L}_Y X = \widetilde{Y}$, with Y and $\widetilde{Y}$ vertical, we get

$$\frac{\partial}{\partial s}(\Phi^Y_{-s})_* X = (\Phi^Y_{-s})_* \widetilde{Y}$$

by Lemma 7.31. Alternatively, we can observe that, in local coordinates $\{x^1, \dots, x^n, y^1, \dots, y^{m-n}\}$, where the ys are the vertical coordinates, a vector field $X(x, y) = \sum_{i=1}^n X^i(x, y)\frac{\partial}{\partial x^i} + \sum_{i=1}^{m-n} \widetilde{X}^i(x, y)\frac{\partial}{\partial y^i}$ is projectable if and only if $\frac{\partial X^i}{\partial y^j} = 0$ for all i, j, as follows from (7.3). In this local picture, the projection $\phi(X)$ is the vector field on N represented by $\sum_{i=1}^n X^i(x)\frac{\partial}{\partial x^i}$.

Notice that π is a $\mathcal{C}^k$-map with respect to this manifold structure and that for $k > 0$ it is a submersion.

The maps

$$A_{\alpha\beta}(q) := A_\beta(q)A_\alpha(q)^{-1}\colon \mathbb{R}^r \to \mathbb{R}^r$$

are $\mathcal{C}^k$ in q (i.e., $A_{\alpha\beta}\colon U_\alpha \cap U_\beta \to \mathrm{End}(\mathbb{R}^r)$ is a $\mathcal{C}^k$-map, where we identify $\mathrm{End}(\mathbb{R}^r)$ with $\mathbb{R}^{r^2}$ with its standard manifold structure) for all $\alpha, \beta \in I$. The transition maps

$$\widetilde{\phi}_{\alpha\beta}(x,u) = (\phi_{\alpha\beta}(x), A_{\alpha\beta}(\phi_\alpha^{-1}(x))u)$$

are linear in the second factor $\mathbb{R}^r$. The point dependent linear maps $A_{\alpha\beta}$ are usually called the transition functions of the vector bundle.

Example 8.2. It is readily verified that the tangent bundle TM of a $\mathcal{C}^k$-manifold M with $k \geq 1$ is a $\mathcal{C}^{k-1}$-vector bundle where we regard the base manifold M as a $\mathcal{C}^{k-1}$-manifold.

Example 8.3. Let $E := M \times V$ where V is an r-dimensional vector space. Then E is a vector bundle of rank r over M with π the projection to M. A choice of basis for V determines an isomorphism $\tau\colon V \to \mathbb{R}^r$. Given an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ for M, we set $\widetilde{\phi}_\alpha(q, v) = (\phi_\alpha(q), \tau v)$. Note that then $A_{\alpha\beta}(q) = \mathrm{Id}$ for all α, β, q.

Example 8.4 (The dual bundle)**.** If E is a vector bundle over M, as in Definition 8.1, then the union of the dual spaces E_q^* is also a vector bundle, called the dual bundle of E. Namely, let $E^* := \cup_{q\in M} E_q^*$. We denote an element of E^* as a pair (q, ω) with $\omega \in E_q^*$. We let $\pi_{E^*}(q, \omega) = q$. To an atlas $\{(\widetilde{U}_\alpha, \widetilde{\phi}_\alpha)\}_{\alpha\in I}$ of E we associate the atlas $\{(\widehat{U}_\alpha, \widehat{\phi}_\alpha)\}_{\alpha\in I}$ of E^* with $\widehat{U}_\alpha = \pi_{E^*}^{-1}(U_\alpha) = \cup_{q\in U_\alpha} E_q^*$ and

$$\begin{array}{cccc} \widehat{\phi}_\alpha\colon & \widehat{U}_\alpha & \to & \mathbb{R}^n \times (\mathbb{R}^r)^* \\ & (q, \omega \in E_q^*) & \mapsto & (\phi_\alpha(q), (A_\alpha(q)^*)^{-1}\, \omega) \end{array}$$

where we regard $(\mathbb{R}^r)^*$ as the manifold $\mathbb{R}^r$ with its standard structure. It follows that we have transitions maps

$$\widehat{\phi}_{\alpha\beta}(x,u) = (\phi_{\alpha\beta}(x), (A_{\alpha\beta}(\phi_\alpha^{-1}(x))^*)^{-1}\, u).$$

Note that actually any linear construction on vector spaces can be carried over to vector bundles. For example, starting from a vector bundle $E \to M$ we may define its endomorphism bundle $\mathrm{End}(E)$ with $\mathrm{End}(E)_p := \mathrm{End}(E_p)$. If $F \to M$ is a second vector bundle over M, we may define the direct sum $E \oplus F$ and the tensor product $E \otimes F$ with $(E \oplus F)_p := E_p \oplus F_p$ and $(E \otimes F)_p := E_p \otimes F_p$. We leave it to the reader to construct the corresponding trivializing atlases.

Example 8.5 (Pullback bundle)**.** Let $F\colon N \to M$ be a $\mathcal{C}^k$-map and $E \xrightarrow{\pi} M$ a $\mathcal{C}^k$-vector bundle. One defines $F^*E := \{(q,e) \in N \times E \mid F(q) = \pi(e)\}$. One can readily see that F^*E is a $\mathcal{C}^k$-vector bundle over M with projection map $\pi_{F^*E}(q,e) = q$. In practice, the fiber of F^*E at q is given by the fiber of E^* at $F(q)$ and the fiber transition maps of F^*E at q are given by the fiber transition maps of E at $F(q)$. More precisely, we pick an atlas $\{(V_j, \psi_j)\}_{j\in J}$ of N. To the atlas in Definition 8.1, we then associate a new atlas $\{(V_{\alpha j}, \psi_{\alpha j})\}_{(\alpha,j)\in I\times J}$ of N with $V_{\alpha j} := F^{-1}(U_\alpha) \cap V_j$ and $\psi_{\alpha j} := \psi_{j|_{V_{\alpha j}}}$. The atlas of F^*E is then given by $\widehat{V}_{\alpha j} = \pi^{-1}_{F^*E}(V_{\alpha j}) = \cup_{q\in V_{\alpha j}} E_{F(q)}$ and

$$\begin{array}{rccc} \widehat{\psi}_{\alpha j}\colon & \widehat{V}_{\alpha j} & \to & \mathbb{R}^s \times \mathbb{R}^r \\ & (q, v \in E_{F(q)}) & \mapsto & (\psi_{\alpha j}(q), A_\alpha(F(q))\, v) \end{array}$$

where s is the dimension of N. It follows that we have transitions maps

$$\widehat{\psi}_{(\alpha j)(\beta j')}(x,u) = (\psi_{(\alpha j)(\beta j')}(x), A_{\alpha\beta}(F(\phi^{-1}_{\alpha j}(x)))u).$$

8.1.1. *Sections.* We now come to the generalization of the notion of vector fields to other vector bundles.

Definition 8.6. A section (also called a global section) of a $\mathcal{C}^k$-vector bundle $E \xrightarrow{\pi} M$ is a $\mathcal{C}^k$-map $\sigma\colon M \to E$ with $\pi \circ \sigma = \mathrm{Id}_M$. We denote by $\Gamma(E)$ the space of sections of E. A section of the restriction of E to an open subset U is also called a local section on U.

Remark 8.7. If σ is a section of $E \to M$, then $\sigma(q)$ is of the form (q, σ_q) where $\sigma_q \in E_q$ for all $q \in M$. We will use this notation throughout.

Remark 8.8. Notice that $\Gamma(E)$ is a vector space. Moreover, if σ is a section of E and f a function on M, we can define a new section $f\sigma$ by letting $(f\sigma)_q$ be the product of $f(q) \in \mathbb{R}$ and $\sigma_q \in E_q$. Hence $\Gamma(E)$ is also a module over $\mathcal{C}^k(M)$.

Example 8.9. A section of the tangent bundle TM is then the same as a vector field on M.

Example 8.10. A section of the vector bundle $M \times V \to M$ of Example 8.3 is a map $q \mapsto (q, f(q))$ where f is a map $M \to V$. Thus, the module of sections of $M \times V \to V$ is canonically isomorphic to the module of maps from M to V.

Example 8.11. A section of the dual bundle E^* of Example 8.4 associates to a point $q \in M$ a vector in E^*_q. Hence, if ω is a section of E^* and σ a section of E, we can evaluate ω_q on σ_q for each q. This produces a function $\omega(\sigma)$ by $\omega(\sigma)(q) := \omega_q(\sigma_q)$. Notice that the map

$\Gamma(E^*) \times \Gamma(E) \to \mathcal{C}^k(M)$, $(\omega, \sigma) \mapsto \omega(\sigma)$ is $\mathcal{C}^k(M)$-bilinear. This is also called the pairing of $\Gamma(E^*)$ with $\Gamma(E)$. For this reason $\omega(\sigma)$ is also often denoted as (ω, σ) or $\langle \omega, \sigma \rangle$.

Example 8.12. A section of the pullback bundle F^*E of Example 8.5 associates to a point $q \in M$ a vector in $E_{F(q)}$.

If one picks a trivializing atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha \in I}$ as in Definition 8.1, then a section of E is the same as a collection of $\mathcal{C}^k$-maps[21] $\sigma_\alpha \colon \phi_\alpha(U_\alpha) \to \mathbb{R}^r$ such that

$$\boxed{\sigma_\beta(\phi_{\alpha\beta}(x)) = A_{\alpha\beta}(\phi_\alpha^{-1}(x))\, \sigma_\alpha(x)} \tag{8.1}$$

for all $\alpha, \beta \in I$ and for all $x \in \phi_\alpha(U_\alpha \cap U_\beta)$.

8.1.2. *Vector subbundles.* We now come to the generalization of the notion of plane distributions to general vector bundles.

Definition 8.13. A vector subbundle of a vector bundle $E \to M$ is a collection $\{F_p\}_{p \in M}$ of subspaces $F_p \subset E_p$ such that $F := \cup_{p \in M} F_p$ is also a vector bundle.

Lemma 8.14. *A collection $\{F_p\}_{p \in M}$ of k-dimensional subspaces is a vector subbundle of $E \to M$ if and only if every point p in M possesses an open neighborhood U and sections $\sigma_1, \dots, \sigma_k$ of the restriction of E to U such that*

$$F_x = \operatorname{span}\{(\sigma_1)_x, \dots, (\sigma_k)_x\}$$

for all $x \in U$. These sections are called (local) generators of F on U.

Proof. If F is a vector bundle, around each point p of M we may take a chart (U_α, ϕ_α) as in Definition 8.1. In $\phi_\alpha(U_\alpha)$ we may take k linearly indipendent maps $\tau_1, \dots, \tau_k$ from $\phi_\alpha(U_\alpha)$ to $\mathbb{R}^k$. The sections $(\sigma_i)_x := A_\alpha^{-1}(x)\tau_i(\phi_\alpha(x))$ are then generators of F on U_α.

For the other implication the assumption is that for each p we have an open neighborhood U_p with generators $\sigma_{p1}, \dots, \sigma_{pk}$. We may also assume that each U_p is a chart domain, with a chart map denoted by ϕ_p. We then consider the atlas $\{(U_p, \phi_p)\}_{p \in M}$. We consider $G := \cup_{p \in M} F_p^*$. On $\widetilde{U}_p := \cup_{x \in U_p} F_x^*$ we may define the chart map

$$\widetilde{\phi}_p(x, \omega) = (\phi_p(x), A_p(x)\omega)$$

with

$$A_p(x)\omega = ((\omega, (\sigma_{p1})_x), \dots, (\omega, (\sigma_{pk})_x)).$$

This shows that G is a vector bundle and hence that its dual F is a vector subbundle of E. □

[21]These maps are also called local sections.

Corollary 8.15. *A plane distribution D on M is smooth if and only $\cup_{p\in M} D_p$ is a vector subbundle of TM.*

8.1.3. *Morphisms.* We now come to the definition of maps compatible with vector bundle structures.

Definition 8.16. Let $E \xrightarrow{\pi_E} M$ and $F \xrightarrow{\pi_F} N$ be $\mathcal{C}^k$-vector bundles. A pair of $\mathcal{C}^k$-maps $\Psi\colon E \to F$ and $\psi\colon M \to N$ is said to be compatible with the bundle structure if $\pi_F \circ \Psi = \psi \circ \pi_E$.

$$\begin{array}{ccc} E & \xrightarrow{\Psi} & F \\ {\scriptstyle \pi_E}\downarrow & & \downarrow{\scriptstyle \pi_F} \\ M & \xrightarrow[\psi]{} & N \end{array}$$

This means that Ψ maps E_q to $F_{\psi(q)}$ for all $q \in M$. We denote by Ψ_q this map. The compatible pair (Ψ, ψ) is called a morphism of vector bundles (or a vector bundle map) if Ψ_q is linear for all $q \in M$. Usually one simply writes $\Psi\colon E \to F$ to denote the morphism. One says that Ψ is a morphism over ψ.

Example 8.17. If F is a vector subbubndle of E, the inclusion map is a morphism.

Example 8.18. If ψ is map from M to M, then we have a morphism $\Psi\colon M \times V \to M \times V$, $(q, v) \mapsto (\psi(q), v)$.

Notice that a composition of morphisms is also a morphism. A morphism is called an isomorphism if it possesses an inverse. Note if Ψ is an isomorphism, so is also ψ.

Definition 8.19. If (Ψ, ψ) is an isomorphism, we can push forward sections of E to sections of F: for $\sigma \in \Gamma(E)$, we define

$$(\Psi_*\sigma)_y := \Psi_{\psi^{-1}(y)}\sigma_{\psi^{-1}(y)}$$

for $y \in N$.

Remark 8.20. If $\psi\colon M \to N$ is a $\mathcal{C}^{k+1}$-map, then we have a morphism, called the tangent lift, $\Psi\colon TM \to TN$ by setting $\Psi_q = \mathrm{d}_q\psi$. Notice that Ψ is an isomorphism if and only ψ is a diffeomorphism and that in this case the push-forward Ψ_* is what we have denoted by ψ_* so far.

Remark 8.21. If $\{(\widetilde{U}_\alpha, \widetilde{\phi}_\alpha)\}_{\alpha\in I}$ is an adapted atlas of a vector bundle $E \to M$ over the trivializing atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ of M, then $\widetilde{\phi}_\alpha$ is an isomorphism from $\pi^{-1}(U_\alpha)$ to $\phi_\alpha(U_\alpha) \times \mathbb{R}^r$ for all α and $\widetilde{\phi}_{\alpha\beta}$ is an isomorphism from $\phi_\alpha(U_\alpha \cap U_\beta) \times \mathbb{R}^r$ to $\phi_\beta(U_\alpha \cap U_\beta) \times \mathbb{R}^r$ for all distinct α and β.

$$\begin{array}{ccc} \pi_E^{-1}(U_\alpha) & \xrightarrow{\widetilde{\phi}_\alpha} & \phi_\alpha(U_\alpha) \times \mathbb{R}^r \\ \downarrow & & \downarrow \\ U_\alpha & \xrightarrow[\phi_\alpha]{} & \phi_\alpha(U_\alpha) \end{array} \qquad \begin{array}{ccc} \phi_\alpha(U_\alpha \cap U_\beta) \times \mathbb{R}^r & \xrightarrow{\widetilde{\phi}_{\alpha\beta}} & \phi_\beta(U_\alpha \cap U_\beta) \times \mathbb{R}^r \\ \downarrow & & \downarrow \\ \phi_\alpha(U_\alpha \cap U_\beta) & \xrightarrow[\phi_{\alpha\beta}]{} & \phi_\beta(U_\alpha \cap U_\beta) \end{array}$$

If σ is a section of E, then the representation σ_α of σ can be written as $\sigma_\alpha = (\widetilde{\phi}_\alpha)_* \sigma|_{U_\alpha}$. The compatibility relations (8.1) now read

$$\boxed{\sigma_\beta = (\widetilde{\phi}_{\alpha\beta})_* \sigma_\alpha}$$

where again, by abuse of notation, σ_α actually denotes the restriction of σ_α to $\phi_\alpha(U_\alpha \cap U_\beta)$.

Example 8.22. Let F^*E be the pullback bundle of Example 8.5. Then we have a morphism $\Psi\colon F^*E \to E$ by setting Ψ_q to be the identity map for all q.

Definition 8.23. A vector bundle $E \to M$ is called trivial (or trivializable) if it is isomorphic to a vector bundle of the form $M \times V \to M$ (see Example 8.3).

Note that the base map ψ of the isomorphism $\Psi\colon E \to M \times V$ is also an isomorphism. So we may compose Ψ with the inverse of the morphism defined in Example 8.18. This means that we can always assume the trivializing isomorphism to be over the identity map.

Proposition 8.24. *A rank r vector bundle is trivial if and only if it has r $\mathbb{R}$-linearly independent sections.*

Proof. Choose a basis of V to identify it with $\mathbb{R}^n$. The vector bundle $M \times \mathbb{R}^r \to M$ has the r linearly independent sections $\sigma_1, \dots, \sigma_r$ given by the maps from M to each of the r components of $\mathbb{R}^r$. If Ψ is an isomorphism from E to $M \times \mathbb{R}^r$, then $\Psi_*^{-1}\sigma_1, \dots, \Psi_*^{-1}\sigma_r$ are r linearly independent sections of E.

Conversely, if $\tau_1, \dots, \tau_r$ are r sections of E, we have a map $\Psi\colon M \times \mathbb{R}^r \to E$, with $\psi(q) = q$ and $\Psi_q(e_i) = (\tau_i)_q$, where $(e_1, \dots, e_r)$ is the canonical basis of $\mathbb{R}^r$. It the sections are linearly independent, this is an isomorphism. Hence, E is trivial by the isomorphism Ψ^{-1}. □

A manifold whose tangent bundle is trivial is called parallelizable. Each open subset of $\mathbb{R}^n$ is clearly parallelizable.

Lemma 8.25. *The circle S^1 is parallelizable.*

Proof. The easiest way to see this is by viewing S^1 as a submanifold of $\mathbb{R}^2$. A point q in S^1 is then a unit vector in $\mathbb{R}^2$ and there is a unique rotation R_q that maps q to the point $(1, 0)$. We then define $\Psi\colon TS^1 \to S^1 \times \mathbb{R}$ by $(q, v) \mapsto (q, R_q v)$. □

In Remark 8.30 we will give another proof. We will see, Lemma 8.31, that also the 3-sphere is parallelizable. In Section 10.1, we will recover these statements as special cases of the fact that Lie groups are parallelizable, see Lemma 10.6. On the other hand, e.g., the sphere S^2 is not parallelizable: actually, one can show that S^2 does not even possess a single nowhere vanishing vector field, see Lemma 8.32.

8.1.4. *Vector bundles from local data.* The linear maps $A_{\alpha\beta}$ are actually enough to specify a vector bundle. Namely, let M be a $\mathcal{C}^k$-manifold. Assume we have an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ for M and, for some fixed vector space V, $\mathcal{C}^k$-maps $A_{\alpha\beta}\colon U_\alpha \cap U_\beta \to \mathrm{End}(V)$ for all $\alpha, \beta \in I$, $\alpha \neq \beta$, such that for all distinct $\alpha, \beta, \gamma \in I$:

(1) $A_{\alpha\beta}(q)A_{\beta\alpha}(q) = \mathrm{Id}_V$ for all $q \in U_\alpha \cap U_\beta$;
(2) $A_{\beta\gamma}(q)A_{\alpha\beta}(q) = A_{\alpha\gamma}(q)$ for all $q \in U_\alpha \cap U_\beta \cap U_\gamma$.

Then we can repeat the construction of subsection 6.1 verbatim. Namely, we define a coordinatized vector at $q \in M$ as a triple $(U_\alpha, \phi_\alpha, v)$ where (U_α, ϕ_α) is a chart containing q and v is a vector in V. We define two coordinatized vectors $(U_\alpha, \phi_\alpha, v)$ and (U_β, ϕ_β, w) at q to be equivalent if $w = A_{\alpha\beta}(q)v$; notice that properties (1) and (2) above ensure that this is an equivalence relation. Finally, we define E_q as the set of equivalence classes of coordinatized vectors at q. A choice of basis on V yields an isomorphism $B\colon V \to \mathbb{R}^r$, $r = \dim V$, and we can define the bijection

$$\begin{aligned} A_\alpha(q)\colon \quad E_q &\to \mathbb{R}^r \\ [(U_\alpha, \phi_\alpha, v)] &\mapsto Bv \end{aligned}$$

and use it to define the vector space structure on E_q that makes $A_\alpha(q)$ into a linear isomorphism. Exactly as in the proof of Lemma 6.6, we show that this vector space structure is canonical (i.e., independent of α). Finally, we see that $E := \cup_{q\in M} E_q$ is a vector bundle over M with $\widetilde{U}_\alpha = \cup_{q\in U_\alpha} E_q$.

Remark 8.26. We can define the dual bundle, see Example 8.4, also by local data. Namely, we start from the transition functions $A_{\alpha\beta}$ for E and define $A^{E^*}_{\alpha\beta}(q) := ((A_{\alpha\beta}(q))^*)^{-1}$ for all $q \in U_\alpha \cap U_\beta$. Conditions (1) and (2) of Section 8.1.4 are automatically satisfied.

Remark 8.27. Note the difference in the conventions in the construction of a vector bundles from local data on a given manifold and the construction of a manifold from local data of Section 4.1. Namely, here we consider the transition functions $A_{\alpha\beta}$s as depending on a point on $U_\alpha \cap U_\beta$, whereas there we considered the transition maps $\phi_{\alpha\beta}$ as functions on $V_{\alpha\beta}$ which, later on, turns out to be $\phi_\alpha(V_{\alpha\beta})$. The reason is that conditions (1) and (2) above read much better with this convention. Equivalently, we may define $\widetilde{A}_{\alpha\beta}(x) := A_{\alpha\beta}(\phi_\alpha^{-1}(x))$ for $x \in \phi_\alpha(U_\alpha \cap U_\beta)$. With these new notations the condition take the uglier form

(1) $\widetilde{A}_{\alpha\beta}(x)\widetilde{A}_{\beta\alpha}(\phi_{\alpha\beta}(x)) = \mathrm{Id}_{\mathbb{R}^r}$ for all $x \in \phi_\alpha(U_\alpha \cap U_\beta)$;
(2) $\widetilde{A}_{\beta\gamma}(\phi_{\alpha\beta}(x))\widetilde{A}_{\alpha\beta}(x) = \widetilde{A}_{\alpha\gamma}(x)$ for all $x \in \phi_\alpha(U_\alpha \cap U_\beta \cap U_\gamma)$.

We can express the triviality of a vector bundle in terms of local data.

Proposition 8.28. *Let E be a vector bundle defined by local data $A_{\alpha\beta}$ as above. Then E is trivial if and only if for each α there is a map $A_\alpha\colon U_\alpha \to \mathrm{Aut}(\mathbb{R}^n)$ such that, on $U_\alpha \cap U_\beta$,*

$$A_{\alpha\beta} = A_\beta^{-1}A_\alpha$$

for all distinct α and β.

Proof. Let $\Psi\colon E \to M \times \mathbb{R}^r$ be an isomorphism over the identity. If $\{(\widetilde{U}_\alpha, \widetilde{\phi}_\alpha)\}_{\alpha\in I}$ is an adapted atlas of E, we set $\Psi_\alpha := \Psi \circ \widetilde{\phi}_\alpha^{-1}$. This is an isomorphism from $\phi_\alpha(U_\alpha) \times \mathbb{R}^r$ to $U_\alpha \times \mathbb{R}^r$. We then have, on $U_\alpha \cap U_\beta$, $\Psi_\beta^{-1} \circ \Psi_\alpha = \widetilde{\phi}_\beta \circ \widetilde{\phi}_\alpha^{-1} = \widetilde{\phi}_{\alpha\beta}$. If we define $A_\alpha(q) := (\Psi_\alpha)_{\phi_\alpha(q)}$, $q \in U_\alpha$, we then get $A_{\alpha\beta} = A_\beta^{-1}A_\alpha$.

Conversely, we can use the given A_αs to define the isomorphisms Ψ_α from $\phi_\alpha(U_\alpha) \times \mathbb{R}^r$ to $U_\alpha \times \mathbb{R}^r$. The relations $A_{\alpha\beta} = A_\beta^{-1}A_\alpha$ show that they patch together to an isomorphism $E \to M \times \mathbb{R}^r$. □

Remark 8.29. If we define $\widetilde{A}_\alpha(x) := A_\alpha(\phi_\alpha^{-1}(x))$, the triviality relations reads

$$\widetilde{A}_{\alpha\beta}(x) = \widetilde{A}_\beta^{-1}(\phi_{\alpha\beta}(x))\widetilde{A}_\alpha(x), \qquad \forall x \in \phi_\alpha(U_\alpha \cap U_\beta).$$

Example 8.30. We may use this to give another proof of Lemma 8.25 on the triviality of TS^1. Define S^1 as in Example 4.12. Note the in this case we can simply write $\phi_{12}(x) = \phi_{21}(x) = 1/x$. We then have

$\widetilde{A}_{12}(x) = \mathrm{d}_x\phi_{12}(x) = -1/x^2$. By setting $\widetilde{A}_1(x) = 1/x$ and $\widetilde{A}_2(y) = -1/y$, we see that TS^1 is trivial.

We may similarly prove the triviality of TS^3:

Lemma 8.31. *The* 3*-sphere* S^3 *is parallelizable.*

Proof. Here we use the the diffeomorphism ψ between $\mathbb{R}^3$ and the space V of self-adjoint 2×2 complex matrices:

$$\psi(\mathbf{x}) = \begin{pmatrix} z & x+\mathrm{i}y \\ x-\mathrm{i}y & -z \end{pmatrix} =: X,$$

with $\mathbf{x} = (x,y,z)$. One can easily see that $\det X = -||x||^2$ and that $X^2 = -\det X\ \mathrm{Id}$. The first consequence is that $V_{12} = V_{21} = \mathbb{R}^3\setminus\{0\}$ is identified with $V\setminus\{0\}$. The second is that the maps $\phi_{12}(\mathbf{x}) = \phi_{21}(\mathbf{x}) = \frac{\mathbf{x}}{||\mathbf{x}||^2}$ get identified with the maps $\widehat{\psi}_{ij} := \psi\circ\phi_{ij}\circ\psi^{-1}$, $\widehat{\psi}_{ij}(X) = -\frac{X}{\det X}$. By the second identity above we get

$$\widehat{\psi}_{12}(X) = \widehat{\psi}_{21}(X) = X^{-1}.$$

An easy way to compute the differential of $\widehat{\psi}_{12}$ is to consider the path $\gamma(t) := X + tB$, for fixed $X\in V\setminus\{0\}$ and $B\in V$. We then have

$$\widehat{\psi}_{12}(\gamma(t)) = (X+tB)^{-1} = (X(\mathrm{id}+tX^{-1}B))^{-1} = (\mathrm{id}+tX^{-1}B)^{-1}X^{-1}.$$

Using the geometric series $\sum_{n=0}^{\infty} t^n A^n = (\mathrm{id}-tA)^{-1}$, we get

$$\widehat{\psi}_{12}(\gamma(t)) = (\mathrm{id} - tX^{-1}B + O(t^2))X^{-1}.$$

It follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}_{|t=0}\widehat{\psi}_{12}(\gamma(t)) = -X^{-1}\dot{\gamma}(0)X^{-1}.$$

In other words, $\widetilde{A}_{12}(X) = \mathrm{d}_X\widehat{\psi}_{12}$ is the automorphism of V defined by $\widetilde{A}_{12}(X)B = -X^{-1}BX^{-1}$. For $X\in V\setminus\{0\}$ we now consider the automorphism L_X and R_X of V defined by

$$L_XB := XB, \qquad R_XB := BX.$$

Note that $(L_X)^{-1} = L_{X^{-1}}$ and $(R_X)^{-1} = R_{X^{-1}}$. Finally define

$$\widetilde{A}_1(X) := L_{X^{-1}} \quad\text{ and }\quad \widetilde{A}_2(Y) := -R_{Y^{-1}}$$

and check that $\widetilde{A}_{12}(X) = \widetilde{A}_2^{-1}(\phi_{12}(X))\widetilde{A}_1(X)$ for all $X\in V\setminus\{0\}$. □

On the other hand, we have the following

Lemma 8.32. *There is no nowhere vanishing vector field on* S^2*.*

The following simple proof, adapted from [2], is based on the notion of the winding number and its invariance under homotopy. (We will introduce this in Section 9.5.3 in the case of differentiable curves, but the result holds also for continous curves and homotopies. In particular, the Lemma above holds also in the case of continuous vector fields.)

Proof. We use again the description of Example 4.12, which corresponds to the stereographic projections. Let X be a vector a field and assume it has no zeros apart possibly at the north pole: i.e., its lower emisphere representation X_2 in $V_2 = \mathbb{R}^2$ has no zeros.

Next consider the upper hemisphere representation X_1 of X in $V_1 = \mathbb{R}$. Outside of the origins of V_1 and V_2, we have $X_1 = (\phi_{21})_* X_2$, i.e., $X_1(\phi_{21}(\mathbf{x})) = \mathrm{d}_\mathbf{x}\phi_{21} X_2(\mathbf{x})$. We may compute

$$\mathrm{d}_\mathbf{x}\phi_{21} = \frac{1}{||x||^4}\begin{pmatrix} y^2 - x^2 & -2xy \\ -2xy & x^2 - y^2 \end{pmatrix},$$

with $\mathbf{x} = (x, y)$. We now restrict to $u = (\cos\theta, \sin\theta) \in S^1$ and get

$$X_1(u) = \begin{pmatrix} \cos(2\theta) & -\sin(2\theta) \\ \sin(2\theta) & \cos(2\theta) \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} X_2(u).$$

We may regard this as a curve $\gamma_1 \colon S^1 \to \mathbb{R}^2 \setminus \{0\}$, $\gamma_1(u) := X_1(u)$. We also define $\Gamma \colon [0,1] \times S^1 \to \mathbb{R}^2 \setminus \{0\}$ by

$$\Gamma(s, u) := \begin{pmatrix} \cos(2\theta) & -\sin(2\theta) \\ \sin(2\theta) & \cos(2\theta) \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} X_2(su),$$

where we use the fact that X_2 has no zeros. Note that Γ is a homotopy from γ_1 to

$$\gamma_0(u) = \begin{pmatrix} \cos(2\theta) & -\sin(2\theta) \\ \sin(2\theta) & \cos(2\theta) \end{pmatrix}\mathbf{v}$$

with $\mathbf{v} = \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right) X_2(0)$. As γ_0 is a circle winding twice in the anticlockwise direction around the origin, we get that its winding number is 2. By homotopy invariance we then get that also the winding number of γ_1 is 2. This shows that X_1 must have a zero inside the disk of radius 1 centered at the origin, since otherwise $\widetilde{\Gamma}(s, u) := X_1(su)$ would be a homotopy to a constant loop, which would imply that the winding number of γ_1 is zero. □

8.2. **Densities and integration.** We consider a simple, but very important class of examples of line bundles (i.e., rank 1 vector bundles) associated to every manifold: the bundles of s-densities, where s is a real number. We will show that 1-densities, usually called just densities, are the natural objects to integrate and that nonnegative 1-densities define a measure. Densities with other weights $s = 1/p$ are needed to define the notion of L^p-spaces associated to manifolds.

Let M be a $\mathcal{C}^{k+1}$-manifold, $k \geq 0$. Fix an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha \in I}$. The differential $\mathrm{d}_x\phi_{\alpha\beta}$ of a transition map is a linear map $T_x\phi_\alpha(U_\alpha \cap U_\beta) \to T_{\phi_{\alpha\beta}(x)}\phi_\beta(U_\alpha \cap U_\beta)$. However, since the chart images are open subsets of $\mathbb{R}^n$, we can canonically identify their tangent spaces with $\mathbb{R}^n$ itself. It follows that $\mathrm{d}_x\phi_{\alpha\beta}$ is canonically given as an $n \times n$ matrix (the Jacobian matrix of the map $\phi_{\alpha\beta}$ with respect to the given coordinates), so we can compute its determinant. Next we fix a real number s and take $V := \mathbb{R}$ and $A_{\alpha\beta}(q) := |\det \mathrm{d}_{\phi_\alpha(q)}\phi_{\alpha\beta}|^{-s}$. The requirements (1) and (2) of Section 8.1.4 are satisfied, so we get a $\mathcal{C}^k$-vector bundle of rank 1 denoted by $|\Lambda M|^s$ over M.

Sections of $|\Lambda M|^s$ are called s-densities. The representation σ_α of an s-density σ in the chart (U_α, ϕ_α) is just a $\mathcal{C}^k$-function on $\phi_\alpha(U_\alpha)$. What distinguishes it from the representation of a function are the transition rules:

$$\sigma_\beta(\phi_{\alpha\beta}(x)) = |\det \mathrm{d}_x\phi_{\alpha\beta}|^{-s}\, \sigma_\alpha(x), \tag{8.2}$$

for all $\alpha, \beta \in I$ and for all $x \in \phi_\alpha(U_\alpha \cap U_\beta)$.

Remark 8.33. Notice that the transition functions are positive, so it makes sense to define nonnegative densities as densities that are nonnegative in each representation (i.e., $\sigma_\alpha \geq 0$ for all α) and positive densities as densities that are positive in each representation (i.e., $\sigma_\alpha > 0$ for all α).

Remark 8.34. There are several immediate consequences of (8.2). The first is that 0-densities are just functions. The second is that the product of an s_1-density σ_1 and an s_2-density σ_2 yields an $(s_1 + s_2)$-density $\sigma_1\sigma_2$. If σ is a nonnegative s-density and $r > 0$, then σ^r is a nonnegative rs-density. Finally, if σ is a positive s-density and $r \in \mathbb{R}$, then σ^r is a positive rs-density.

Remark 8.35. By looking directly into the definition of $|\Lambda M|^s$, by the discussion in Section 8.1.4, one immediately sees that $|\Lambda M|^0 = M \times \mathbb{R}$ and $|\Lambda M|^{-s} = (|\Lambda M|^s)^*$.

8.2.1. *Integration.* For $s = 1$, one simply speaks of densities. Densities are the natural objects one can integrate on a manifold (which possesses a partition of unity).

Let us consider first the case of a compact Hausdorff manifold M. Let σ be a density on M and $\{(U_\alpha, \phi_\alpha)\}_{\alpha \in I}$ an atlas on M. By Theorem 5.6 we can find a finite partition of unity $\{\rho_j\}_{j \in J}$ subordinate to $\{U_\alpha\}_{\alpha \in I}$ (i.e., for each j we have an α_j with $\operatorname{supp} \rho_j \subset U_{\alpha_j}$). Notice that $\operatorname{supp}\rho_j$ is compact by Lemma A.9. Since ϕ_{α_j} is a homeomorphism, by Lemma A.8 we see that also $\phi_{\alpha_j}(\operatorname{supp}\rho_j)$ is compact (in $\mathbb{R}^n$, $n = \dim M$). Finally, observe that the representation $(\rho_j\sigma)_{\alpha_j}$ of the density $\rho_j\sigma$ in the chart (U_α, ϕ_α) is a $\mathcal{C}^k$-function, so at least continuous, in $\phi_{\alpha_j}(U_{\alpha_j})$, so it is integrable on $\phi_{\alpha_j}(\operatorname{supp}\rho_j)$. One then defines

$$\int_{M;\{(U_\alpha,\phi_\alpha)\},\{\rho_j\}} \sigma = \sum_{j\in J} \int_{\phi_{\alpha_j}(\operatorname{supp}\rho_j)} (\rho_j\sigma)_{\alpha_j}\, d^n x.$$

(Here $d^n x$ stands for the Lebesgue measure on $\mathbb{R}^n$, also otherwise denoted by $dx^1\cdots dx^n$.)

Lemma 8.36. *This integral does not depend on the choice of atlas and of partition of unity.*

Proof. Consider an atlas $\{(\bar U_{\bar\alpha}, \bar\phi_{\bar\alpha})\}_{\bar\alpha\in\bar I}$ and a finite partition of unity $\{\bar\rho_{\bar\jmath}\}_{\bar\jmath\in\bar J}$ subordinate to it. From $\sum_{\bar\jmath\in\bar J}\bar\rho_{\bar\jmath} = 1$, it follows that $\sigma = \sum_{\bar\jmath\in\bar J}\bar\rho_{\bar\jmath}\,\sigma$, so we have

$$\int_{M;\{(U_\alpha,\phi_\alpha)\},\{\rho_j\}} \sigma = \sum_{j\in J} \int_{\phi_{\alpha_j}(\operatorname{supp}\rho_j)} (\rho_j\sigma)_{\alpha_j}\, d^n x = \\ = \sum_{\bar\jmath\in\bar J}\sum_{j\in J} \int_{\phi_{\alpha_j}(\operatorname{supp}\rho_j)} (\rho_j\bar\rho_{\bar\jmath}\,\sigma)_{\alpha_j}\, d^n x,$$

where we have taken out the finite sum $\sum_{\bar\jmath\in\bar J}$. Next observe that

$$S_{j\bar\jmath} := \int_{\phi_{\alpha_j}(\operatorname{supp}\rho_j)} (\rho_j\bar\rho_{\bar\jmath}\,\sigma)_{\alpha_j}\, d^n x = \\ = \int_{\phi_{\alpha_j}(\operatorname{supp}\rho_j\cap\operatorname{supp}\bar\rho_{\bar\jmath})} (\rho_j\bar\rho_{\bar\jmath}\,\sigma)_{\alpha_j}\, d^n x = \\ = \int_{\phi^{-1}_{\bar\alpha_{\bar\jmath}\alpha_j}(\bar\phi_{\bar\alpha_{\bar\jmath}}(\operatorname{supp}\rho_j\cap\operatorname{supp}\bar\rho_{\bar\jmath}))} (\rho_j\bar\rho_{\bar\jmath}\,\sigma)_{\alpha_j}\, d^n x,$$

where $\phi_{\bar\alpha_{\bar\jmath}\alpha_j}$ denotes the transition map $\phi_{\alpha_j}(U_{\alpha_j}) \to \bar\phi_{\bar\alpha_{\bar\jmath}}(\bar U_{\bar\alpha_{\bar\jmath}})$. Since $\rho_j\bar\rho_{\bar\jmath}\,\sigma$ is a density, we have

$$(\rho_j\bar\rho_{\bar\jmath}\,\sigma)_{\alpha_j}(x) = |\det \mathrm{d}_x\phi_{\bar\alpha_{\bar\jmath}\alpha_j}|(\rho_j\bar\rho_{\bar\jmath}\,\sigma)_{\bar\alpha_{\bar\jmath}}(\bar x),$$

with $\bar x = \phi_{\bar\alpha_{\bar\jmath}\alpha_j}(x)$ and $x \in \phi_{\alpha_j}(\operatorname{supp}\rho_j \cap \operatorname{supp}\bar\rho_{\bar\jmath})$. By the change-of-variables formula, we then have

$$S_{j\bar\jmath} = \int_{\bar\phi_{\bar\alpha_{\bar\jmath}}(\operatorname{supp}\rho_j\cap\operatorname{supp}\bar\rho_{\bar\jmath})} (\rho_j\bar\rho_{\bar\jmath}\,\sigma)_{\bar\alpha_{\bar\jmath}}\, d^n\bar x = \int_{\bar\phi_{\bar\alpha_{\bar\jmath}}(\operatorname{supp}\bar\rho_{\bar\jmath})} (\rho_j\bar\rho_{\bar\jmath}\,\sigma)_{\bar\alpha_{\bar\jmath}}\, d^n\bar x.$$

Hence

$$\sum_{j\in J} S_{j\bar{\jmath}} = \int_{\bar{\phi}_{\bar{\alpha}_{\bar{\jmath}}}(\operatorname{supp}\bar{\rho}_{\bar{\jmath}})} (\bar{\rho}_{\bar{\jmath}}\,\sigma)_{\bar{\alpha}_{\bar{\jmath}}}\; d^n\bar{x}$$

and

$$\int_{M;\{(U_\alpha,\phi_\alpha)\},\{\rho_j\}} \sigma = \sum_{\bar{\jmath}\in\bar{J}}\sum_{j\in J} S_{j\bar{\jmath}} = \\ = \sum_{\bar{\jmath}\in\bar{J}} \int_{\bar{\phi}_{\bar{\alpha}_{\bar{\jmath}}}(\operatorname{supp}\bar{\rho}_{\bar{\jmath}})} (\bar{\rho}_{\bar{\jmath}}\,\sigma)_{\bar{\alpha}_{\bar{\jmath}}}\; d^n\bar{x} = \int_{M;\{(\bar{U}_{\bar{\alpha}},\bar{\phi}_{\bar{\alpha}})\},\{\bar{\rho}_{\bar{\jmath}}\}} \sigma.$$

□

We can hence drop the choice of atlas and partition of unity from the notation and simply write:

$$\boxed{\int_M \sigma = \sum_{j\in J} \int_{\phi_{\alpha_j}(\operatorname{supp}\rho_j)} (\rho_j\sigma)_{\alpha_j}\; d^n x} \tag{8.3}$$

This formula is also well defined in the case when M is a Hausdorff, second-countable manifold and σ is a density with compact support. The main point is Theorem 5.7 that ensures the existence of a partition of unity subordinate to the given trivialing atlas. The sum over j is well defined as in a neighborhood of each point only finitely many ρ_j are different from zero. The integrals on the right hand side converge as they are actually defined on $\phi_{\alpha_j}(\operatorname{supp}\rho_j \cap \operatorname{supp}\sigma)$ and $\operatorname{supp}\rho_j \cap \operatorname{supp}\sigma$ is compact.[22] The proof of the independence on the choice of trivializing atlas and partition of unity is exactly as above.

Remark 8.37. In many a situation a density σ on M is chosen once and for all. In this case, one can define the integral of a function f (compactly supported if σ is not compactly supported) as the integral of the density $f\sigma$.

The integral on manifolds is additive not only with respect to the integrating densities but also with respect to the union of integration domains. We present a simple case that is useful for explicit computations. First, we say that a subset N of M has measure zero if, for some atlas $\{(U_\alpha,\phi_\alpha)\}_{\alpha\in I}$, we have that $\phi_\alpha(N\cap U_\alpha)$ is a Lebesgue null

[22]This is true since the supports are, by definition, closed. Hence their intersection is also closed. As this is a subset of $\operatorname{supp}\sigma$, which is compact, it is also compact.

set for all $\alpha \in I$. Let M_1 and M_2 be disjoint open subsets of M such that $M \setminus (M_1 \cup M_2)$ has measure zero. Then we have

$$\int_M \sigma = \int_{M_1} \sigma + \int_{M_2} \sigma,$$

where M_1 and M_2 are now regarded as manifolds.

This property is very useful for actual computations. Suppose that we can find finitely many mutually disjoint open subsets M_k, $k \in K$, such that their union differs from M by a set of measure zero and such that each M_k is entirely contained in a chart U_{α_k}. Then we can choose an atlas for M_k consisting of the single chart $(M_k, (\phi_{\alpha_k})_{|M_k})$. As a partition of unity subordinate to it we take the function 1 on M_k. We then have $\int_{M_k} \sigma = \int_{\phi_{\alpha_k}(M_k)} \sigma_{\alpha_k}\, d^n x$ for all $k \in K$. Hence

$$\int_M \sigma = \sum_{k\in K} \int_{\phi_{\alpha_k}(M_k)} \sigma_{\alpha_k}\, d^n x.$$

Remark 8.38. A nonnegative density σ (i.e., $\sigma_\alpha \geq 0$ for all α in some atlas) defines a measure on M, which we assume to be Hausdorff and second countable. Namely, let $\mathcal{B}(M)$ be the Borel algebra of M (i.e., the σ-algebra generated by the open subsets of M). Pick an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ and a partition of unity $\{\rho_j\}_{j\in J}$ subordinate to it. If $A \in \mathcal{B}(M)$, then $A \cap \operatorname{supp} \rho_j$ is also Borel. Since ϕ_{α_j} is a homeomorphism, also $\phi_{\alpha_j}(A \cap \operatorname{supp} \rho_j)$ is Borel. Since $(\rho_j\sigma)_{\alpha_j}$ is continuous, it is Lebesgue-measurable; since it is nonnegative, we can define $\int_{\phi_{\alpha_j}(A\cap \operatorname{supp} \rho_j)} (\rho_j\sigma)_{\alpha_j} d^n x$, which is allowed to be infinity. We then define

$$\mu_\sigma(A) := \sum_{j\in J} \int_{\phi_{\alpha_j}(A\cap \operatorname{supp} \rho_j)} (\rho_j\sigma)_{\alpha_j} d^n x.$$

As in the proof to Lemma 8.36, one can see that μ_σ does not depend on the choice of atlas and of partition of unity. One can also verify that μ_σ is a measure on $\mathcal{B}(M)$. Finally, if $f \in \mathcal{C}^k(M)$ is such that $f\sigma$ has compact support, we have

$$\int_M f\sigma = \int_M f \mathrm{d}\mu_\sigma.$$

We can now use the right hand side to extend the integral to all μ_σ-integrable functions. Another advantage is that in the right hand side we can use all measure theoretic techniques; for example, we can write M as a countable union of mutually disjoint Borel subsets M_k and write $\int_M f\sigma = \sum_k \int_{M_k} f \mathrm{d}\mu_\sigma$. Notice that we no longer have to require that the M_ks be open.

Remark 8.39. Densities appear naturally in several instances, as we will see. For example, a Riemannian metric naturally defines a positive density, see Remark 8.78. On the other hand, every top differential form on a connected orientable manifold naturally defines two densities (one for each orientation), see Section 9.4.

Remark 8.40. Using Remark 8.34, one can define the L^p-space associated to a Hausdorff, second-countable manifold as the completion of the space of the $(1/p)$-densities σ such $\int_M |\sigma|^p < \infty$. If p and and q are conjugate, i.e. $1/p + 1/q = 1$, we have a pairing of $\sigma_1 \in L^p$ with $\sigma_2 \in L^q$ by integrating $\sigma_1\sigma_2$. In particular, $L^2(M)$, the completion of the space of half-densities on M, is a Hilbert space.

8.2.2. *Densities on vector spaces.* We now give a more conceptual description of densities. The idea is to define s-densities directly on vector spaces and then to apply this to the case of the tangent space.

Definition 8.41. For $s \neq 0$, an s-density on a real vector space V of dimension n is a function μ on V^n satisfying

$$\mu(Av_1, \dots, Av_n) = |\det A|^s \mu(v_1, \dots, v_n) \tag{8.4}$$

for all $v_1, \dots, v_n \in V$ and all automorphisms A of V. A 0-density is defined to be a constant function on V^n.

Note that a linear combination of s-densitites is also an s-density. We denote by $|\Lambda V|^s$ the vector space of s-densities on V. We want to show that this space is one-dimensional. First we show that it is not zero-dimensional.

Example 8.42. Let $\mathcal{F} = (e_1, \dots, e_n)$ be a frame, i.e., an ordered basis, of V. To an n-tuple $(v_1, \dots, v_n)$ of vectors we associate the endomorphism $T_{v_1,\dots,v_n}$ that maps e_i to v_i for all i. Then we define

$$\mu_{\mathcal{F},s}(v_1, \dots, v_n) := |\det T_{v_1,\dots,v_n}|^s,$$

which is clearly an s-density. Note that the expansions $v_i = \sum_{j=1}^n v_i^j e_j$ imply that the representing matrix of $T_{v_1,\dots,v_n}$ in the frame $\mathcal{F}$ is (v_i^j). We can then write

$$\mu_{\mathcal{F},s}(v_1, \dots, v_n) = \left|\det\left(v_i^j\right)\right|^s.$$

We also need the following observation.

Lemma 8.43. *Let μ be an an s-density on V, $s \neq 0$. If the vectors $v_1, \dots, v_n$ are linearly dependent, then $\mu(v_1, \dots, v_n) = 0$.*

Proof. Let W be the span of $v_1, \dots, v_n$. Let W' be a complement, which by assumption is not zero-dimensional. For some $\lambda \in \mathbb{R} \setminus \{0, 1, -1\}$, let A be the automorphism defined by $Aw = w$ if $w \in W$ and $Aw' = \lambda w'$ if $w' \in W'$. Then we get $\mu(v_1, \dots, v_n) = \mu(Av_1, \dots, Av_n) = |\lambda|^{ks}\mu(v_1, \dots, v_n)$, $k = \dim W' > 0$. This yields $\mu(v_1, \dots, v_n) = 0$, since $|\lambda| \neq 1$. □

From the example and the Lemma we get the

Proposition 8.44. *Let μ be an s-density on V. Then, for any frame $\mathcal{F} = (e_1, \dots, e_n)$ of V, we have*

$$\mu(v_1, \dots, v_n) = \lambda\mu_{\mathcal{F},s}(v_1, \dots, v_n)$$

with $\lambda = \mu(e_1, \dots, e_n)$.

Proof. If $s = 0$, both sides of the equations are constant functions and they are equal since $\mu_{\mathcal{F},0} = 1$. If $s \neq 0$ and the vectors are linearly dependent, then both sides vanish. If they are linearly independent, we have

$$\mu(v_1, \dots, v_n) = \mu(T_{v_1,\dots,v_n}e_1, \dots, T_{v_1,\dots,v_n}e_n) = |\det T_{v_1,\dots,v_n}|^s\mu(e_1, \dots, e_n),$$

which concludes the proof. □

As a consequence, a frame $\mathcal{F}$ of V induces a basis $(\mu_{\mathcal{F},s})$ of $|\Lambda V|^s$. Thus,

$$\dim |\Lambda V|^s = 1$$

for all s.

Let now $\phi\colon V \to W$ be a linear map between n-dimensional vector spaces. If μ is an s-density on W we define, on V^n,

$$\phi^*\mu(v_1, \dots, v_n) := \mu(\phi v_1, \dots, \phi v_n).$$

Lemma 8.45. *$\phi^*\mu$ is an s-density on V.*

Proof. For $s = 0$ this is obvious. Consider $s \neq 0$. If ϕ not an isomorphism, then the vectors $\phi v_1, \dots, \phi v_n$ are linearly dependent, so $\phi^*\mu$ is identically equal to zero. If ϕ is an isomorphism, we have

$$\begin{aligned}\phi^*\mu(Av_1, \dots, Av_n) &= \mu(\phi Av_1, \dots, \phi Av_n) = \\ &= \mu(\phi A\phi^{-1}\phi v_1, \dots, \phi A\phi^{-1}\phi v_n) = |\det \phi A\phi^{-1}|^s\mu(\phi v_1, \dots, \phi v_n) = \\ &= |\det A|^s\phi^*\mu(v_1, \dots, v_n).\end{aligned}$$

□

We then have a linear map $\phi^*\colon |\Lambda W|^s \to |\Lambda V|^s$, called the pullback of densities. Note that the product of an s_1-density μ_1 with an s_2-density μ_2 is an (s_1+s_2)-density $\mu_1\mu_2$ and we have

$$\phi^*(\mu_1\mu_2) = \phi^*\mu_1\phi^*\mu_2.$$

Moreover, if $\psi\colon W \to Z$ is also a linear map, and $\dim Z = n$, we have

$$(\psi\circ\phi)^* = \phi^*\psi^*.$$

Remark 8.46. As an s-density vanishes on linearly dependent vectors, some textbooks prefer to define it only on linearly independent ones. Namely, they define an s-density, for any s, as a function μ on the space $F(V)$ of frames of V satisfying (8.4) for all $(v_1,\dots,v_n)\in F(V)$ and all automorphisms A of V. Note that with this definition the pullback ϕ^* is only defined when ϕ is an isomorphism.

Suppose now that we have chosen a frame $\mathcal{F} = (e_1,\dots,e_n)$ of V and a frame $\widetilde{\mathcal{F}} = (\widetilde{e}_1,\dots,\widetilde{e}_n)$ of W. Then we can compute

$$\phi^*\mu_{\widetilde{\mathcal{F}},s}(e_1,\dots,e_n) = |\det \mathsf{A}_{\phi,\mathcal{F},\widetilde{\mathcal{F}}}|^s,$$

where $\mathsf{A}_{\phi,\mathcal{F},\widetilde{\mathcal{F}}}$ is the matrix representing ϕ in the given frames. As a consequence, by Proposition 8.44, we have

$$\phi^*\mu_{\widetilde{\mathcal{F}},s} = |\det \mathsf{A}_{\phi,\mathcal{F},\widetilde{\mathcal{F}}}|^s\mu_{\mathcal{F},s}. \tag{8.5}$$

Finally, if $\phi\colon V\to W$ is an isomorphism, we define the push-forward by ϕ as

$$\phi_* := (\phi^*)^{-1}.$$

8.2.3. *Back to densities on manifolds.* We may apply the above construction to give an abstract definition of the density bundles. Let M be a manifold. We now define $|\Lambda M|^s$ as $\cup_{q\in M}|\Lambda T_qM|^s$. Given an atlas $\{(U_\alpha,\phi_\alpha)\}_{\alpha\in I}$, we have, for all $q\in U_\alpha$, the linear map

$$(\mathrm{d}_q\phi_\alpha)_*\colon |\Lambda T_qM|^s \to |\Lambda\mathbb{R}^n|^s.$$

We also have an isomorphism $|\Lambda\mathbb{R}^n|^s\to\mathbb{R}$ given by the frame $(\mu_{\mathcal{F},s})$ of $|\Lambda\mathbb{R}^n|^s$ induced by the canonical frame $\mathcal{F}$ of $\mathbb{R}^n$. The fiber component $A_\alpha(q)$ of the resulting adapted chart map is the composition of $(\mathrm{d}_q\phi_\alpha)_*$ with this isomorphism.

To check that this construction defines the same line bundle as we have used before, we just have to compute the transition functions. For $q\in U_\alpha\cap U_\beta$ we have that $A_{\alpha\beta}(q)$ is the composition of $(\mathrm{d}_{\phi_\alpha(q)}\phi_{\alpha\beta})_*$ with the isomorphisms to $\mathbb{R}$ given by the canonical frame. By (8.5), we get that $A_{\alpha\beta}(q) = |\det \mathrm{d}_{\phi_\alpha(q)}\phi_{\alpha\beta}|^{-s}$, as expected.

Remark 8.47. Note that we can avoid identifying $|\Lambda\mathbb{R}^n|^s$ with $\mathbb{R}$ and think of the chart map on $\widetilde{U}_\alpha$ as taking values in $\phi_\alpha(U_\alpha) \times |\Lambda\mathbb{R}^n|^s$. In this case the transition functions are directly given by the $(\mathrm{d}_{\phi_\alpha(q)}\phi_{\alpha\beta})_*$s.

Remark 8.48 (Standard density)**.** The s-density on $\mathbb{R}^n$ determined by the canonical basis as in Example 8.42 is called the standard s-density and is denoted by the symbol $|d^nx|^s$. The standard s-density on an open subset U of $\mathbb{R}^n$, also denoted by $|d^nx|^s$, is the positive s-density that takes the value $|d^nx|^s$ at each point. If $s = 1$, we can integrate it and we clearly have $\int_U |d^nx| = \int_U d^nx$, where d^nx denotes the Lebesgue measure. Note that the standard 1-density and the Lebesgue measure are conceptually two different objects that are identified in integration just because of the definition of the integral of a density. We will keep different notations for the sake of clarity, but there is no risk of confusion in using the same notation.

8.2.4. *Pullback and push-forward.* Let M and N be manifolds of the same dimension. Let F be a map from M to N and let σ be an s-density on N. We define the pullback $F^*\sigma$ of σ as the s-density on M given by

$$(F^*\sigma)_q = (\mathrm{d}_qF)^*\sigma_{F(q)}, \quad q \in M.$$

In coordinates, the pullback is then described as follows. We first pick an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ of M and an atlas $\{(V_j, \psi_j)\}_{j\in J}$ of N. If $\{\sigma_j\}$ denotes the representation of σ in the atlas for N and $\{F_{\alpha j}\}$ denotes the representation of F with respect to the two atlases, we have

$$(F^*\sigma)_\alpha(x) = |\det \mathrm{d}_xF_{\alpha j}|^s\,\sigma_j(F_{\alpha j}(x)),$$

for $x \in \phi_\alpha(U_\alpha)$.

Also notice that, by construction F^* is linear and that

$$F^*(\sigma_1\sigma_2) = F^*\sigma_1\,F^*\sigma_2.$$

Finally, if M, N, Z are manifolds of the same dimension and we have maps $F\colon M \to N$ and $G\colon N \to Z$, then

$$(G \circ F)^* = F^*G^*.$$

Remark 8.49. In the particular case of a map F of open subsets of $\mathbb{R}^n$, we have the following useful formula for the standard densities:

$$\boxed{F^*|d^nx|^s = |\det \mathrm{d}F|^s\,|d^nx|^s} \tag{8.6}$$

In this formula $\mathrm{d}F$ stands for d_xF where x is the point where we want to evaluate the formula.

In case $F\colon M \to N$ is a diffeomorphism, we can also push forward densities from M to N, just by setting

$$F_* := (F^{-1})^*.$$

Of course we have that

$$F_*(\sigma_1\sigma_2) = F_*\sigma_1 F_*\sigma_2,$$

and, given diffeomorphisms $F\colon M \to N$ and $G\colon N \to Z$, that

$$(G \circ F)_* = G_* F_*.$$

Remark 8.50. Using push-forwards, we can write the representation of an s-density σ in a chart (U, ϕ_U) as

$$(\phi_U)_*\sigma|_U = \sigma_U |d^n x|^s.$$

To understand this formula just notice that the left and right hand side of this equation are both s-densities on $\phi_U(U)$ that are represented in the standard chart of $\phi_U(U)$ by the same function σ_U. If we pick an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$, we then have

$$(\phi_\alpha)_*\sigma|_{U_\alpha} = \sigma_\alpha |d^n x|^s, \tag{8.7}$$

for all α, and we can rewrite the compatibility equation (8.2) as

$$\boxed{\sigma_\beta |d^n x|^s = (\phi_{\alpha\beta})_*(\sigma_\alpha |d^n x|^s)}$$

for all $\alpha, \beta \in I$, where, by abuse of notation, σ_α denotes here the restriction of σ_α to $\phi_\alpha(U_\alpha \cap U_\beta)$ and σ_β denotes the restriction of σ_β to $\phi_\beta(U_\alpha \cap U_\beta)$. Since the push-forward is an algebra morphism, we have that $(\phi_{\alpha\beta})_*(\sigma_\alpha |d^n x|^s) = (\phi_{\alpha\beta})_*\sigma_\alpha (\phi_{\alpha\beta})_* |d^n x|^s$. By (8.6), we have that $(\phi_{\alpha\beta})_* |d^n x|^s = |\det \mathrm{d}\phi|^{-s} |d^n x|^s$, which then yields the transformation rule (8.2) for the coefficients σ_α.

Remark 8.51. If $F\colon M \to N$ is a diffeomorphism and σ a density on M, then the change-of-variables formula immediately implies

$$\boxed{\int_M \sigma = \int_N F_*\sigma} \tag{8.8}$$

assuming that the integrals converge. If f is a function on M, we also have

$$\int_M f\sigma = \int_N F_* f\, F_*\sigma = \int_N f \circ F^{-1}\, F_*\sigma.$$

Remark 8.52. Recall that, if M and N are measurable spaces and $F\colon M \to N$ is a measurable map, then we can define the push-forward of a measure μ on M to a measure $F_*\mu$ on N by

$$(F_*\mu)(A) := \mu(F^{-1}(A))$$

for any measurable subset A of N. If M and N are manifolds, F a diffeomorphism and σ a nonnegative density on M, then we immediately get

$$F_*\mu_\sigma = \mu_{F_*\sigma}.$$

8.2.5. *The Lie derivative.* If M is Hausdorff, we may define the Lie derivative of an s-density in complete analogy with what we did for functions and vector fields (see also Remark 7.30). For simplicity we are now going to consider only smooth manifolds. Namely, we set

$$\mathsf{L}_X\sigma := \left.\frac{\partial}{\partial t}\right|_{t=0}(\Phi^X_{-t})_*\,\sigma = \left.\frac{\partial}{\partial t}\right|_{t=0}(\Phi^X_t)^*\sigma,$$

for $X \in \mathfrak{X}(M)$ and σ an s-density. The Lie derivative is again $\mathbb{R}$-linear. Moreover, we have

$$\left.\frac{\partial}{\partial t}\right|_{t=s}(\Phi^X_{-t})_*\sigma = (\Phi^X_{-s})_*\mathsf{L}_X\sigma = \mathsf{L}_X(\Phi^X_{-s})_*\sigma \tag{8.9}$$

for all s for which the flow is defined. The proof is exactly the same as in Lemma 7.31. This also implies that σ is preserved by the flow of X if and only if the Lie derivative vanishes:

$$(\Phi^X_t)_*\sigma = \sigma\ \forall t \quad \text{if and only if} \quad \mathsf{L}_X\sigma = 0.$$

Moreover, the Lie derivative satisfies the Leibniz rule

$$\mathsf{L}_X(\sigma_1\sigma_2) = \mathsf{L}_X\sigma_1\,\sigma_2 + \sigma_1\,\mathsf{L}_X\sigma_2.$$

Finally if $F\colon M \to N$ is a diffeomorphism, by the properties of the push-forward and by Proposition 7.28, we get

$$F_*\mathsf{L}_X\sigma = \mathsf{L}_{F_*X}F_*\sigma, \tag{8.10}$$

for all $X \in \mathfrak{X}(M)$ and all $\sigma \in \Gamma(|\Lambda M|^s)$. Finally, exactly as in the proof of Lemma 7.35, we get

$$\begin{aligned}\mathsf{L}_{X+Y}\sigma &= \mathsf{L}_X\sigma + \mathsf{L}_Y\sigma,\\ \mathsf{L}_X\mathsf{L}_Y\sigma - \mathsf{L}_Y\mathsf{L}_X\sigma &= \mathsf{L}_{[X,Y]}\sigma,\end{aligned}$$

for any two vector fields X and Y.

A useful interpretation of the Lie derivative of densities is given by an application of the change-of-variables formula (8.8).

Lemma 8.53. *Let U be an open subset of M, which we regard as a submanifold. We consider Φ^X_t as a diffeomorphism from U to $\Phi^X_t(U)$. Then*

$$\frac{\partial}{\partial t}\int_{\Phi^X_t(U)}\sigma = \int_{\Phi^X_t(U)}\mathsf{L}_X\sigma. \tag{8.11}$$

In particular, this shows that $\int_{\Phi_t^X(U)} \sigma$ is constant for all Us, for which the integral converges, if and only if $\mathsf{L}_X\sigma = 0$.

Proof. By the change-of-variables formula (say, for t_1 close to t and t_2 close to 0), we have

$$\int_{\Phi_{t_1}^X(U)} (\Phi_{t_2}^X)_*\sigma = \int_{\Phi_{t_1-t_2}^X(U)} \sigma.$$

From the right hand side we see that this depends only on the difference between t_1 and t_2, so the derivative with respect to t_1 must be equal to the derivative with respect to t_2. Hence, also using (8.9), we get

$$\frac{\partial}{\partial t_1}\int_{\Phi_{t_1}^X(U)} (\Phi_{t_2}^X)_*\sigma = -\frac{\partial}{\partial t_2}\int_{\Phi_{t_1}^X(U)} (\Phi_{t_2}^X)_*\sigma = \int_{\Phi_{t_1}^X(U)} (\Phi_{t_2}^X)_*\mathsf{L}_X\sigma.$$

By applying the change-of-variables formula again, we then get

$$\frac{\partial}{\partial t_1}\int_{\Phi_{t_1-t_2}^X(U)} \sigma = \int_{\Phi_{t_1-t_2}^X(U)} \mathsf{L}_X\sigma,$$

which is the equation we wanted to prove by setting $t = t_1 - t_2$. (More precisely, you can make the change of variables $t = t_1 - t_2$, $s = t_1 + t_2$, and observe that $\frac{\partial}{\partial t_1} = \frac{\partial}{\partial t} + \frac{\partial}{\partial s}$ and that the expression does not depend on s.) □

This result has an even better formulation if σ is nonnegative, for, as explained in Remark 8.38, it defines a measure μ_σ. We define

$$\mathrm{Vol}(U) = \mu_\sigma(U)$$

where U is a Borel subset of M. If we also have a vector field X, we define the volume at time t by

$$\mathrm{Vol}_t(U) := \mu_\sigma(\Phi_t^X(U)).$$

Here the intuition comes from fluid dynamics where X is interpreted as a velocity field that prescribes the fluid motion; under this motion the fluid region U is transformed into $\Phi_t^X(U)$, and we are interested in measuring the volume as the region changes with time. With these new notations equation (8.11) becomes

$$\frac{\partial}{\partial t}\mathrm{Vol}_t(U) = \int_{\Phi_t^X(U)} \mathsf{L}_X\sigma$$

for all open subsets U. The vector field X is called volume preserving if $\frac{\partial}{\partial t}\Big|_{t=0}\mathrm{Vol}_t(U) = 0$ for every open subset U. By the above formula one sees that

$$X \text{ is volume preserving} \quad \text{if and only} \quad \mathsf{L}_X\sigma = 0.$$

In fluid mechanics where X is the velocity field of a fluid, one speaks instead of an incompressible flow when this condition holds; typically, liquids are incompressible as a good approximation.

We finally come to the computation of the Lie derivative of a density. We start with the

Lemma 8.54. *Let U be an open subset of $\mathbb{R}^n$ and $X = \sum_{i=1}^n X^i\partial_i$ a vector field on U. Then, for every s,*

$$\mathsf{L}_X|d^nx|^s = s\sum_{i=1}^n \partial_i X^i |d^nx|^s.$$

Proof. By (8.6), we have

$$\mathsf{L}_X|d^nx|^s = \frac{\partial}{\partial t}\Big|_{t=0}(\Phi_t^X)^*|d^nx|^s = \frac{\partial}{\partial t}\Big|_{t=0}|\det \mathrm{d}\Phi_t^X|^s|d^nx|^s.$$

Since

$$(\Phi_t^X(x))^i = x^i + tX^i(x) + O(t^2),$$

we have

$$\partial_j(\Phi_t^X)^i = \delta_j^i + t\partial_j X^i + O(t^2).$$

Hence, by the Leibniz formula for determinants,

$$\det \mathrm{d}\Phi_t^X = 1 + t\sum_{i=1}^n \partial_i X^i + O(t^2).$$

Finally,

$$|\det \mathrm{d}\Phi_t^X|^s = (\det \mathrm{d}\Phi_t^X)^s = 1 + st\sum_{i=1}^n \partial_i X^i + O(t^2),$$

so

$$\frac{\partial}{\partial t}\Big|_{t=0}|\det \mathrm{d}\Phi_t^X|^s = s\sum_{i=1}^n \partial_i X^i.$$

□

Proposition 8.55. *Let σ be an s-density and X a vector field on M, $\dim M = n$. In an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ of M, we have*

$$(\mathsf{L}_X\sigma)_\alpha = \sum_{i=1}^n (X_\alpha^i\partial_i\sigma_\alpha + s\sigma_\alpha\partial_i X_\alpha^i).$$

In particular, if $s = 1$ we have

$$(\mathsf{L}_X\sigma)_\alpha = \sum_{i=1}^n \partial_i(\sigma_\alpha X_\alpha^i). \tag{8.12}$$

Proof. By (8.7) and (8.10), we have $(\mathsf{L}_X\sigma)_\alpha|d^nx|^s = \mathsf{L}_{X_\alpha}(\sigma_\alpha|d^nx|^s)$. But by the Leibniz rule we have

$$\mathsf{L}_{X_\alpha}(\sigma_\alpha|d^nx|^s) = X_\alpha(\sigma_\alpha)\,|d^nx|^s + \sigma_\alpha\mathsf{L}_{X_\alpha}|d^nx|^s.$$

Finally, we use Lemma 8.54. □

Remark 8.56. The local formula may be used to define the Lie derivative of a density also when the manifold is not Hausdorff or not smooth.

Notice that (8.12) immediately implies that, if σ is a density, X a vector field and f a function, we then have

$$\mathsf{L}_{fX}\sigma = \mathsf{L}_X(f\sigma). \tag{8.13}$$

8.2.6. *Positive densities and the divergence of vector fields.* We now concentrate on positive densities, for which the above results can be refined. The crucial point is that, if σ is a positive density on M, then for every density τ on M there is a uniquely defined function f such that $\tau = f\sigma$ (this is easily shown using the chart representations). As a consequence, for every vector field X, there is a uniquely defined function $\operatorname{div}_\sigma X$, called the divergence[23] of X with respect to σ, such that

$$\boxed{\mathsf{L}_X\sigma = \operatorname{div}_\sigma X\,\sigma}$$

Remark 8.57. Notice that a vector field X is then volume preserving if and only its divergence vanishes.

Remark 8.58. Notice that positive densities are rather general. An example is the standard density $|d^nx|$ on an open subset of $\mathbb{R}^n$. More generally, every Hausdorff, second-countable manifold admits a positive density. This is easily proved. Let $\{\rho_j\}_{j\in J}$ be a partition of unity subordinate to an atlas $\{(U_\alpha,\phi_\alpha)\}_{\alpha\in I}$ of M, with $\dim M = n$. Then we define $\sigma = \sum_{j\in J}\rho_j\,(\phi_{\alpha_j}^{-1})_*|d^nx|$. By Proposition 8.24, $|\Lambda M|$, and hence $|\Lambda M|^s$ for all s, is then a trivializable line bundle for every manifold M that admits partitions of unity (so, e.g., compact Hausdorff of Hausdorff second-countable manifolds).

Remark 8.59. From the Leibniz rule $\mathsf{L}_X(f\sigma) = X(f)\sigma + f\mathsf{L}_X\sigma$, where f is a function, we get

$$\operatorname{div}_{f\sigma}X = \frac{1}{f}X(f) + \operatorname{div}_\sigma X$$

[23] This term also comes from fluid dynamics: if $\operatorname{div}_\sigma X \neq 0$, the volume of open subsets changes. This means that the fluid has to move in or out of U and hence that the flow lines of X, for an appropriate time direction, move apart, i.e., diverge.

for every positive function f. From the Leibniz rule and from (8.13), we also have

$$\mathrm{div}_\sigma(fX) = X(f) + f\mathrm{div}_\sigma X.$$

for every function f. For f positive, we then have

$$\boxed{\mathrm{div}_{f\sigma} X = \frac{1}{f}\mathrm{div}_\sigma(fX)}$$

Remark 8.60. A nice application of these formulae occurs in fluid dynamics where X is the velocity field. Let ρ be a physical density (of matter, of charge,...). Despite the name ρ is a function, which is in general also allowed to be time dependent. In addition one has a given 1-density σ (in examples, often the standard density on $\mathbb{R}^n$).[24] The integral $M_t(U) = \int_{\Phi_t^X(U)} \rho\sigma$ then represents the total quantity (of mass, charge,...) in the region U at time t. If the quantity is conserved, then $\frac{\partial}{\partial_t} M_t(U)$ must be zero for every open subset U. Again by the change-of-variable formula (8.8) and by (8.9), as in the proof to Lemma 8.53, we see that this happens if and only if $\frac{\partial\rho}{\partial t}\sigma + \mathsf{L}_X(\rho\sigma) = 0$. By the formulae above, $\mathsf{L}_X(\rho\sigma) = \mathsf{L}_{\rho X}\sigma = \mathrm{div}_\sigma(\rho X)\,\sigma$. Writing $J := \rho X$, which represents the current of the transported quantity, we see that the quantity is conserved if and only if

$$\frac{\partial\rho}{\partial t} + \mathrm{div}_\sigma J = 0.$$

This is called the continuity equation.

We finally come to the computation of the divergence of a vector field. Pick an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ on M. Then by (8.7) and (8.12), we have

$$\boxed{(\mathrm{div}_\sigma X)_\alpha = \sum_{i=1}^n \frac{1}{\sigma_\alpha}\partial_i(\sigma_\alpha X_\alpha^i)} \tag{8.14}$$

Remark 8.61. If U is an open subset of $\mathbb{R}^n$ and $X = \sum_{i=1}^n X^i\partial_i$ is a vector field on U, then the divergence of X with respect to the standard density $|d^n x|$, which we denote simply by div, has the usual form from calculus:

$$\mathrm{div} X = \sum_{i=1}^n \partial_i X^i.$$

[24]A physical density, like e.g. mass per unit volume, is a ratio of 1-denstities (in this example mass density over volume density), and hence a 0-density, i.e., a function.

Remark 8.62. The local formula (8.14) may be used to define the divergence of a vector field also when the manifold is not Hausdorff.

8.3. **The cotangent bundle and 1-forms.** The dual bundle (see Example 8.4) of the tangent bundle TM is denoted by T^*M and is called the cotangent bundle of M. Its fiber at q is denoted by T^*_qM. Its sections are called 1-forms. The space of 1-forms on M is denoted by

$$\Omega^1(M) := \Gamma(T^*M).$$

If one picks a trivializing atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$, then a 1-form is the same as a collection of $\mathcal{C}^k$-maps $\omega_\alpha\colon \phi_\alpha(U_\alpha) \to \mathbb{R}^n$ such that

$$\boxed{(\mathrm{d}_x\phi_{\alpha\beta})^*\omega_\beta(\phi_{\alpha\beta}(x)) = \omega_\alpha(x)} \tag{8.15}$$

for all $\alpha, \beta \in I$ and for all $x \in \phi_\alpha(U_\alpha \cap U_\beta)$.

If $F\colon M \to N$ is a $\mathcal{C}^k$-map and $\omega \in \Omega^1(N)$, we may define a 1-form $F^*\omega$ on M by

$$(F^*\omega)_q := (\mathrm{d}_qF)^*\omega_{F(q)}.$$

This is called the pullback of 1-forms and extends the pullback of functions defined in subsection 3.1. The pullback of 1-forms is clearly $\mathbb{R}$-linear. Moreover, if $f \in \mathcal{C}^k(N)$ and $\omega \in \Omega^1(N)$, then

$$F^*(f\omega) = F^*f F^*\omega.$$

Finally, again, if $G\colon N \to Z$ is also a $\mathcal{C}^k$-map, then

$$(G \circ F)^* = F^*G^*.$$

Using pullbacks, we can write the representation of ω in a chart (U, ϕ_U) as $\omega_U = (\phi_U^{-1})^*\omega|_U$. Moreover, we can rewrite the compatibility equation (8.15) as

$$\boxed{\omega_\alpha = \phi^*_{\alpha\beta}\omega_\beta}$$

for all $\alpha, \beta \in I$, where, by abuse of notation, ω_α denotes here the restriction of ω_α to $\phi_\alpha(U_\alpha \cap U_\beta)$ and ω_β denotes the restriction of ω_β to $\phi_\beta(U_\alpha \cap U_\beta)$. This just extends Remark 3.12 from functions to 1-forms.

The pairing (ω, X) of a 1-form ω and a vector field X (see Example 8.11) is often denoted by $\iota_X\omega$ and called the contraction of X with ω. If $F\colon M \to N$ is a diffeomorphism, then we clearly have

$$(F^*\omega, X) = F^*(\omega, F_*X)$$

for all $X \in \mathfrak{X}(M)$ and $\omega \in \Omega^1(N)$. This may be put in a more symmetric way if we introduce the push-forward F_* of a 1-form by a diffeomorphism F. Namely, for $\omega \in \Omega^1(M)$ one defines

$$F_*\omega := (F^{-1})^*\omega = (F^*)^{-1}\omega.$$

One then has

$$F_*(\omega, X) = (F_*\omega, F_*X) \tag{8.16}$$

for all $X \in \mathfrak{X}(M)$ and $\omega \in \Omega^1(M)$. Using the notation with contraction we have

$$F_*(\iota_X\omega) = \iota_{F_*X}F_*\omega.$$

A large class of 1-forms arise by differentiating functions. Namely, if f is a $\mathcal{C}^k$-function on M, then $\mathrm{d}_q f\colon T_qM \to \mathbb{R}$ can also be read as an element of T_q^*M and hence defines a 1-form of class $\mathcal{C}^{k-1}$ denoted by $\mathrm{d}f$. We prefer not to bother with the shift in k, so for simplicity we only consider smooth manifolds now. The map

$$\begin{array}{rccc} \mathrm{d}\colon & \mathcal{C}^\infty(M) & \to & \Omega^1(M) \\ & f & \mapsto & \mathrm{d}f \end{array}$$

(a.k.a. the de Rham differential) is $\mathbb{R}$-linear and satisfies the Leibniz rule[25]

$$\mathrm{d}(fg) = \mathrm{d}f\, g + f\, \mathrm{d}g$$

for all $f, g \in \mathcal{C}^\infty(M)$. Elements in the image of d are called exact 1-forms. If X is a vector field, we also clearly have

$$\iota_X \mathrm{d}f = (\mathrm{d}f, X) = X(f) = \mathsf{L}_X f.$$

By the chain rule, the de Rham differential commutes with pullbacks:

Lemma 8.63. *If $F\colon M \to N$ is a smooth map, then*

$$F^*\mathrm{d}f = \mathrm{d}F^*f$$

for all $f \in \mathcal{C}^\infty(N)$.

Proof. We have

$$\mathrm{d}_q(F^*f) = \mathrm{d}_q(f \circ F) = \mathrm{d}_{F(q)}f\mathrm{d}_qF = (\mathrm{d}_qF)^*\mathrm{d}_{F(q)}f = (F^*\mathrm{d}f)_q.$$

□

If F is a diffeomorphism, then we also have

$$F_*\mathrm{d}f = \mathrm{d}F_*f$$

for all $f \in \mathcal{C}^\infty(M)$.

---

[25]This innocent looking formula requires some explanation. One way to get it is to observe that for every vector field X and every function f we have $X(f)(q) = (\mathrm{d}_qf, X_q)$. From $X(fg) = X(f)g + fX(g)$ evaluated at q, noticing that we have enough vector fields to span all directions, we get $\mathrm{d}_q(fg) = \mathrm{d}_qfg(q) + f(q)\mathrm{d}_qg$. Another way of proving the Leibniz rule consists in observing that on open subsets of $\mathbb{R}^n$ it is just the rule for deriving a product. Then one uses Lemma 8.63 to transfer it from the images of the charts to the manifold.

If M is Hausdorff, we may define the Lie derivative of a 1-form in complete analogy with what we did for functions and vector fields in subsection 7.4 (see also Remark 7.30). Namely, we set

$$\mathsf{L}_X\omega := \frac{\partial}{\partial t}\Big|_{t=0}(\Phi_{-t}^X)_*\omega = \frac{\partial}{\partial t}\Big|_{t=0}(\Phi_t^X)^*\omega,$$

for $X \in \mathfrak{X}(M)$ and $\omega \in \Omega^1(M)$. The Lie derivative is again $\mathbb{R}$-linear. Moreover, we have

$$\frac{\partial}{\partial t}\Big|_{t=s}(\Phi_{-t}^X)_*\omega = (\Phi_{-s}^X)_*\mathsf{L}_X\omega = \mathsf{L}_X(\Phi_{-s}^X)_*\omega$$

for all s for which the flow is defined. The proof is exactly the same as in Lemma 7.31. As a corollary, we then have that ω is preserved by the flow of X, i.e.,

$$(\Phi_t^X)_*\omega = \omega\ \forall t \quad \text{if and only if} \quad \mathsf{L}_X\omega = 0.$$

If f is a function, the properties of pullback imply

$$\mathsf{L}_X(f\omega) = \mathsf{L}_Xf\,\omega + f\,\mathsf{L}_X\omega. \tag{8.17}$$

Moreover, Lemma 8.63 implies that

$$\boxed{\mathsf{L}_X\mathrm{d}f = \mathrm{d}\mathsf{L}_Xf} \tag{8.18}$$

for every vector field X and every function f. From (8.16) with F the flow of a vector field Y, we also get

$$\mathsf{L}_Y(\omega, X) = (\mathsf{L}_Y\omega, X) + (\omega, \mathsf{L}_YX), \tag{8.19}$$

or all $X \in \mathfrak{X}(M)$ and $\omega \in \Omega^1(M)$. Using the notation with contraction and Lemma 7.29, we can rewrite this as

$$\boxed{\iota_X\mathsf{L}_Y\omega - L_Y\iota_X\omega = \iota_{[X,Y]}\omega} \tag{8.20}$$

Finally if F is a diffeomorphism, by the properties of the push-forward and by Proposition 7.28, we get

$$\mathsf{L}_XF^*\omega = F^*\mathsf{L}_{F_*X}\omega, \tag{8.21}$$

for all $X \in \mathfrak{X}(M)$ and all $\omega \in \Omega^1(N)$. Equivalently,

$$F_*\mathsf{L}_X\omega = \mathsf{L}_{F_*X}F_*\omega,$$

for all $X \in \mathfrak{X}(M)$ and all $\omega \in \Omega^1(M)$. Finally, we have

$$\mathsf{L}_{X+Y}\omega = \mathsf{L}_X\omega + \mathsf{L}_Y\omega,$$
$$\mathsf{L}_X\mathsf{L}_Y\omega - \mathsf{L}_Y\mathsf{L}_X\omega = \mathsf{L}_{[X,Y]}\omega,$$

for any two vector fields X and Y. If we work in the image of a chart, the proof is exactly as in the local chart proof of Lemma 7.35. Alternatively, we may observe that since L commutes with d, in the case of exact 1-forms these identities immediately follow from those for functions. The next remark is that in a local chart every 1-form is a linear combination of products of a function and an exact 1-form, so the result for general 1-forms in a chart image follows from linearity and from the Leibniz rule. The global result follows from the commutativity of L with pullbacks. (If a partition of unity subordinate to an atlas is available, one can easily see that also globally a 1-form is a linear combination of a product of a function and an exact 1-form.)

If V is an open subset of $\mathbb{R}^n$, then we can consider the differentials $\mathrm{d}x^i$ of the coordinate functions x^i. Notice that we have (using the notation of Remark 7.16 on page 38, which comes in very handy here)

$$\iota_{\partial_j}\mathrm{d}x^i = \frac{\partial x^i}{\partial x^j} = \delta^i_j.$$

Since $(\partial_i)_{i=1,\dots,n}$ is a basis of $\mathfrak{X}(V)$ as a module over $\mathcal{C}^\infty(V)$, we see that $(\mathrm{d}x^i)_{i=1,\dots,n}$ is the dual basis of $\Omega^1(V)$. It follows that for every $\omega \in \Omega^1(V)$, we have

$$\omega = \sum_{i=1}^n \omega_i \mathrm{d}x^i,$$

where $\omega_1, \dots, \omega_n$ are uniquely determined functions. If f is a function on U, we also clearly have

$$\mathrm{d}f = \sum_{i=1}^n \partial_i f\, \mathrm{d}x^i.$$

The Lie derivative is easy to compute in local coordinates:

Lemma 8.64. *Let U be an open subset of $\mathbb{R}^n$. If $\omega = \sum_{i=1}^n \omega_i \mathrm{d}x^i$ is a 1-form and $X = \sum_{i=1}^n X^i\partial_i$ is a vector field, then $\mathsf{L}_X\omega = \sum_{i=1}^n (\mathsf{L}_X\omega)_i \mathrm{d}x^i$ with*

$$(\mathsf{L}_X\omega)_i = \sum_{j=1}^n (X^j\partial_j\omega_i + \omega_j\partial_i X^j).$$

Proof. By (8.17) and (8.18), we have

$$\mathsf{L}_X \sum_{i=1}^n \omega_i \mathrm{d}x^i = \sum_{i=1}^n (\mathsf{L}_X\omega_i \mathrm{d}x^i + \omega_i \mathrm{d}\mathsf{L}_X x^i).$$

From $\mathsf{L}_X\omega_i = \sum_{j=1}^n X^j\partial_j\omega_i$, $\mathsf{L}_X x^i = X^i$ and $\mathrm{d}X^i = \sum_{j=1}^n \partial_j X^i \mathrm{d}x^j$ we get the result. □

Remark 8.65. One can define the Lie derivative of a 1-form also on a non-Hausdorff manifold simply using the above formula on the representations in each chart. Namely, let ω and X be a 1-form and a vector field on M. Let ω_α and X_α be their representations in the atlas $\{(U_\alpha,\phi_\alpha)\}_{\alpha\in I}$. Since the intersection chart images $\phi_\alpha(U_\alpha\cap U_\beta)$ are Hausdorff, as subsets of $\mathbb{R}^n$, we can use (8.21) to conclude that the collection $\mathsf{L}_{X_\alpha}\omega_\alpha$ represents a 1-form that we call the Lie derivative of ω by X and denote by $\mathsf{L}_X\omega$.

If $F\colon M\to U$ is a smooth map to an open subset of $\mathbb{R}^n$, by using all the properties above we have

$$F^*\omega=\sum_{i=1}^n F^*\omega_i\,\mathrm{d}F^*x^i=\sum_{i=1}^n \omega_i\circ F\,\mathrm{d}F^i$$

with $F^i:=x^i\circ F$ the ith component of the map F.

8.4. **The tensor bundle.** If E is a vector bundle over M, we define $T^k_s(E)$ as the vector bundle whose fiber at q is $T^k_s(E_q)$. (We use the notations of Appendix B.1.) Namely, to an adapted atlas $\{(\widetilde{U}_\alpha,\widetilde{\phi}_\alpha)\}_{\alpha\in I}$ of E over the trivializing atlas $\{(U_\alpha,\phi_\alpha)\}_{\alpha\in I}$ of M, we associate the atlas $\{(\widehat{U}_\alpha,\widehat{\phi}_\alpha)\}_{\alpha\in I}$ of $T^k_s(E)$ with $\widehat{U}_\alpha=\pi^{-1}_{T^k_s(E)}(U_\alpha)=\cup_{q\in U_\alpha}T^k_s(E_q)$ and

$$\begin{array}{cccc}\widehat{\phi}_\alpha\colon & \widehat{U}_\alpha & \to & \mathbb{R}^n\times T^k_s(\mathbb{R}^r)\\ & (q,\omega\in T^k_s(E_q)) & \mapsto & (\phi_\alpha(q),(A_\alpha(q))^k_s\,\omega)\end{array}$$

where we identify $T^k_s(\mathbb{R}^r)$ with $\mathbb{R}^{r(k+s)}$. It follows that we have transitions maps

$$\widehat{\phi}_{\alpha\beta}(x,u)=(\phi_{\alpha\beta}(x),(A_{\alpha\beta}(\phi_\alpha^{-1}(x)))^k_s\,u).$$

We specialize this construction to the case when the vector bundle E is the tangent bundle TM. In this case, the transition maps are

$$\widehat{\phi}_{\alpha\beta}(x,u)=(\phi_{\alpha\beta}(x),(\mathrm{d}_x\phi_{\alpha\beta})^k_s\,u).$$

A section of $T^k_sM:=T^k_s(TM)$ is called a tensor field of type (k,s). If $\{(U_\alpha,\phi_\alpha)\}_{\alpha\in I}$ is an atlas on M, a (k,s)-tensor field Ξ is represented by a collection of maps $\Xi_\alpha\colon\phi_\alpha(U_\alpha)\to T^k_s(\mathbb{R}^n)=\mathbb{R}^{n(k+s)}$ such that

$$\Xi_\beta(\phi_{\alpha\beta}(x))=(\mathrm{d}_x\phi_{\alpha\beta})^k_s\,\Xi_\alpha(x)\tag{8.22}$$

for all $\alpha,\beta\in I$ and for all $x\in\phi_\alpha(U_\alpha\cap U_\beta)$.

Tensor fields of type $(0,s)$ are also called covariant tensor fields of order s, whereas tensor fields of type $(k,0)$ are also known as contravariant tensor fields of order k. Notice that $(0,0)$-tensor fields are the same as functions, $(1,0)$-tensor fields are the same as vector fields and $(0,1)$-tensor fields are the same as 1-forms. The tensor product of tensors induces, by pointwise multiplication, the tensor product of tensor fields

$$\Gamma(T^{k_1}_{s_1}M)\times\Gamma(T^{k_2}_{s_2}M)\to\Gamma(T^{k_1+k_2}_{s_1+s_2}M),$$

which is bilinear over the ring of functions. If Ξ_1 and Ξ_2 are tensor fields, their tensor product is denoted by $\Xi_1 \otimes \Xi_2$.

Remark 8.66. The notion of tensor product, which we have recalled in Appendix B for vector spaces, holds more general for modules. If M_1 and M_2 are modules over a ring R, one uses the notation $M_1 \otimes_R M_2$ for their tensor product (as M_1 and M_2 may often be regarded as modules for some subring). The tensor product of tensor fields, being bilinear over $C^\infty(M)$, induces a $C^\infty(M)$-linear map

$$\Gamma(T^{k_1}_{s_1}M) \otimes_{C^\infty(M)} \Gamma(T^{k_2}_{s_2}M) \to \Gamma(T^{k_1+k_2}_{s_1+s_2}M).$$

In particular, we have the canonical $C^\infty(M)$-isomorphism

$$T^k_s\mathfrak{X}(M) \simeq \Gamma(T^k_sM),$$

where T^k_s denotes on the left hand side the $C^\infty(M)$-tensor power and on the right hand side the fiberwise $\mathbb{R}$-tensor power. As tensor fields are tensors for the $C^\infty(M)$-module $\mathfrak{X}(M)$, they are often simply called tensors.

If $F\colon M \to N$ is a diffeomorphism, then we can define the push-forward F_* of tensor fields by pointwise application of (B.4): namely, if Ξ is a tensor field of type (k, s), we define

$$(F_*\Xi)_y := (\mathrm{d}_{F^{-1}(y)}F)^k_s\, \Xi_{F^{-1}(y)},$$

for $y \in N$. Notice that in the case of functions, vector fields and 1-forms the push-forward coincides with the one that we have already defined. Moreover, for any two tensor fields Ξ_1 and Ξ_2 we have

$$F_*(\Xi_1 \otimes \Xi_2) = F_*\Xi_1 \otimes F_*\Xi_2. \tag{8.23}$$

Also notice that if G is a diffeomorphism from N to Z, then $(G \circ F)_* = G_* \circ F_*$.

Using push-forwards, we can write the representation of a tensor field Ξ in a chart (U, ϕ_U) as $\Xi_U = (\phi_U)_*\Xi|_U$. Moreover, we can rewrite the compatibility equation (8.22) as

$$\Xi_\beta = (\phi_{\alpha\beta})_*\Xi_\alpha$$

for all $\alpha, \beta \in I$, where, by abuse of notation, Ξ_α denotes here the restriction of Ξ_α to $\phi_\alpha(U_\alpha \cap U_\beta)$ and Ξ_β denotes the restriction of Ξ_β to $\phi_\beta(U_\alpha \cap U_\beta)$.

If M is Hausdorff, we may define the Lie derivative of a tensor field in complete analogy with what we did for functions, vector fields, densities and 1-forms (see also Remark 7.30). For simplicity we are now going to consider only smooth manifolds. Namely, we set

$$\mathsf{L}_X\Xi := \frac{\partial}{\partial t}\Big|_{t=0}(\Phi^X_{-t})_*\,\Xi,$$

for $X \in \mathfrak{X}(M)$ and Ξ a tensor field. The Lie derivative is again $\mathbb{R}$-linear. Moreover, we have

$$\frac{\partial}{\partial t}\Big|_{t=s}(\Phi^X_{-t})_*\Xi = (\Phi^X_{-s})_*\mathsf{L}_X\Xi = \mathsf{L}_X(\Phi^X_{-s})_*\Xi$$

for all s for which the flow is defined. The proof is exactly the same as in Lemma 7.31. This implies the important

Corollary 8.67. *A tensor field Ξ is preserved by the flow of a vector field X, i.e.,*

$$(\Phi^X_t)_*\Xi = \Xi \ \forall t \quad \textit{if and only if} \quad \mathsf{L}_X\Xi = 0.$$

If Ξ_1 and Ξ_2 are tensor fields, then (8.23) implies the Leibniz rule

$$\mathsf{L}_X(\Xi_1 \otimes \Xi_2) = \mathsf{L}_X\Xi_1 \otimes \Xi_2 + \Xi_1 \otimes \mathsf{L}_X\Xi_2. \tag{8.24}$$

Finally if $F\colon M \to N$ is a diffeomorphism, by the properties of the push-forward and by Proposition 7.28, we get

$$F_*\mathsf{L}_X\Xi = \mathsf{L}_{F_*X}F_*\Xi, \tag{8.25}$$

for all $X \in \mathfrak{X}(M)$ and all tensor fields Ξ. Finally, we have

$$\mathsf{L}_{X+Y}\Xi = \mathsf{L}_X\Xi + \mathsf{L}_Y\Xi,$$
$$\mathsf{L}_X\mathsf{L}_Y\Xi - \mathsf{L}_Y\mathsf{L}_X\Xi = \mathsf{L}_{[X,Y]}\Xi.$$

for any two vector fields X and Y. Again, in a chart image this may be proved either as in the local chart proof of Lemma 7.35 or observing that a tensor field is a linear combination of tensor products of functions, vector fields and exact 1-forms, as explained in the following remark.

Remark 8.68. On an open subset U of $\mathbb{R}^n$ we have a basis of the $\mathcal{C}^\infty(U)$-module of (k,s)-tensor fields given by

$$\partial_{i_1} \otimes \cdots \otimes \partial_{i_k} \otimes \mathrm{d}x^{j_1} \otimes \cdots \otimes \mathrm{d}x^{j_s}.$$

As a consequence, a tensor field on U is a finite linear combination of tensor products of functions, vector fields (∂_i) and exact 1-forms ($\mathrm{d}x^i$). This is also true globally for tensor fields on a manifold with partitions of unity (such as a Hausdorff second-countable manifold), for we can write a tensor field Ξ as

$$\Xi = \sum_{i\in J}\rho_i\,\Xi|_{U_{\alpha_i}} = \sum_{i\in J}\rho_i\,(\phi^{-1}_{\alpha_i})_*\Xi_{\alpha_i},$$

where $\{\rho_i\}_{i\in J}$ is a partition of unity subordinate to the atlas $\{(U_\alpha,\phi_\alpha)\}_{\alpha\in I}$. Hence by (8.24) it is enough to know the action of the Lie derivative on functions and on vector fields as the action on exact 1-forms can be deduced from equation (8.18).

Remark 8.69 (Pullback of covariant tensors fields)**.** The push-forward of tensor fields requires having a diffeomorphism. If Ξ is a $(0,s)$-tensor field on N, however, we can pull it back by any map $F\colon M\to N$ by

$$(F^*\Xi)_q := ((\mathrm{d}_qF)^*)^{\otimes s}\,\Xi_{F(q)},$$

for $q\in M$. Notice that in the case of functions and 1-forms this coincides with the already defined pullback. Also notice that, in case F is a diffeomorphism, then $F^*=(F^{-1})_*$.

8.5. **Digression: Riemannian metrics.** A very important example of tensor field is the so-called Riemannian tensor, which is used to introduce geometric properties on a manifold (this specialized part of the theory of manifolds goes under the name of differential geometry).

Let us start recalling the notion of (Euclidean) length in $\mathbb{R}^n$. Let $\gamma\colon I\to\mathbb{R}^n$ be a piecewise differentiable curve.[26] Then one sets

$$\ell(\gamma)=\int_I\sqrt{\eta(\dot\gamma(t),\dot\gamma(t))}\,dt,$$

where η denotes the Euclidean scalar product: $\eta(v,w)=\sum_{i=1}^n v^iw^i$. This definition extends the notion of Euclidean length of segments by Pythagoras' theorem. Actually, if one approximates γ by a piecewise linear curve, whose length is defined as the sum of the lengths of its composing segments, then one gets $\ell(\gamma)$ in the limit. The main idea of Riemannian geometry consists in replacing the fixed bilinear form η by a point-depending one.

Definition 8.70. A Riemannian metric on a differentiable manifold is a positive definite, symmetric $(0,2)$-tensor field. If g is a Riemannian metric on M, the pair (M,g) is called a Riemannian manifold.

More explictly, a Riemannian metric g on M is a section of $T^0_2M=T^*M\otimes T^*M$ such that, for each $q\in M$,

$$g_q(v,w)=g_q(w,v)\quad\text{for all } v,w\in T_qM$$

and

$$g_q(v,v)>0\quad\text{for all } v\in T_qM\setminus\{0\}.$$

[26]This means that γ is continuous and that I can be written as a disjoint finite union of intervals I_k such that γ is continuously differentiable in the interior of each I_k and the limits of its derivatives to the endpoints of each I_k are finite.

Remark 8.71. In an open subset of $\mathbb{R}^n$ a metric g can be expanded as $g = \sum_{i=1}^n g_{ij}\,\mathrm{d}x^i \otimes \mathrm{d}x^j$, where for each x the matrix with entries $g_{ij}(x)$ is symmetric and positive definite. Typically one omits writing the tensor product symbol, so a metric is simply written as

$$g = \sum_{i=1}^n g_{ij}\,\mathrm{d}x^i \mathrm{d}x^j. \tag{8.26}$$

The Euclidean metric η then reads $\eta = \sum_{i=1}^n (\mathrm{d}x^i)^2$. If one uses coordinate functions without indices, e.g., $x, y, z, \ldots$, then one customarily writes $\mathrm{d}x^2$ instead of $(\mathrm{d}x)^2$, and so on, so the Euclidean metric reads $\eta = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 + \cdots$.

Remark 8.72. Another notation for a metric g is by the "infinitesimal line length" $\mathrm{d}s^2$. Namely, instead of (8.26) one writes

$$\mathrm{d}s^2 = \sum_{i=1}^n g_{ij}\,\mathrm{d}x^i \mathrm{d}x^j.$$

In the case of the Euclidean metric,

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 + \cdots,$$

this takes the form of an "infinitesimal Pythagorean theorem." This notation also fits well with the general notion of product Riemannian metric, of which this was an example. More generally, if g_M and g_N are Riemannian metrics on M and N, respectively, then one defines

$$g_{M\times N} := \pi_M^* g_M + \pi_N^* g_N,$$

where π_M and π_N are the projections from $M \times N$ to M and N. It is readily verified that $g_{M\times N}$ is also a Riemannian metric. If we write $\mathrm{d}s^2$ instead of g and let the pullbacks be understood, then we have

$$\mathrm{d}s^2_{M\times N} = \mathrm{d}s^2_M + \mathrm{d}s^2_N,$$

another instance of the "infinitesimal Pythagorean theorem."

Notice that Riemannian metrics are a very general concepts:

Lemma 8.73. *Every manifold with partitions of unity (e.g., a Hausdorff, second-countable one) admits a Riemannian metric.*

Proof. Let $\{\rho_j\}_{j\in J}$ be a partition of unity subordinate to an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ of M. Then define $g = \sum_{j\in J} \rho_j\,(\phi_{\alpha_j}^{-1})_*\eta$, where η is the Euclidean metric and $\rho_j\,(\phi_{\alpha_j}^{-1})_*\eta$ is extended by zero outside of U_{α_j}. Since a convex linear combination of positive definite bilinear forms is positive definite, it is readily checked that g is a Riemannian metric. $\square$

Definition 8.74. The Riemannian length of a piecewise differentiable curve $\gamma\colon I \to \mathbb{R}^n$ is

$$\ell_g(\gamma) := \int_I \sqrt{g_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t))}\, dt.$$

Lemma 8.75. *The length of a curve does not depend on its parametrization.*

Proof. Let $\phi\colon I \to J$ be a diffeomorphism and let $\widetilde\gamma := \gamma \circ \phi^{-1}\colon J \to M$ be a different parametrization of γ. Then

$$\begin{aligned}\ell_g(\gamma) = \int_I \sqrt{g_{\widetilde\gamma(\phi(t))}(\dot\phi(t)\dot{\widetilde\gamma}(\phi(t)), \dot\phi(t)\dot{\widetilde\gamma}(\phi(t)))}\, dt =\\ = \int_I \sqrt{g_{\widetilde\gamma(\phi(t))}(\dot{\widetilde\gamma}(\phi(t)), \dot{\widetilde\gamma}(\phi(t)))}\, |\dot\phi(t)|\, dt =\\ = \int_J \sqrt{g_{\widetilde\gamma(s)}(\dot{\widetilde\gamma}(s), \dot{\widetilde\gamma}(s))}\, ds = \ell_g(\widetilde\gamma),\end{aligned}$$

with $s = \phi(t)$. □

For those who know calculus of variations, we can introduce the

Definition 8.76. A geodesic on (M, g) from q_1 to q_2 is an extremal path for ℓ_g on the space of immersed curves joining q_1 to q_2.

One can use the calculus of variations to show that a geodesic is a solution of a second-order differential equation. If M is Hausdorff and the endpoints q_1 and q_2 are close enough, there is then a unique geodesic joining them that minimizes length; also notice that the length of an immersed curve is necessarily strictly positive. This result is fundamental to prove the

Theorem 8.77. *Let (M, g) be a connected Hausdorff Riemannian manifold. Let*

$$d_g(q_1, q_2) := \inf_\gamma \ell_g(\gamma),$$

where the infimum is taken over the set of all piecewise differentiable curves joining q_1 to q_2. Then d_g is a distance on M which induces the same topology as the atlas topology.

Sketch of the proof. The function d_g is clearly symmetric and nonnegative. We also have $d_g(q, q) = 0$ as we can take the constant path to join q to q. The triangle inequality also follows immediately from the definition: if we have three points q_1, q_2 and q_3, then any curve γ_1 that joins q_1 to q_2 may be joined to any curve γ_2 that joins q_2 to q_3 to produce a curve γ that joins q_1 to q_3. By additivity of the integral we have $\ell_g(\gamma) = \ell_g(\gamma_1) + \ell_g(\gamma_2)$. On the other hand, by definition we have $d_g(q_1, q_3) \leq \ell_g(\gamma)$. Taking the infimum over γ_1 and over γ_2, we finally get $d_g(q_1, q_3) \leq d_g(q_1, q_2) + d_g(q_2, q_3)$. What is left to prove is that if $q_1 \neq q_2$, then $d_g(q_1, q_2) > 0$. The idea is to take the preimage under a chart map of a ball around q_1 not containing q_2 such that there is a unique length-minimizing geodesic joining q_1 to any point inside this ball. Since every path joining q_1 to q_2 must go through this ball, the infimum of the lengths is not smaller than the length of a geodesic inside the ball, which is strictly positive. We leave to the reader to check that the topology induced by this metric is the same as the original topology. For more details, we refer to any book on differential geometry, e.g., [4, Prop. 8.19]. □

Remark 8.78 (The Riemannian density)**.** Riemannian metrics can be used not only to define lengths, but also volumes. First, observe that if g is a $(0,2)$-tensor field then its determinant $\det g$ is a 2-density. More precisely, pick an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ and let g_α be the representation of g in the chart α. Then

$$((\mathrm{d}_x\phi_{\alpha\beta})^*)^{\otimes 2} g_\beta(\phi_{\alpha\beta}(x)) = g_\alpha(x)$$

for all $\alpha, \beta \in I$ and for all $x \in \phi_\alpha(U_\alpha \cap U_\beta)$. As $g_\alpha(x)$ is an element of $((\mathbb{R}^n)^*)^{\otimes 2}$, i.e., a bilinear form on $\mathbb{R}^n$ or, more concretely, an $n \times n$ matrix, we can take its determinant. We then have

$$(\det \mathrm{d}_x\phi_{\alpha\beta})^2 \det g_\beta(\phi_{\alpha\beta}(x)) = \det g_\alpha(x)$$

for all $\alpha, \beta \in I$ and for all $x \in \phi_\alpha(U_\alpha \cap U_\beta)$. This shows that the collection $\{\det g_\alpha\}_{\alpha\in I}$ represents a 2-density. If g is a Riemannian metric, then $\det g_\alpha > 0$ for all α, so we can take its square root, which defines a density since

$$|\det \mathrm{d}_x\phi_{\alpha\beta}| \sqrt{\det g_\beta(\phi_{\alpha\beta}(x))} = \sqrt{\det g_\alpha(x)}$$

for all $\alpha, \beta \in I$ and for all $x \in \phi_\alpha(U_\alpha \cap U_\beta)$. We denote by v_g the positive density that in the chart (U_α, ϕ_α) is represented by $\sqrt{\det g_\alpha}$. This is called the Riemannian density. Finally, we can define the Riemannian volume of M as

$$\mathrm{Vol}_g(M) := \int_M v_g \in \mathbb{R}_{>0} \cup \{+\infty\}.$$

Remark 8.79. If U is an open subset of $\mathbb{R}^n$ and we take the restriction to U of the Euclidean metric η as a Riemannian metric, then $\mathrm{Vol}_\eta(U) = \int_U d^n x$, the usual Lebesgue volume.

Remark 8.80. Let $F\colon M \to N$ be a diffeomorphism and let g be a Riemannian metric on M. Then we have $F_* v_g = v_{F_* g}$, where on the left hand side we use the push-forward of densities introduced in subsubsection 8.2.4.

Remark 8.81 (The divergence)**.** Since the Riemannian density v_g is positive definite, it defines a divergence operator as explained in subsubsection 8.2.6. For simplicity of notation, we write div_g instead of div_{v_g}. Notice that by (8.14), in an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ we have

$$(\mathrm{div}_g X)_\alpha = \sum_{i=1}^n \frac{1}{\sqrt{\det g_\alpha}}\,\partial_i\left(\sqrt{\det g_\alpha}\,X^i_\alpha\right). \tag{8.27}$$

The Riemannian divergence operator will play a central role in the theorem of Gauss.

Remark 8.82 (Gradient and Laplacian)**.** A $(0,2)$-tensor g establishes a linear map $T_qM \to T^*_qM$ at each $q \in M$ by $v \mapsto g_q(v,\ )$. These maps may be assembled into a morphism

$$\Phi_g\colon TM \to T^*M$$

as in Definition 8.16. This morphism can be used to push forward vector fields to 1-forms. If g is a Riemannian metric, this is actually an isomorphism. In particular, to every function f we can associate a uniquely determined vector field

$$\mathrm{grad}_g f := \Phi_g^{-1}\mathrm{d}f,$$

called the gradient of f. The gradient of f is uniquely characterized by the property that

$$g\left(\mathrm{grad}_g f, Y\right) = Y(f)$$

for every vector field Y. The divergence of $\mathrm{grad}_g f$ is called the Laplacian of f and is denoted by $\Delta_g f$, so

$$\Delta_g f := \mathrm{div}_g \mathrm{grad}_g f.$$

In an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$, we have

$$\left(\mathrm{grad}_g f\right)_\alpha = \sum_{ij} g^{ij}_\alpha\,\partial_i f_\alpha\,\partial_j,$$

where (g^{ij}_α) denotes the inverse of the matrix $((g_\alpha)_{ij})$. The Laplacian then reads

$$(\Delta_g f)_\alpha = \sum_{ij} \frac{1}{\sqrt{\det g_\alpha}}\,\partial_j\left(\sqrt{\det g_\alpha}\,g^{ij}_\alpha\,\partial_i f_\alpha\right).$$

Notice that on an open subset of $\mathbb{R}^n$ with Euclidean metric we recover the usual formulae from calculus:

$$\operatorname{grad} f = \sum_i \partial_i f\, \partial_i \quad \text{and} \quad \Delta f = \sum_i \partial_i^2 f.$$

Finally, recall that, as in Remark 8.69, covariant tensor fields can be pulled back. The pullback of a Riemannian metric is in general degenerate, so not a Riemannian metric. However, if $\iota\colon S \to M$ is an immersion, then $\mathrm{d}_q\iota(T_qS)$ is a subspace of $T_{\iota(q)}M$. The restriction of a positive definite, symmetric bilinear form to a subspace is still positive definite. Hence, if g is a Riemannian metric on M, then $g_S := \iota^*g$ is a Riemannian metric on S. In particular, this is the case when S is a submanifold and ι the inclusion map. The Riemannian metric g_S is called the restriction of g to S. The volume of S is then defined as

$$\mathrm{Vol}_g(S) := \int_S v_{g_S}.$$

Remark 8.83. The Riemannian density v_g is positive, so it defines a positive measure μ_{v_g} on M, see Remark 8.38. If U is an open subset of M, then $\mathrm{Vol}_g(U) = \mu_{v_g}(U)$. If on the other hand S is a submanifold of strictly lower dimension, we have $\mu_{v_g}(S) = 0$, but $\mathrm{Vol}_g(S) > 0$. The volume defined by the induced metric generalizes, e.g., the notion of areas of surfaces in three-dimensional Euclidean space.

Example 8.84 (Length)**.** Let $\gamma\colon I \to M$ be an immersed curve. If g is a Riemannian metric on M, the induced metric g_γ is given by

$$g_\gamma(t) = g_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t))\, \mathrm{d}t^2.$$

Hence

$$\mathrm{Vol}_g(\gamma) = \int_I \sqrt{g_\gamma} = \ell_g(\gamma).$$

Example 8.85 (Area)**.** Let S be an open subset of $\mathbb{R}^2$, M an open subset of $\mathbb{R}^n$ and $\sigma\colon S \to M$ an immersion. We denote by x^i the coordinates on S and by z^μ the coordinates on M. We then have $\sigma^*z^\mu = \sigma^\mu(x)$, where the σ^μs are components of the map σ. It follows that

$$\sigma^*\mathrm{d}z^\mu = \mathrm{d}\sigma^*z^\mu = \sum_{i=1}^2 \partial_i\sigma^\mu\, \mathrm{d}x^i.$$

Thus,

$$(g_S)_x = \sum_{i,j=1}^2 (\partial_i\sigma, \partial_j\sigma)_x\, \mathrm{d}x^i \mathrm{d}x^j,$$

with $(\partial_i\sigma,\partial_j\sigma)_x := \sum_{\mu,\nu=1}^n g_{\mu\nu}(\sigma(x))\,\partial_i\sigma^\mu(x)\,\partial_j\sigma^\nu(x)$. Hence

$$\mathrm{Vol}_g(S) = \int_S \sqrt{(\partial_1\sigma,\partial_1\sigma)_x(\partial_2\sigma,\partial_2\sigma)_x - ((\partial_1\sigma,\partial_2\sigma)_x)^2}\, d^2x.$$

Example 8.86. Pullback (or push-forward) by a diffeomorphism can also be understood as a change of variables. Suppose for example that we want to use polar coordinates in $\mathbb{R}^2$. We actually have a map

$$\mathbb{R}_{>0}\times(0,2\pi)\to\mathbb{R}^2\setminus D,\quad (r,\theta)\mapsto(r\cos\theta, r\sin\theta),$$

where $D=\{(x,y)\in\mathbb{R}^2 \mid y=0,\ x>0\}$. This map can be extended to a diffeomorphism

$$F\colon \mathbb{R}_{>0}\times S^1\to\mathbb{R}^2\setminus\{0\}.$$

Denoting by x and y the coordinates on $\mathbb{R}^2\setminus\{0\}$, we have

$$F^*\mathrm{d}x = \mathrm{d}F^*x = \cos\theta\,\mathrm{d}r - r\sin\theta\,\mathrm{d}\theta,$$
$$F^*\mathrm{d}y = \mathrm{d}F^*y = \sin\theta\,\mathrm{d}r + r\cos\theta\,\mathrm{d}\theta.$$

With the notations of Remark 8.71, we write the Euclidean metric as $\eta=\mathrm{d}x^2+\mathrm{d}y^2$. Since F^* is an algebra morphism, we have

$$F^*\mathrm{d}x^2 = \cos^2\theta\,\mathrm{d}r^2 - r\sin\theta\cos\theta\,(\mathrm{d}r\mathrm{d}\theta+\mathrm{d}\theta\mathrm{d}r) + r^2\sin^2\theta\,\mathrm{d}\theta^2,$$
$$F^*\mathrm{d}y^2 = \sin^2\theta\,\mathrm{d}r^2 + r\sin\theta\cos\theta\,(\mathrm{d}r\mathrm{d}\theta+\mathrm{d}\theta\mathrm{d}r) + r^2\cos^2\theta\,\mathrm{d}\theta^2.$$

Thus, the Euclidean metric in polar coordinates reads

$$F^*\eta = \mathrm{d}r^2 + r^2\mathrm{d}\theta^2.$$

This means that in these coordinates the metric is diagonal with diagonal entries 1 and r^2, so $\det F^*\eta = r^2$ and by (8.27) the polar coordinate expression for the divergence of a vector field $X = X_r\frac{\partial}{\partial r} + X_\theta\frac{\partial}{\partial\theta}$ is

$$\mathrm{div}_{F^*\eta}X = \frac{1}{r}\left(\frac{\partial}{\partial r}(rX_r)+\frac{\partial}{\partial\theta}(rX_\theta)\right) = \frac{1}{r}\frac{\partial}{\partial r}(rX_r)+\frac{\partial}{\partial\theta}X_\theta.$$

If f is a function, and $\mathrm{d}f = \partial_r f\mathrm{d}r + \partial_\theta f\mathrm{d}\theta$ its differential, by Remark 8.82 we compute its gradient as $\mathrm{grad}_{F^*\eta}f = \partial_r f\partial_r + \frac{1}{r^2}\partial_\theta f\partial_\theta$. It then follows that its Laplacian is

$$\Delta_{F^*\eta}f = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 f}{\partial\theta^2}.$$

Remark 8.87 (Pseudo-Riemannian metrics)**.** Notice that for most of the constructions in this subsection what was really needed was just that g be nondegenerate; namely, that g_q establish an isomorphism between T_qM and T_q^*M at each q or, equivalently, that $\det g_\alpha\neq 0$ in every representation. A symmetric $(0,2)$-tensor field with this property is called a pseudo-Riemannian metric. A pseudo-Riemannian metric g

defines a positive density v_g by $\sqrt{|\det g_\alpha|}$ in each chart. By this, one may define the notions of volume of regions and of divergence of vector fields. Since we have an isomorphism $TM \to T^*M$, we may also define the gradient and the "Laplacian" of functions (in the case when g has exactly one positive eigenvalue, or exactly one negative eigenvalue, this operator is usually called d'Alembertian). One can also define a "length" functional $\ell_g(\gamma) := \int_I \sqrt{|g_{\gamma(t)}(\dot\gamma(t),\dot\gamma(t))|}\,dt$, which may now however vanish on curves joining distinct points and does no longer induce a distance function. One can also define geodesics as extremal immersions for this functional, and they still turn out to be solutions of a second-order differential equation. Notice on the other hand that, unlike in Lemma 8.73, the existence of a pseudo-Riemannian metric, with a prescribed signature, is in general not guaranteed. A particular case, of great importance in physics (general relativity), is that of a Lorentzian metric, i.e., a pseudo-Riemannian metric g such that g_q has exactly one positive eigenvalue (or, according to another convention, exactly one negative eigenvalue) at each q. The standard example on an open subset of $\mathbb{R}^n$ is that of the Minkowski metric $(\mathrm{d}x^1)^2 - \sum_{i=2}^n (\mathrm{d}x^i)^2$.

9. Differential forms, integration and Stokes theorem

In subsection 8.3 we have seen that some 1-forms arise as the differential of a function and we called such 1-forms exact. A natural question is how one can characterize 1-forms. A simple answer occurs on an open subset U of $\mathbb{R}^n$. Let $\omega = \sum_{i=1}^n \omega_i \mathrm{d}x^i$ be an exact 1-form: i.e., $\omega = \mathrm{d}f$ for some function f. This means that the components satisfy $\omega_i = \partial_i f$. Now, even without knowing f, we can affirm that $\partial_i\omega_j = \partial_j\omega_i$ for all i,j. This is a necessary condition for a 1-form to be exact (we will see that it is not sufficient in general though). This suggests defining the skew-symmetric tensor field $\sum_{ij}(\partial_i\omega_j - \partial_j\omega_i)\mathrm{d}x^i \otimes \mathrm{d}x^j$ which vanishes if ω is exact. It turns out that this construction makes sense also for manifolds and that it can be futher extended. This leads to the concept of differential forms (i.e., sections of the exterior algebra of the cotangent bundle) and of the de Rham differential (a.k.a. the exterior derivative).

A further reason for studying the de Rham differential is its intimate connection with Stokes theorem, the higher dimensional version of the fundamental theorem of analysis $\int_a^b f'(x)\,dx = f(b) - f(a)$. Just to give a glimpse of it, consider the integration of 1-forms. If $\omega \in \Omega^1(M)$

and $\gamma\colon [a,b] \to M$ is a piecewise differentiable curve, one defines

$$\int_\gamma \omega = \int_a^b \omega_{\gamma(t)}(\dot\gamma(t))\,dt.$$

If M is an open subset of $\mathbb{R}^n$ and $\omega = \sum_{i=1}^n \omega_i\,\mathrm{d}x^i$, we just have

$$\int_\gamma \omega = \int_a^b \sum_{i=1}^n \omega_i(\gamma(t))\,\dot\gamma^i(t)\,dt.$$

If $\omega = \mathrm{d}f$, we then get

$$\int_\gamma \omega = f(\gamma(b)) - f(\gamma(a)).$$

On a manifold this is also true: simply split γ into portions each lying in a single chart and apply the result in charts. Finally, if γ is closed (i.e., $\gamma(a) = \gamma(b)$) we get $\int_\gamma \omega = 0$ if ω is exact.

For simplicity in this section we will only consider smooth manifolds an smooth differential forms on them.[27]

9.1. **Differential forms.** If E is a vector bundle over M (see Definition 8.1 on page 53), we define $\Lambda^k E$ as the vector bundle whose fiber at q is $\Lambda^k E_q$. (We use the notations of Appendix B.2.) Namely, to an adapted atlas $\{(\widetilde{U}_\alpha, \widetilde{\phi}_\alpha)\}_{\alpha\in I}$ of E over the atlas $\{(U_\alpha,\phi_\alpha)\}_{\alpha\in I}$ of M, we associate the atlas $\{(\widehat{U}_\alpha, \widehat{\phi}_\alpha)\}_{\alpha\in I}$ of $\wedge^m E$ with $\widehat{U}_\alpha = \pi^{-1}_{\wedge^m E}(U_\alpha) = \cup_{q\in U_\alpha} \wedge^m E_q$ and

$$\begin{array}{cccc} \widehat{\phi}_\alpha\colon & \widehat{U}_\alpha & \to & \mathbb{R}^n \times \wedge^m\mathbb{R}^r \\ & (q, \omega \in \wedge^m E_q) & \mapsto & (\phi_\alpha(q), \wedge^m A_\alpha(q)\,\omega) \end{array}$$

where we regard $\wedge^m\mathbb{R}^r$ as the manifold $\mathbb{R}^{\binom{r}{m}}$ with its standard structure. It follows that we have transitions maps

$$\widehat{\phi}_{\alpha\beta}(x,u) = (\phi_{\alpha\beta}(x), \wedge^m A_{\alpha\beta}(\phi_\alpha^{-1}(x))\,u).$$

Definition 9.1. An m-form on a smooth manifold M is a section of $\wedge^m T^*M$. We denote by $\Omega^m(M)$ the $\mathcal{C}^\infty(M)$-module of m-forms and by $\Omega^\bullet(M) = \oplus_m \Omega^m(M)$. An element of $\Omega^\bullet(M)$ is called a differential form on M.

Notice that if σ is an m-form, then (8.1) now reads

$$(\wedge^m \mathrm{d}_x\phi_{\alpha\beta})^*\sigma_\beta(\phi_{\alpha\beta}(x)) = \sigma_\alpha(x).$$

[27]By inspection one may see how certain definitions may be changed otherwise and that certain results only require continuity or continuous differentiability in low degree.

The exterior product of the exterior algebra induces, by pointwise multiplication, the exterior product of differential forms: $(\alpha \wedge \beta)_q := \alpha_q \wedge \beta_q$. It follows that, if α is a k-form and β an l-form, we have

$$\beta \wedge \alpha = (-1)^{kl} \alpha \wedge \beta.$$

Since we regard the exterior algebra as a subspace of the tensor algebra, we have that differential forms are a special case of covariant tensor fields. In particular, we may restrict the pullback and the Lie derivative to differential forms. More explicitly, if F is a map $M \to N$ and $\omega \in \Omega^k(N)$, we have

$$(F^*\omega)_q = (\Lambda^k \mathrm{d}_q F)^* \omega_{F(q)}.$$

We also have $F^*(\alpha \wedge \beta) = F^*\alpha \wedge F^*\beta$.

Using pull-backs, we can write the representation of a differential form σ in a chart (U, ϕ_U) as $\sigma_U = (\phi_U^{-1})^* \omega|_U$. Moreover, we can rewrite the compatibility equation as

$$\boxed{\sigma_\alpha = \phi_{\alpha\beta}^* \sigma_\beta} \tag{9.1}$$

for all $\alpha, \beta \in I$, where, by abuse of notation, σ_α denotes here the restriction of σ_α to $\phi_\alpha(U_\alpha \cap U_\beta)$ and σ_β denotes the restriction of σ_β to $\phi_\beta(U_\alpha \cap U_\beta)$.

If M is Hausdorff, the Lie derivative is defined as usual:

$$\mathsf{L}_X \omega := \frac{\partial}{\partial t}\Big|_{t=0} (\Phi_{-t}^X)_* \omega = \frac{\partial}{\partial t}\Big|_{t=0} (\Phi_t^X)^* \omega,$$

for $X \in \mathfrak{X}(M)$ and $\omega \in \Omega^\bullet(M)$. This immediately implies

$$\frac{\partial}{\partial t}\Big|_{t=s} (\Phi_{-t}^X)_* \omega = (\Phi_{-s}^X)_* \mathsf{L}_X \omega = \mathsf{L}_X (\Phi_{-s}^X)_* \omega \tag{9.2}$$

The Lie derivative has all the properties so far discussed for tensor fields. We list them here:

$$\mathsf{L}_{X+Y} \omega = \mathsf{L}_X \omega + \mathsf{L}_Y \omega, \tag{9.3a}$$

$$\mathsf{L}_X \mathsf{L}_Y \omega - \mathsf{L}_Y \mathsf{L}_X \omega = \mathsf{L}_{[X,Y]} \omega, \tag{9.3b}$$

for every two vector fields X, Y and every differential form ω. In addition, it satisfies the Leibniz rule

$$\mathsf{L}_X(\alpha \wedge \beta) = \mathsf{L}_X \alpha \wedge \beta + \alpha \wedge \mathsf{L}_X \beta$$

for every vector field X and every two differential forms α and β. Finally, if $F \colon M \to N$ is a diffeomorphism, then

$$F_* \mathsf{L}_X \omega = \mathsf{L}_{F_* X} F_* \omega, \tag{9.4}$$

for all $X \in \mathfrak{X}(M)$ and all $\omega \in \Omega(M)$.

Another useful operation is that of contraction. This is also defined pointwise using the contraction of vectors with forms explained in Appendix B.2.1. Namely, if X is a vector field and ω a differential form, one defines

$$(\iota_X\omega)_q = \iota_{X_q}\omega_q,$$

for all $q \in M$. From Lemma B.14 we immediately get

$$\iota_X(\alpha \wedge \beta) = \iota_X\alpha \wedge \beta + (-1)^k \alpha \wedge \iota_X\beta,$$

for all $X \in \mathfrak{X}(M)$, $\alpha \in \Omega^k(M)$ and $\beta \in \Omega(M)$, as well as

$$\iota_X\iota_Y\alpha = -\iota_Y\iota_X\alpha \tag{9.5}$$

for all $X, Y \in \mathfrak{X}(M)$ and $\alpha \in \Omega(M)$. Finally, if $F\colon M \to N$ is a diffeomorphism, by (B.9) we get

$$F_*\iota_X\omega = \iota_{F_*X}F_*\omega, \tag{9.6}$$

for all $X \in \mathfrak{X}(M)$ and all $\omega \in \Omega(M)$.

9.2. **The de Rham differential.** We now return to the problem of extending the differential to higher forms.

We start with the case when U is an open subset of $\mathbb{R}^n$. Then $\mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}$, $i_1, \dots, i_k \in \{1, \dots, n\}$, is a system of generators of $\Omega^k(U)$ over $\mathcal{C}^\infty(U)$ (a basis if we take only $i_1 < \cdots < i_k$). We can then expand $\alpha \in \Omega^k(U)$ as $\alpha = \sum_{i_1,\dots,i_k=1}^n \alpha_{i_1\cdots i_k}\,\mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}$. The de Rham differential of α is then defined as the $(k+1)$-form

$$\mathrm{d}\alpha = \sum_{j=1}^n \sum_{i_1,\dots,i_k=1}^n \partial_j\alpha_{i_1\cdots i_k}\,\mathrm{d}x^j \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}.$$

Notice that if α is a top form, i.e., $k = n$, then automatically $\mathrm{d}\alpha = 0$.

Lemma 9.2. *The de Rham differential on an open subset U of $\mathbb{R}^n$ is a collection of linear maps $\mathrm{d}\colon \Omega^k(U) \to \Omega^{k+1}(U)$ for all k and satisfies the following three properties:*

1. *$\mathrm{d}\colon \Omega^0(U) \to \Omega^1(U)$ is the usual differential of functions;*
2. *$\mathrm{d}^2 = 0$;*
3. *$\mathrm{d}(\alpha \wedge \beta) = \mathrm{d}\alpha \wedge \beta + (-1)^k \alpha \wedge \mathrm{d}\beta$ for $\alpha \in \Omega^k(U)$ and $\beta \in \Omega(U)$.*

Proof. Linearity and property (1) are clear from the definition. For property (2) we compute

$$\mathrm{d}^2\alpha = \sum_{l=1}^n \sum_{j=1}^n \sum_{i_1,\dots,i_k=1}^n \partial_l\partial_j\alpha_{i_1\cdots i_k}\,\mathrm{d}x^l \wedge \mathrm{d}x^j \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}.$$

This vanishes since $\partial_l\partial_j\alpha_{i_1\cdots i_k}$ is symmetric in the exchange of l and j whereas $\mathrm{d}x^l \wedge \mathrm{d}x^j$ is skew-symmetric.

Property (3) also follows from a direct computation. For simplicity we assume $\alpha = f\,\mathrm{d}x^{i_1}\wedge\ldots\mathrm{d}x^{i_k}$ and $\beta = g\,\mathrm{d}x^{r_1}\wedge\ldots\mathrm{d}x^{r_l}$, the general case following by linearity. We have $\alpha\wedge\beta = fg\,\mathrm{d}x^{i_1}\wedge\ldots\mathrm{d}x^{i_k}\wedge\mathrm{d}x^{r_1}\wedge\ldots\mathrm{d}x^{r_l}$. Hence

$$\begin{aligned}\mathrm{d}(\alpha\wedge\beta) &= \sum_{j=1}^n \partial_j(fg)\,\mathrm{d}x^j\wedge\mathrm{d}x^{i_1}\wedge\ldots\mathrm{d}x^{i_k}\wedge\mathrm{d}x^{r_1}\wedge\ldots\mathrm{d}x^{r_l} = \\ &= \sum_{j=1}^n \partial_j f g\,\mathrm{d}x^j\wedge\mathrm{d}x^{i_1}\wedge\ldots\mathrm{d}x^{i_k}\wedge\mathrm{d}x^{r_1}\wedge\ldots\mathrm{d}x^{r_l} + \\ &+ (-1)^k\sum_{j=1}^n f\partial_j g\,\mathrm{d}x^{i_1}\wedge\ldots\mathrm{d}x^{i_k}\wedge\mathrm{d}x^j\wedge\mathrm{d}x^{r_1}\wedge\ldots\mathrm{d}x^{r_l} = \\ &= \mathrm{d}\alpha\wedge\beta + (-1)^k\alpha\wedge\mathrm{d}\beta.\end{aligned}$$

□

Lemma 9.3. *If U is an open subset of $\mathbb{R}^n$, the de Rham differential is uniquely determined by properties (1), (2) and (3).*

Proof. Since every differential form on U is a linear combination of wedge products of functions and exact 1-forms, (3) implies that d is completely determined by its action on functions and exact 1-forms. On the other hand, (1) defines d on functions and (2) says that d vanishes on exact 1-forms. □

Corollary 9.4. *Let F be a smooth map $U \to V$ where U and V are open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Then*

$$\mathrm{d}F^* = F^*\mathrm{d}.$$

Proof. We have already proved, Lemma 8.63, that $\mathrm{d}F^*f = F^*\mathrm{d}f$ for f a function. Applying d and using $\mathrm{d}^2 = 0$, we get $0 = \mathrm{d}F^*\mathrm{d}f$. On the other hand, we also have $F^*\mathrm{d}\mathrm{d}f = 0$. This shows that $\mathrm{d}F^*\alpha = F^*\mathrm{d}\alpha$ if α is a function or an exact 1-form.

Next notice that $\mathrm{d}F^*(\alpha\wedge\beta) = \mathrm{d}F^*\alpha\wedge F^*\beta + (-1)^k F^*\alpha\wedge\mathrm{d}F^*\beta$, for $\alpha\in\Omega^k(U)$ and $\beta\in\Omega(U)$, so it is enough to compute $\mathrm{d}F^*$ on functions and on exact 1-forms. □

We now turn to the general case. Let M be a smooth manifold and ω a k-form. Pick an atlas $\{(U_\alpha,\phi_\alpha)\}_{\alpha\in I}$ and denote by ω_α the representation of ω in the chart α. By Corollary 9.4, the collection $\mathrm{d}\omega_\alpha$ defines a $(k+1)$-form on M that we denote by $\mathrm{d}\omega$. It then immediately follows that d is a collection of linear maps $\Omega^k(M)\to\Omega^{k+1}(M)$ satisfying the conditions

(1) $\mathrm{d}\colon \Omega^0(M) \to \Omega^1(M)$ is the usual differential of functions;
(2) $\mathrm{d}^2 = 0$;
(3) $\mathrm{d}(\alpha\wedge\beta) = \mathrm{d}\alpha\wedge\beta + (-1)^k\alpha\wedge\mathrm{d}\beta$ for $\alpha \in \Omega^k(M)$ and $\beta \in \Omega(M)$.

If M is Hausdorff and second countable, then d is uniquely determined by these properties. The proof is as for Corollary 9.3 after noticing, by using a partition of unity, that every differential form on M is a linear combination of wedge products of functions and exact 1-forms.

Finally, if $F\colon M \to N$ is a smooth map, we have

$$\mathrm{d}F^* = F^*\mathrm{d}. \tag{9.7}$$

This follows imediately from Corollary 9.4 using the chart representations of F.

Definition 9.5. The collection $\Omega^k(M)$ together with the de Rham differential d is called the de Rham complex of M.

Remark 9.6. To simplify notations it makes sense to extend the definitions to negative degrees. Namely, one defines $\Omega^k(M) := \{0\}$ for k a negative integer. Then $\mathrm{d}\colon \Omega^k(M) \to \Omega^{k+1}(M)$ is the zero map if $k < 0$. Notice that we still have $\mathrm{d}^2 = 0$.

Remark 9.7 (Vector calculus)**.** If U is an open subset of $\mathbb{R}^3$ the de Rham differential corresponds, in the various degrees, to gradient, curl and divergence. Divergence and gradient are defined in terms of the Euclidean metric as a special case of Remarks 8.81 and 8.82. They simply read

$$\mathrm{grad} f = \sum_{i=1}^3 \partial_i f\,\partial_i, \qquad \mathrm{div} X = \sum_{i=1}^3 \partial_i X^i,$$

where we use the expansion $X = \sum_{i=1}^3 X^i\partial_i$. In addition one defines

$$\mathrm{curl} X = \sum_{i,j,k=1}^3 \epsilon_{ijk}\,\partial_i X^j\,\partial_k.$$

One can easily verify that $\mathrm{curl}\circ\mathrm{grad} = 0$ and $\mathrm{div}\circ\mathrm{curl} = 0$. All this is actually equivalent to de Rham. The point is that $\dim\Lambda^2\mathbb{R}^3 = 3 = \dim\Lambda^1\mathbb{R}^3$. This allows us to define an isomorphism from the $C^\infty(U)$-module of 2-forms to that of 1-forms and to that of vector fields. More generally, we define isomorphisms $\phi_1\colon \Omega^1(U) \to \mathfrak{X}(U)$, $\phi_2\colon \Omega^2(U) \to \mathfrak{X}(U)$, $\phi_3\colon \Omega^3(U) \to \mathcal{C}^\infty(U)$ by specifying them on pure forms:

$$\phi_1(\mathrm{d}x^i) := \partial_i, \quad \phi_2(\mathrm{d}x^i\wedge\mathrm{d}x^j) := \sum_{k=1}^3 \epsilon_{ijk}\partial_k, \quad \phi_3(\mathrm{d}x^i\wedge\mathrm{d}x^j\wedge\mathrm{d}x^k) := \epsilon_{ijk}.$$

One can easily see that grad $= \phi_1 \mathrm{d}$, curl $= \phi_2 \mathrm{d}\phi_1^{-1}$, div $= \phi_3 \mathrm{d}\phi_2^{-1}$ and one sees that the identities of vector calculus are equivalent to $\mathrm{d}^2 = 0$.

9.2.1. *The de Rham cohomology.* The equation $\mathrm{d}^2 = 0$ implies that the image of $\mathrm{d}\colon \Omega^{k-1}(M) \to \Omega^k(M)$ lies in the kernel of $\mathrm{d}\colon \Omega^k(M) \to \Omega^{k+1}(M)$. Differential forms in the image of d are called exact, those in the kernel of d are called closed. Thus, every exact form is closed. To measure the failure of the converse statement, one introduces the de Rham cohomology groups:

$$H^k(M) := \ker(\mathrm{d}\colon \Omega^k(M) \to \Omega^{k+1}(M)) / \operatorname{im}(\mathrm{d}\colon \Omega^{k-1}(M) \to \Omega^k(M)).$$

These groups have a lot of interesting properties. One can e.g. show (even though we will not do it here) that, under mild assumptions, they are finite dimensional, which makes them very managable.

Since the pullback $F^*\colon \Omega(N) \to \Omega(M)$ by a map $F\colon M \to N$ commutes with d, it descends to the quotient: i.e., we can define $F^*[\omega] := [F^*\omega]$ where ω is any closed form representing the class $[\omega]$.

If F is a diffeomorphism, then F^* is an isomorphism. This implies that if $\dim H^k(M) \neq \dim H^k(N)$ for some k, then M and N cannot be diffeomorphic. One can actually prove that it is enough to have a homeomorphism for the cohomology groups to be isomorphic.

There are several techniques to compute cohomology groups. We refer to texts on algebraic topology for this. In particular, in the context of differential forms we recommend [1]. We present just a fundamental result known as the Poincaré Lemma.

Definition 9.8. A subset U of $\mathbb{R}^n$ is called star shaped if it possesses a distinguished point x_0 such that for every $x \in U$ the segment joining x_0 to x is entirely contained in U.

Lemma 9.9 (Poincaré Lemma). *Let U be an open star-shaped subset of $\mathbb{R}^n$ (e.g. $\mathbb{R}^n$ itself). Then*

$$H^0(U) = \mathbb{R}, \qquad H^k(U) = \{0\} \text{ for } k \neq 0.$$

Proof. For simplicity we assume $q_0 = 0$ (otherwise just observe that a translation that moves q_0 to 0 is a diffeomorphism). The star shape condition means that for every $x \in U$ also $tx \in U$ for all $t \in [0, 1]$.

If $f \in \Omega^0(U)$ is closed, then $\partial_i f(x) = 0$ for all i and all x. Then

$$f(x) - f(0) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} f(tx) dt = \int_0^1 \sum_i \partial_i f(tx) x^i dt = 0.$$

This means that a closed zero form is completely specified by its value at 0. Hence the linear map $H^0(U) \to \mathbb{R}$, $f \mapsto f(0)$, is an isomorphism.

To explain the general method we consider first the case of a 1-form $\omega = \sum_i \omega_i \mathrm{d}x^i$ as the general case will be just a generalization of this. First observe that

$$\omega_i(x) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\left(t\omega_i(tx)\right) dt = \int_0^1 \left(\omega_i(tx) + t\sum_j \partial_j\omega_i(tx)x^j\right) dt.$$

If ω is closed, then we have $\partial_j\omega_i = \partial_i\omega_j$. Hence

$$\omega_i(x) = \int_0^1 \left(\omega_i(tx) + t\sum_j \partial_i\omega_j(tx)x^j\right) dt.$$

The idea now is to define $K\colon \Omega^1(U) \to \Omega^0(U)$ by

$$K\omega(x) = \int_0^1 \sum_j \omega_j(tx)\, x^j\, dt.$$

We then have $\partial_i K\omega = \omega_i$, which shows that ω is exact.

The extension to the general case is similar but requires more computations and some guesswork. This will become more clear with the Cartan calculus. We present it here anyway for completeness, but will return to it in subsubsection 9.3.3. Let $\omega = \sum_{i_1,\dots,i_k} \omega_{i_1\cdots i_k}\mathrm{d}x^{i_1}\wedge\cdots\wedge\mathrm{d}x^{i_k}$ be a k-form with $k > 0$. We have

$$\begin{aligned}\omega_{i_1\cdots i_k}(x) &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\left(t^k\omega_{i_1\cdots i_k}(tx)\right) dt = \\ &= \int_0^1 \left(kt^{k-1}\omega_{i_1\cdots i_k}(tx) + t^k\sum_j \partial_j\omega_{i_1\cdots i_k}(tx)x^j\right) dt.\end{aligned}$$

If ω is closed, assuming we already took its components to be skew-symmetric in the indices, we have

$$\partial_j\omega_{i_1\cdots i_k} = \sum_{r=1}^k \partial_{i_r}\omega_{i_1\cdots i_{r-1}j i_{r+1}\cdots i_k}$$

This suggests defining $K\colon \Omega^k(U) \to \Omega^{k-1}(U)$ by

(9.8)

$$K\omega = \sum_{i_1,\dots,i_k}\sum_{r=1}^k (-1)^{r-1}\left(\int_0^1 t^{k-1}\omega_{i_1\cdots i_k}(tx)x^{i_r}\,dt\right)\mathrm{d}x^{i_1}\wedge\cdots\wedge\widehat{\mathrm{d}x^{i_r}}\wedge\mathrm{d}x^{i_k}.$$

One can then easily check that $\mathrm{d}K\omega = \omega$, which shows that ω is exact.

□

Example 9.10. As a simple application we may show that $\mathbb{R}^2$ is not diffeomorphic to $\mathbb{R}^2 \setminus \{0\}$ (actually, they are not even homeomorphic). By the Poincaré Lemma $H^1(\mathbb{R}^2) = \{0\}$. On other hand, the 1-form $\omega = (x\mathrm{d}y - y\mathrm{d}x)/(x^2 + y^2)$ on $\mathbb{R}^2 \setminus \{0\}$ is closed but $\int_\gamma \omega = 2\pi$, where γ is a circle around the origin. This shows that ω is not exact, so $\dim H^1(\mathbb{R}^2 \setminus \{0\}) > 0$.

9.3. **Graded linear algebra and the Cartan calculus.** We have seen several relations satified by the Lie derivative, the contractions and the de Rham differential. Even more hold and they make working with differential forms much handier than with general tensor fields. One can summarize these relations nicely by using the language of graded Lie algebras. We thus make a short digression on graded linear algebra.

9.3.1. *Graded linear algebra.* In graded linear algebra one generalizes the usual concepts of linear algebras to collections of vector spaces.

Definition 9.11. A graded vector space $V^\bullet$ is a collection $\{V^k\}_{k\in\mathbb{Z}}$ of vector spaces. A morphism $\phi\colon V^\bullet \to W^\bullet$ is a collection of linear maps $\phi^k\colon V^k \to W^k$ for all k. A graded morphism $\phi\colon V^\bullet \to W^\bullet$ of degree r is a collection of linear maps $\phi^k\colon V^k \to W^{k+r}$ for all k. If $W^\bullet = V^\bullet$, we call ϕ a graded endomorphism.

Example 9.12. The vector spaces $V^{\otimes k}$, $\Lambda^k V$ and $\Omega^k(M)$ define graded vector spaces $T^\bullet(V)$, $\Lambda^\bullet V$ and $\Omega^\bullet(M)$. The de Rham differential is an example of graded morphism of degree $+1$, the contraction by a vector (field) is an example of graded morphism of degree -1, the Lie derivative by a vector field is an example of graded morphism of degree 0.

Example 9.13. Notice that the set $\mathrm{Hom}^r(V^\bullet, W^\bullet)$ of graded morphisms of degree r is a vector space for each $r \in \mathbb{Z}$. Hence we have a new graded vector space $\mathrm{Hom}^\bullet(V^\bullet, W^\bullet)$. In case $V^\bullet = W^\bullet$, we write $\mathrm{End}^\bullet(V^\bullet)$.

Remark 9.14. Sometimes one also uses the realization $V := \bigoplus_{k\in\mathbb{Z}} V^k$ of a graded vector space $V^\bullet$. We have used this notation, e.g., for $\Omega(M)$. Notice that, if infinitely many V^k are not zero dimensional, a graded morphism in general does not define a linear map between the realizations.

Definition 9.15. A graded endomorphism of degree -1 that squares to zero is called a boundary operator. A graded endomorphism of degree $+1$ that squares to zero is called a coboundary operator. A graded vector space endowed with a boundary or a coboundary operator is called a complex.

We have already seen the example of the de Rham complex $(\Omega^\bullet(M), \mathrm{d})$.

Definition 9.16. A graded algebra is a graded vector space $A^\bullet$ together with a collection of bilinear maps $A^k \times A^l \to A^{k+l}$, $(a,b) \mapsto ab$, for all k,l. The graded algebra is called associative if $(ab)c = a(bc)$ for all a,b,c. It is called graded commutative if

$$ab = (-1)^{kl} ba, \qquad \text{for } a \in A^k,\ b \in A^l.$$

It is called graded skew-commutative if

$$ab = -(-1)^{kl} ba, \qquad \text{for } a \in A^k,\ b \in A^l.$$

Example 9.17. $(\Lambda^\bullet V, \wedge)$ and $(\Omega^\bullet(M), \wedge)$ are examples of associative graded commutative algebras.

Example 9.18. The composition $\phi\psi$ of two graded endomorphisms ϕ and ψ of $V^\bullet$ of degree r and s, respectively, is defined as the collection of linear maps $\phi^{k+s}\psi^k \colon V^k \to V^{k+r+s}$. It is then a graded endomorphism of degree $r+s$. This makes $\mathrm{End}^\bullet(V^\bullet)$ into an associative graded algebra.

Definition 9.19. A graded derivation D of degree r on a graded algebra $A^\bullet$ is a graded morphism of degree r satisfying the graded Leibniz rule

$$D(ab) = Da\, b + (-1)^{kr} a\, Db, \qquad \text{for } a \in A^k,\ b \in A^l.$$

Notice that r enters both as the degree of D and in the sign factor.

Example 9.20. The de Rham differential is a graded derivation of degree $+1$ on $\Omega^\bullet(M)$, the contraction by a vector (field) is a graded derivation of degree -1, and the Lie derivative by a vector field is a graded derivation of degree 0.

Definition 9.21. A coboundary operator that is also a derivation is called a differential.

The standard example for us is the de Rham differential.

Remark 9.22. Notice that the set $\mathrm{Der}^r(A^\bullet)$ of graded derivations of degree r on a graded algebra $A^\bullet$ is a vector space, actually a subspace of $\mathrm{End}^r(A^\bullet)$, for all r, so we have a new graded vector space $\mathrm{Der}^\bullet(A^\bullet)$.

Definition 9.23. A graded Lie algebra is a graded algebra $\mathfrak{g}^\bullet$ whose product, usually denoted by $[\ ,\ ]$, is graded skew-commutative and satisfies the graded Jacobi identity

$$[a,[b,c]] = [[a,b],c] + (-1)^{kl}[b,[a,c]], \qquad \text{for } a \in \mathfrak{g}^k,\ b \in \mathfrak{g}^l,\ c \in \mathfrak{g}^m.$$

Example 9.24. The graded commutator of $a \in A^k$ and $b \in A^l$, where $A^\bullet$ is an associative graded algebra, is defined by

$$[a, b] := ab - (-1)^{kl} ba.$$

One can easily verify that $(A^\bullet, [\ ,\ ])$ is a graded Lie algebra. In particular, $(\mathrm{End}^\bullet(V^\bullet), [\ ,\ ])$, where $V^\bullet$ is a graded vector space, is a graded Lie algebra.

Remark 9.25. An element of even/odd degree of a graded vector space is called even/odd. Notice then that, in the above example, if both a and b are odd, then we have

$$[a, b] = ab + ba,$$

the so-called anti-commutator; otherwise $[a, b] = ab - ba$, the usual commutator.

Definition 9.26. A graded subspace $W^\bullet$ of a graded vector space $V^\bullet$ is a collection $W^k \subset V^k$ of subspaces for all k. A graded subalgebra $B^\bullet$ of a graded algebra $A^\bullet$ is a graded subspace that is closed under the product. Usually a graded subalgebra of a graded Lie algebra is called a graded Lie subalgebra.

Example 9.27. One can verify, exactly as in the non graded case, that $(\mathrm{Der}^\bullet(A^\bullet), [\ ,\ ])$ is a graded Lie subalgebra of $(\mathrm{End}^\bullet(A^\bullet), [\ ,\ ])$, where $A^\bullet$ is a graded algebra.

Remark 9.28. The choice of $\mathbb{Z}$ for the grading and the sign conventions we have used are those needed for differential forms. More generally, one may define a G-graded vector space as a collection $\{V^k\}_{k\in G}$ of vector spaces, where G is a set. In order to define graded morphisms we have to assume that a composition law $G \times G \to G$ is given. If we want to view morphisms as a special case of graded morphisms we have to assume that G possesses a special element. In order to define the notion of associative graded algebra (and also to give the space of graded endomorphisms the structure of a graded algebra), we need G to be a group. In order to define graded commutivity, for whatever choice of signs, we have to assume that G is abelian. The choice of signs in the definition of graded commutativity is a map $s\colon G \times G \to \{-1, 1\}$. Compatibility between associativity and graded commutativity give conditions on this map s. Typical examples for G and s are $G = \mathbb{Z}$ and $s(k, l) = (-1)^{kl}$ (which is what we have done here), $G = \mathbb{Z}$ and $s(k, l) = 1$, $G = \mathbb{Z}/2\mathbb{Z}$ and $s(k, l) = (-1)^{kl}$, $G = \mathbb{Z}/2\mathbb{Z}$ and $s(k, l) = 1$, $G = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}$ and $s((k, k'), (l, l')) = (-1)^{kl}$.

9.3.2. *Cartan calculus.* We now come to the fundamental

Theorem 9.29. *The span over* $\mathbb{R}$ *of the set* $\{\mathrm{d}, \iota_X, \mathsf{L}_X : X \in \mathfrak{X}(M)\}$ *is a graded Lie subalgebra of* $\mathrm{Der}^\bullet(\Omega^\bullet(M))$. *More precisely,*

$$\begin{aligned} [\mathrm{d}, \mathrm{d}] &= 0, \\ [\mathrm{d}, \iota_X] &= \mathsf{L}_X, \\ [\mathrm{d}, \mathsf{L}_X] &= 0, \\ [\iota_X, \iota_Y] &= 0, \\ [\iota_X, \mathsf{L}_Y] &= \iota_{[X,Y]}, \\ [\mathsf{L}_X, \mathsf{L}_Y] &= \mathsf{L}_{[X,Y]}, \end{aligned}$$

for all $X, Y \in \mathfrak{X}(M)$.

The second relation is known as Cartan's formula and is very useful to compute the Lie derivative of a differential form.

Proof. The proof is very simple since these are all identities between derivations, so it is enough to check them on functions and exact 1-forms.[28] Some identities are actually already known for all differential forms or follow easily from the other identities.

The first identity is just $\mathrm{d}^2 = 0$. The second identity for functions is just the fact that $\mathsf{L}_X f = X(f) = \iota_X \mathrm{d}f = [\mathrm{d}, \iota_X]f$. On exact 1-forms we have

$$[\mathrm{d}, \iota_X]\mathrm{d}f = \mathrm{d}\iota_X\mathrm{d}f = \mathrm{d}\mathsf{L}_X f$$

which is the same as $\mathsf{L}_X\mathrm{d}f$ by (8.18) on page 79. The third identity follows from the second and the graded Jacobi identity

$$[\mathrm{d}, \mathsf{L}_X] = [\mathrm{d}, [\mathrm{d}, \iota_X]] = [[\mathrm{d}, \mathrm{d}], \iota_X] - [\mathrm{d}, [\mathrm{d}, \iota_X]] = -[\mathrm{d}, \mathsf{L}_X].$$

One can also directly prove it using Corollary 9.4 with F the flow of X.

The fourth identity is equation (9.5). The fifth identity is obvious on functions (as the contraction kills functions) and for 1-forms is equation (8.20) on page 79.[29] Finally, the last identity is just (9.3b), but it can

[28] Even if M does not have a partition of unity subordinate to an atlas, this is still ok, as we can actually first prove this theorem on open subsets of $\mathbb{R}^n$, where it is true that a differential form is a linear combination of wedge products of functions and exact 1-forms, and then use the transformation properties (9.4), (9.6) and (9.7) under push-forward by the chart maps.

[29] We can also obtain it on exact 1-forms using the third and fifth equations:

$$\begin{aligned} [\iota_X, \mathsf{L}_Y]\mathrm{d}f = \iota_X\mathsf{L}_Y\mathrm{d}f - \mathsf{L}_Y\iota_X\mathrm{d}f = \iota_X\mathrm{d}\mathsf{L}_Y f - \mathsf{L}_Y\mathsf{L}_X f = \\ = \mathsf{L}_X\mathsf{L}_Y f - \mathsf{L}_Y\mathsf{L}_X f = [\mathsf{L}_X, \mathsf{L}_Y]f = \mathsf{L}_{[X,Y]}f = \iota_{[X,Y]}\mathrm{d}f. \end{aligned}$$

also be obtained from the second, the third and the fifth by using the graded Jacobi identity:

$$\mathsf{L}_{[X,Y]} = [\mathrm{d}, \iota_{[X,Y]}] = [\mathrm{d}, [\iota_X, \mathsf{L}_Y]] = [[\mathrm{d}, \iota_X], \mathsf{L}_Y] - [\iota_X, [\mathrm{d}, \mathsf{L}_Y]] = [\mathsf{L}_X, \mathsf{L}_Y].$$

□

We will see several application of the Cartan calculus in general and of Cartan's formula in particular.

Remark 9.30. A simple consequence is that the Lie derivative of a closed differential form ω is simply

$$\mathsf{L}_X\omega = \mathrm{d}\iota_X\omega.$$

In particular this is true when ω is a top form.

Another application is a very explicit formula for computing the de Rham differential. We start with an iteration of the fifth identity of the Cartan calculus.

Lemma 9.31. *Given $k+1$ vector fields $X_0, X_1, \dots X_k$, $k \geq 1$,*

$$[\iota_{X_k} \cdots \iota_{X_1}, \mathsf{L}_{X_0}] = \sum_{i=1}^{k} (-1)^{i+1} \iota_{X_k} \cdot \widehat{\iota_{X_i}} \cdots \iota_{X_1} \iota_{[X_i, X_0]}.$$

Proof. The proof is by induction on the number of vector fields. For $k = 1$ this is the fifth identity in the Cartan calculus. Now suppose it has been proved for $k-1$. Since the graded Lie bracket is a graded commutator, we have

$$[\iota_{X_k} \cdots \iota_{X_1}, \mathsf{L}_{X_0}] = \iota_{X_k} \cdots \iota_{X_2} [\iota_{X_1}, \mathsf{L}_{X_0}] + [\iota_{X_k} \cdots \iota_{X_2}, \mathsf{L}_{X_0}] \iota_{X_1}.$$

Inserting the formulae for two vector fields and for k vector fields yields then the identity for $k+1$ vector fields. □

If ω is a k-form, we use the notation

$$\omega(X_1, \dots, X_k) = \iota_{X_k} \cdots \iota_{X_1} \omega,$$

where on the left hand side ω is regarded as an alternating multilinear form on vector fields. Notice that $\omega(X_1, \dots, X_k)$ is a function.

Proposition 9.32. *Given $k+1$ vector fields $X_0, X_1, \dots X_k$, $k \geq 0$, and a k-form ω, we have*

$$\mathrm{d}\omega(X_0, \cdots, X_k) = \sum_{i=0}^{k} (-1)^i X_i(\omega(X_0, \dots \widehat{X_i}, \dots, X_k)) +$$
$$+ \sum_{0 \leq i < j \leq k} (-1)^{i+j} \omega([X_i, X_j], X_0, \dots, \widehat{X_i}, \dots \widehat{X_j}, \dots, X_k).$$

Proof. This is also proved by induction on k. For $k=0$ it is just the formula $\iota_{X_0}\mathrm{d}\omega = X_0(\omega)$. By induction we can apply the formula to the $(k-1)$-form $\iota_{X_0}\omega$:

$$\begin{aligned}
\mathrm{d}\iota_{X_0}\omega(X_1,\cdots,X_k) &= \sum_{i=1}^{k}(-1)^{i+1}X_i(\iota_{X_0}\omega(X_1,\dots\widehat{X_i},\dots,X_k))+\\
&+\sum_{1\le i<j\le k}(-1)^{i+j}\iota_{X_0}\omega([X_i,X_j],X_1,\dots,\widehat{X_i},\dots\widehat{X_j},\dots,X_k) =\\
&= -\sum_{i=1}^{k}(-1)^{i}X_i(\omega(X_0,\dots\widehat{X_i},\dots,X_k))+\\
&-\sum_{1\le i<j\le k}(-1)^{i+j}\omega([X_i,X_j],X_0,\dots,\widehat{X_i},\dots\widehat{X_j},\dots,X_k).
\end{aligned}$$

On the other hand, by Cartan's formula,

$$\begin{aligned}
\mathrm{d}\iota_{X_0}\omega(X_1,\cdots,X_k) &= \iota_{X_k}\cdots\iota_{X_1}\mathrm{d}\iota_{X_0}\omega =\\
&= -\iota_{X_k}\cdots\iota_{X_0}\mathrm{d}\omega + \iota_{X_k}\cdots\iota_{X_1}\mathsf{L}_{X_0}\omega
\end{aligned}$$

Finally, observe that $\iota_{X_k}\cdots\iota_{X_0}\mathrm{d}\omega$ is the quantity $\mathrm{d}\omega(X_0,\dots,X_k)$ that we want to compute. By Lemma 9.31 the last term finally yields the remaining terms of the formula. □

9.3.3. *The Poincaré Lemma.* As promised, we return to the proof of Lemma 9.9. Its proof reduces to proving the following

Lemma 9.33. *Let U be an open star-shaped subset of $\mathbb{R}^n$. Then there is a graded endomorphism K of degree -1 on $\Omega^\bullet(U)$ such that*

$$\mathrm{d}K\omega + K\mathrm{d}\omega = \omega$$

for all ω of positive degree.

Notice that, if ω is closed, then we get $\omega = \mathrm{d}K\omega$, which shows that it is also exact, thus proving the Poincaré Lemma.

To get a clue at the construction of K in the Lemma, we first consider polynomial forms. A k-form $\omega = \sum_{i_1,\dots,i_k}\omega_{i_1\cdots i_k}\mathrm{d}x^{i_1}\wedge\cdots\wedge\mathrm{d}x^{i_k}$ on $\mathbb{R}^n$ is called polynomial of degree r if all its coefficients $\omega_{i_1\cdots i_k}$ are polynomials of degree r. We let $\Omega^{k,r}(\mathbb{R}^n)$ denote the vector space of k-forms of degree r on $\mathbb{R}^n$ and set $\Omega^{(l)}(\mathbb{R}^n) := \oplus_{k+r=l}\Omega^{k,r}(\mathbb{R}^n)$. The de Rham differential and the contraction ι_E by the Euler vector field $E=\sum_i x^i\partial_i$ are endomorphism of $\Omega^{(l)}$ for every l, and, consequently, so is L_E. Since $\mathsf{L}_E x^i = x^i$ for each coordinate x^i and since it is a derivation commuting with d, we have that $\mathsf{L}_E\omega = l\omega$ for $\omega \in \Omega^{(l)}(\mathbb{R}^n)$. By the Cartan formula we then have

$$\mathrm{d}\iota_E + \iota_E\mathrm{d} = l\,\mathrm{Id}$$

on $\Omega^{(l)}(\mathbb{R}^n)$. We then define $K := \frac{1}{l}\iota_E$ for $l \neq 0$. It follows that

$$\mathrm{d}K + K\mathrm{d} = \mathrm{Id},$$

which in particular shows that the polynomial cohomology is trivial unless $l = 0$, i.e., for constant zero-forms.

We now want to extend K to nonhomogeneous polynomial forms. The trick is to observe that, for $\omega \in \Omega^{(l)}(\mathbb{R}^n)$, we have $(\phi_s^E)^*\omega = \mathrm{e}^{ls}\omega$. This is easily follows from observing that $\phi_s^E(x) = \mathrm{e}^s x$. Using $\frac{1}{l} = \int_{-\infty}^0 \mathrm{e}^{ls}\,ds$, we see that

$$K\omega = \int_{-\infty}^0 (\phi_s^E)^*\iota_E\omega\;ds \tag{9.9}$$

for every polynomial form. (Note that this extends the definition of K to the case $l = 0$ where K must be the zero operator for degree reasons.) The proof of Lemma 9.33 essentially follows from extending such a K to all forms.

Proof of Lemma 9.33. Let U be an open star-shaped subset of $\mathbb{R}^n$. For simplicity assume that we have already translated the distinguished point to 0. Notice that ϕ_s^E is defined for all $s \leq 0$ since U is star shaped, so the integrand in (9.9) is defined. We claim that the integral converges. In fact,

$$\iota_E\omega = \sum_{i_1,\dots,i_k}\sum_{r=1}^{k}(-1)^{r-1}\omega_{i_1\cdots i_k}x^{i_r}\mathrm{d}x^{i_1}\wedge\cdots\wedge\widehat{\mathrm{d}x^{i_r}}\wedge\mathrm{d}x^{i_k}.$$

Hence

$$((\phi_s^E)^*\iota_E\omega)(x) = \sum_{i_1,\dots,i_k}\sum_{r=1}^{k}(-1)^{r-1}\omega_{i_1\cdots i_k}(\mathrm{e}^s x)\,\mathrm{e}^{ks}\,x^{i_r}\mathrm{d}x^{i_1}\wedge\cdots\wedge\widehat{\mathrm{d}x^{i_r}}\wedge\mathrm{d}x^{i_k}.$$

By the change of variable $t = \mathrm{e}^s$, assuming $k > 0$, we get (9.8) which shows that $K\omega$ is the integral of a smooth function on a compact interval; thus, the integral converges, the result is smooth and one can differentiate through the integral.

We now have to check that this K satsifes the identity stated in the Lemma. Since we can differentiate through the integral, by the Cartan formula we have

$$\mathrm{d}K\omega + K\mathrm{d}\omega = \int_{-\infty}^0 (\phi_s^E)^*\mathsf{L}_E\omega\;ds.$$

By (9.2), writing $(\phi_s^E)^*$ instead of $(\phi_{-s}^E)_*$, we then have

$$\mathrm{d}K\omega + K\mathrm{d}\omega = \int_{-\infty}^{0} \frac{\partial}{\partial s}(\phi_s^E)^*\omega \, ds = (\phi_0^E)^*\omega - \lim_{s\to-\infty}(\phi_s^E)^*\omega.$$

Since $\phi_0^E = \mathrm{Id}$, the first term is ω, so we just have to prove that the second term vanishes. If we expand $\omega = \sum_{i_1,\dots,i_k} \omega_{i_1\cdots i_k}\mathrm{d}x^{i_1}\wedge\cdots\wedge\mathrm{d}x^{i_k}$, we get

$$((\phi_s^E)^*\omega)(x) = \sum_{i_1,\dots,i_k} \omega_{i_1\cdots i_k}(\mathrm{e}^s x)\,\mathrm{e}^{ks}\,\mathrm{d}x^{i_1}\wedge\cdots\wedge\mathrm{d}x^{i_k},$$

which implies

$$\lim_{s\to-\infty}(\phi_s^E)^*\omega = 0$$

if $k > 0$.

□

9.3.4. *Quotients.* We return to the setting of Section 7.5.1 to give an algebraic characteration of forms on a quotient. Let $\pi\colon M \to N$ be a surjective submersion with connected fibers. The pullback by π,

$$\pi^*\colon \Omega^\bullet(N) \to \Omega^\bullet(M),$$

is injective. To show this, assume α is a k-form with $\pi^*\alpha = 0$. Then, for all $q \in M$ and all $v_1,\dots,v_k \in T_qM$, we have

$$0 = ((\pi^*\alpha)_q, v_1\wedge\cdots\wedge v_k) = (\alpha_{\pi(q)}, \Lambda^k\mathrm{d}_q\pi\, v_1\wedge\cdots\wedge v_k).$$

Since $\mathrm{d}_q\pi$ is surjective, $\alpha_{\pi(q)} = 0$. Since π is surjective, we conclude that $\alpha = 0$.

A differential form in the image of π^* is called basic. We wish to characterize basic forms.

Definition 9.34. A differential form α on M is called horizontal if $\iota_Y\alpha = 0$ for every vertical vector field Y and invariant if $\mathsf{L}_Y\alpha = 0$ for every vertical vector field Y.

Proposition 9.35. *A differential form on M is basic if and only it is horizontal and invariant.*

Proof. If $\alpha = \pi^*\beta$, then α is clearly horizontal. Moreover, since the flow of a vertical vector field Y preserves the fibers (i.e., $\pi\circ\Phi_t^Y = \pi$), we have $(\Phi_t^Y)^*\alpha = \alpha$, which implies $\mathsf{L}_Y\alpha = 0$. Hence, α is also invariant.

Conversely, assume that α is a horizontal and invariant k-form. Given vector fields $X_1,\dots,X_k$ on N, let $\widetilde{X}_1,\dots,\widetilde{X}_k$ be projectable vector fields on M with $\phi(\widetilde{X}_i) = X_i$ for all i (see Section 7.5.1 for notations). Since α is horizontal, the function $\iota_{\widetilde{X}_k}\cdots\iota_{\widetilde{X}_1}\alpha$ does not depend on the choice of the $\widetilde{X}_i$s. Let Y be now a vertical vector field. Since α is invariant, by Lemma 9.31 we get

$$\mathsf{L}_Y \iota_{\widetilde{X}_k} \dots \iota_{\widetilde{X}_1} \alpha = \sum_{i=1}^{k} (-1)^i \iota_{\widetilde{X}_k} \cdot \widehat{\iota_{\widetilde{X}_i}} \cdots \iota_{\widetilde{X}_1} \iota_{[\widetilde{X}_i, Y]} \alpha.$$

Since $\widetilde{X}_i$ is projectable and α is horizontal, we conclude that the function $\iota_{\widetilde{X}_k} \dots \iota_{\widetilde{X}_1} \alpha$ is invariant. We may then define a k-form β on N by setting $\iota_{X_k} \cdots \iota_{X_1} \beta(z) := \iota_{\widetilde{X}_k} \dots \iota_{\widetilde{X}_1} \alpha(q)$ for any $q \in \pi^{-1}(z)$. It then also follows that $\alpha = \pi^* \beta$. □

By Cartan's formula we then get the following

Corollary 9.36. *A closed form on M is basic if and only it is horizontal.*

On a Hausdorff manifold, invariance under all vertical vector fields actually implies horizontality. In fact, if Y is vertical, then so is fY for every function f. Since $\mathsf{L}_{fY}\omega = f\mathsf{L}_Y\omega + \mathrm{d}f \wedge \iota_Y \omega$, the invariance conditions imply that $\mathrm{d}f \wedge \iota_Y\omega = 0$ for every function f. In particular, for every $p \in M$ we have $\mathrm{d}_p f \wedge (\iota_Y \omega)_p = 0$. By Lemma 7.13, we can get all germs of functions at p and, in particular, $\mathrm{d}_p f$ spans the whole cotangent space, so $(\iota_Y\omega)_p = 0$. The two conditions must however be imposed separately if we restrict ourselves to generators.

Definition 9.37. Let $\mathcal{Y}$ be a family of generators of the vertical vector fields. A differential form α on M is called $\mathcal{Y}$-horizontal if $\iota_Y\alpha = 0$ for every $Y \in \mathcal{Y}$ and $\mathcal{Y}$-invariant if $\mathsf{L}_Y\alpha = 0$ for every $Y \in \mathcal{Y}$. (One often simply says horizontal and invariant when it is clear that a certain family $\mathcal{Y}$ is understood.)

Lemma 9.38. *Let $\mathcal{Y}$ be a family of generators of the vertical vector fields. A differential form is $\mathcal{Y}$-horizontal and $\mathcal{Y}$-invariant if and only if it is horizontal and invariant, and hence if and only if it is basic.*

Proof. If ω is horizontal and invariant, then it is obviously $\mathcal{Y}$-horizontal and $\mathcal{Y}$-invariant.

Conversely, if ω is $\mathcal{Y}$-horizontal, then for every vertical vector field Y with expansion $\sum_i f_i Y_i$, $Y_i \in \mathcal{Y}$, we get $\iota_Y\omega = \sum_i f_i \iota_{Y_i}\omega = 0$; so ω is horizontal. It follows that $\mathsf{L}_Y\omega = \sum_i f_i \mathsf{L}_{Y_i}\omega$, which vanishes if ω is also $\mathcal{Y}$-invariant.

□

9.4. **Orientation and the integration of differential forms.** In subsection 8.2 we defined the integration of densities. We now turn to the problem of extending integration to differential forms. An n-form on an n-dimensional manifold is called a top form. Let v be a top form and v_α its representations in an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$. If we write $v_\alpha = \underline{v_\alpha}\mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n$, by (B.8) the functions $\underline{v_\alpha}$ transform as

$$\underline{v_\alpha}(x) = \det \mathrm{d}_x\phi_{\alpha\beta}\ \underline{v_\beta}(\phi_{\alpha\beta}(x)).$$

for all $\alpha, \beta \in I$ and for all $x \in \phi_\alpha(U_\alpha \cap U_\beta)$. This is almost the same as the transformation rule for a density, see (8.2), apart from the missing absolute value of the determinant.

One way out is to consider, if possible, a special atlas where the determinants of the differentials of all transition maps are positive: in such an atlas a top form is the same as a density. We will return to this approach in Section 9.4.1, where we will also show that, in positive dimension, it is equivalent to the second approach.

The main idea of the second approach consists in taking the absolute value of the representations v_α. More precisely, we define the absolute value $|v|$ as the density with representations $|\underline{v_\alpha}|$. Notice that there are in general two problems if $\underline{v_\alpha}$ changes sign somewhere. The first problem is that then the absolute value does not produce a smooth density in general. This is not a problem as long as only integration is concerned, but without differentiation we lose the power of Stokes theorem. The second problem is that in general we want to keep track of the values with signs as this may be some very relevant information.

To avoid this problems we restrict first to top forms that do not change sign (at least locally).

Definition 9.39. A volume form on a manifold M is a nowhere vanishing top form. A manifold admitting a volume form is called orientable.[30]

If v is a smooth volume form, then $|v|$ is a smooth density. A volume form v establishes a $\mathcal{C}^\infty(M)$-linear isomorphism

$$\phi_v\colon \Omega^{\mathrm{top}}(M) \to \mathrm{Dens}(M)$$

between the module of top forms and the module of densities as follows: since v is nowhere vanishing, for every top form ω there is a uniquely defined function f such that $\omega = fv$; the corresponding density is then defined to be $f|v|$. Formally we may write

$$\phi_v\omega = \omega\frac{|v|}{v}.$$

Notice that two volume forms v and $\widetilde{v}$ yield the same isomorphism (i.e., $\phi_v = \phi_{\widetilde{v}}$) if and only if there is a positive function g such that $\widetilde{v} = gv$. This establishes an equivalence relation on the set of volume forms.

[30] We will see below that there are manifolds on which volume forms do not exist.

Definition 9.40. An equivalence class of volume forms on an orientable manifold M is called an orientation. An orientable manifold with a choice of orientation is called oriented.

We denote by $[v]$ an orientation, by $(M,[v])$ the corresponding oriented manifold and by $\phi_{[v]}$ the isomorphism given by ϕ_v for any $v \in [v]$. If M admits a partition of unity, we can finally define the integral of a top form ω by

$$\int_{(M,[v])} \omega := \int_M \phi_{[v]}\omega,$$

where we use the already defined integration of densities. (We assume the integral on the right hand side to make sense. Typically we work with compactly supported top forms.)

Remark 9.41. Notice that unlike the integration of densities, the integration of top forms requires an orientation. This allows integrating top forms only on orientable manifolds. Moreover, the result of integration in general depends on the choice of orientation.

Lemma 9.42. *A connected orientable manifold admits two orientations*

Proof. Let $[v]$ be an orientation. Then $-[v] = [-v]$ is a different orientation, so every orientable manifold has at least two orientations. Now let M be connected. Let $[v]$ and $[\widetilde{v}]$ be orientations. Choose representatives $v \in [v]$ and $\widetilde{v} \in [\widetilde{v}]$. Since v is a volume form, there is a uniquely defined function g such that $\widetilde{v} = gv$. Since $\widetilde{v}$ is a volume form, then g is also nowhere vanishing. Since M is connected, we then have only two possibilities: (i) $g > 0$, and in this case $[\widetilde{v}] = [v]$, or (ii) $g < 0$, and in this case $[\widetilde{v}] = [-v]$. □

Remark 9.43. Notice that $\phi_{[-v]} = -\phi_{[v]}$, hence

$$\int_{(M,-[v])} \omega = -\int_{(M,[v])} \omega$$

for every ω.

Suppose $F\colon M \to N$ is a diffeomorphism. If $[v_N]$ is an orientation on N, then F^*v_n is a volume form on M for any $v_N \in [v_N]$. Its class does not depend on the choice of representative, so we write

$$F^*[v_N] := [F^*v_N].$$

If M is oriented, we can compare its orientation with $F^*[v_N]$:

Definition 9.44. A diffeomorphism F of connected oriented manifolds $(M, [v_M])$ and $(N, [v_N])$ is called orientation preserving if $F^*[v_N] = [v_M]$ and orientation reversing if $F^*[v_N] = -[v_M]$.

Proposition 9.45 (Change of variables)**.** *Let $(M, [v_M])$ and $(N, [v_N])$ be connected oriented manifolds, $F\colon M \to N$ a diffeomorphism and ω a top form on N. Then*

$$\int_{(M,[v_M])} F^*\omega = \pm \int_{(N,[v_N])} \omega,$$

with the plus sign if F is orientation preserving and the minus sign if F is orientation reversing.

Proof. This follows immediately from the definitions and from the change of variables for densities, see equation (8.8) on page 71. □

Remark 9.46. An interesting application is a technique for proving that the integral of some ω on an oriented manifold M vanishes. The trick consists in finding an orientation-reversing diffeomorphism F of M such that $F^*\omega = \omega$ or an orientation-preserving diffeomorphism F of M such that $F^*\omega = -\omega$. (Of course one has to verify first that the integral converges.)

Example 9.47. Let $M = \{pt\}$ be the connected 0-dimensional manifold. Top forms are in this case the same as functions, i.e., real numbers. There are two orientations: the class of $+1$ and the class of -1. If f a top form, then we have $\phi_{[+1]} f = f$ and $\phi_{[-1]} f = -f$, so

$$\int_{(\{pt\},[+1])} f = f, \qquad \int_{(\{pt\},[-1])} f = -f$$

Notice that orientation in this case corresponds just to evaluation up to sign. This sign, as we will see studying Stokes theorem, is the same that appears in the fundamental theorem of analysis: $\int_a^b f(x)\mathrm{d}x = f(b) - f(a)$. Here $f(b)$ and $f(a)$ are both functions on a point, but we should think of $\{b\}$ as having the positive orientation and of $\{a\}$ as having the negative orientation. These orientations, as we will see, are induced from the orientation of the interval (a, b).

Example 9.48. Let $U = (a, b)$ be a connected open subset of $\mathbb{R}$, with coordinate x^1. We have two orientations: the class of $\mathrm{d}x^1$ and that of $-\mathrm{d}x^1$. We might think of the first as the geometric orientation given by moving along U from left to right (in the usual graphical description of the real line), the second as moving from right to left. You may see that with the first orientation we move away from the boundary point

$\{a\}$ and approach the boundary point $\{b\}$. This is the origin of the two orientations of the two boundary points.

Example 9.49. Let U be a connected open subset of $\mathbb{R}^2$, with coordinates x^1, x^2. Here the two orientations are the classes of $\mathrm{d}x^1 \wedge \mathrm{d}x^2$ and of $-\mathrm{d}x^1 \wedge \mathrm{d}x^2$. As $-\mathrm{d}x^1 \wedge \mathrm{d}x^2 = \mathrm{d}x^2 \wedge \mathrm{d}x^1$, we might think of the orientations as giving the moving directions along the axes but also the ordering of the axes.

As on $\mathbb{R}^n$ we have distinguished coordinates, with a given ordering, we also have a distinguished volume form:

$$\mathrm{d}^n x := \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n,$$

called the standard volume form. Its class is called the standard orientation. The integral of the top form $f\,\mathrm{d}^n x$ on an open subset U of $\mathbb{R}^n$ with the standard orientation is

$$\int_U f\ \mathrm{d}^n x = \int_U f\,|d^n x| = \int_U f\ d^n x.$$

It is worth recalling that in this funny looking formula $\mathrm{d}^n x$ denotes the standard volume form, $|d^n x|$ the standard density and $d^n x$ the Lebesgue measure. The open subset U of $\mathbb{R}^n$ here plays three different roles: in the first integral it is understood as an oriented manifold, in the second as a manifold and in the third as a measure space. In each case its structure is induced from the corresponding standard structure on $\mathbb{R}^n$.

As the pullback of the standard volume form is the standard volume form multiplied by the Jacobian of the transformation, we immediately get the

Lemma 9.50. *Let $F\colon U \to V$ be a diffeomorphism of open subsets of $\mathbb{R}^n$. Then F is orientation preserving (orientation reversing), with respect to the standard orientation, if and only if* $\det \mathrm{d}F > 0$ *(*$\det \mathrm{d}F < 0$*).*

Remark 9.51. For $n = 0$ the statement has to be interpreted. On $\mathbb{R}^0 = \{0\}$, the standard volume form is $+1$. There is a unique map $\mathbb{R}^0 \to \mathbb{R}^0$. This map is linear and orientation preserving with respect to the standard orientation. Its determinant is by definition $+1$. What sets this case apart is that there is no orientation-reversing diffeomorphism of $\mathbb{R}^0$ with the standard orientation. On the other hand, $\mathbb{R}^n$, $n > 0$, has orientation reversing diffeomorphisms, e.g.,

$$T\colon (x^1, x^2, \dots, x^n) \mapsto (-x^1, x^2, \dots, x^n)).$$

9.4.1. *Orientation by atlases.* The last Lemma naturally leads to a second notion of orientation.

Definition 9.52. An atlas of a manifold is called oriented if all its transition maps are orientation preserving.

Lemma 9.53. *An orientable manifold possesses an oriented atlas.*

Proof. Let M be an n-dimensional orientable manifold, v a volume form and $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ an atlas with connected charts. If $n = 0$ there is nothing to prove, so assume $n > 0$. For every α there is a uniquely defined nowhere vanishing function $\underline{v_\alpha} \in C^\infty(\phi_\alpha(U_\alpha))$ such that $v_\alpha = \underline{v_\alpha}\, \mathrm{d}^n x$. Since $\phi_\alpha(U_\alpha)$ is connected, then necessarily either $\underline{v_\alpha} > 0$ or $\underline{v_\alpha} < 0$. In the second case we change the chart (U_α, ϕ_α) to the chart $(U_\alpha, T \circ \phi_\alpha)$ where T is an orientation-reversing diffeomorphism of $\mathbb{R}^n$ (e.g., the one defined in Remark 9.51). Notice that the atlas obtained by changing charts this way is $\mathcal{C}^\infty$-equivalent to the previous one.

Thus, we have an atlas, which we still denote by $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$, such that $v_\alpha = \underline{v_\alpha}\, \mathrm{d}^n x$ with $\underline{v_\alpha} > 0$ for all α. The transition rules for the v_αs imply

$$\underline{v_\alpha}(x) = \det \mathrm{d}_x\phi_{\alpha\beta}\ \underline{v_\beta}(\phi_{\alpha\beta}(x)).$$

for all $\alpha, \beta \in I$ and for all $x \in \phi_\alpha(U_\alpha \cap U_\beta)$. It then follows that $\det \mathrm{d}_x\phi_{\alpha\beta} > 0$ for all α, β, x. □

Using partition of unity, which is necessary anyway if we want to integrate, we also get the converse statement:

Lemma 9.54. *A Hausdorff, second-countable manifold with an oriented atlas is orientable.*

Proof. As pointed out in Remark 8.58 on page 75, under these assumptions we can construct a positive density σ starting from any atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$. If the atlas is oriented, the transition rules for the representations σ_α are the same as for a top form. Hence, the σ_αs define a volume form. □

There is a natural equivalence relation of oriented atlases which turns out to correspond to the previously introduced equivalence relation of volume forms.

Definition 9.55. Two oriented atlases are equivalent if their union is an oriented atlas.

Lemma 9.56. *Let M be a Hausdorff, second-countable, orientable manifold of positive dimension. Then there is a one-to-one correspondence between orientations of M and equivalence classes of oriented atlases of M.*

Proof. The proof of Lemma 9.53 for $n > 0$ produces a map from orientations to equivalence classes of oriented atlases. In fact any two atlases produced by the construction in the proof starting from the same v are equivalent to each other. On the other hand, if we rescale v by a positive function, we produce an oriented atlas in the same class.

The proof of Lemma 9.54 yields a map in the other direction. Notice that an equivalent oriented atlas produces an equivalent volume form.

Finally observe that the two maps are inverse to each other. □

Remark 9.57. In zero dimensions this Lemma fails. We have seen that $\{pt\}$ has two orientations, like every orientable connected manifold, but has only one atlas (which is obviously oriented). What does not work in the proof is that we do not have an orientation-reversing diffeomorphism of $\mathbb{R}^0$, so we cannot make the map from orientations to atlases injective.

Remark 9.58. If one only works with Hausdorff, second-countable manifolds of positive dimension, because of this equivalence of the two notions one may define a manifold to be orientable if it possesses an oriented atlas and one may define an orientation as an equivalence class of oriented atlases. This is done in several textbooks.

Lemma 9.59. *Let $F\colon M \to N$ be a diffeomorphism of oriented manifolds of positive dimension. Then F is orientation preserving (orientation reversing) if and only if $\det \mathrm{d}F > 0$ ($\det \mathrm{d}F < 0$), where these determinants are computed using representations in any oriented atlases defining the orientation.*

Proof. Let v_M and v_N be representatives of the given orientations. Pick oriented atlases $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ of M and $(V_j, \psi_j)_{j\in J}$ of N corresponding to the orientations: i.e., $(v_M)_\alpha = \underline{v_\alpha}\,\mathrm{d}^n x$ and $(v_N)_j = g_j\,\mathrm{d}^n x$ with $\underline{v_\alpha} > 0$ and $g_j > 0$. Let $F_{\alpha j}$s denote the representations of F. We have

$$(F^*(v_N))_\alpha(x) = \det \mathrm{d}_x F_{\alpha j}\, g_j(F_{\alpha j}(x))\,\mathrm{d}^n x$$

for all α, j, x, and we see that F is orientation preserving (reversing) if and only if all the $\det \mathrm{d}_x F_{\alpha j}$ are positive (negative). □

Remark 9.60. Notice that in an oriented atlas the representations of a top form transform like the representations of a density. One can also easily see that the density defined by a top form ω in the oriented atlas corresponding to an orientation $[v]$ is exactly $\phi_{[v]}\omega$. More explicitly, in an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ we have the representations $\omega_\alpha = \underline{\omega_\alpha}\mathrm{d}^n x$, for uniquely defined functions $\underline{\omega_\alpha}$, and we get $(\phi_{[v]}\omega)_\alpha = \underline{\omega_\alpha}|d^n x|$. It then follows that the integral of ω on an oriented manifold M is given by

$$\int_{(M,[v])} \omega = \sum_{j\in J} \int_{\phi_{\alpha_j}(\operatorname{supp}\rho_j)} (\rho_j)_{\alpha_j}\, \underline{\omega_{\alpha_j}}\, d^n x,$$

where $\{(U_\alpha,\phi_\alpha)\}_{\alpha\in I}$ is an oriented atlas corresponding to $[v]$ (i.e., in which any $v\in[v]$ is represented by a positive volume form) and $\{\rho_j\}_{j\in J}$ is a partition of unity $\{\rho_j\}_{j\in J}$ subordinate to $\{U_\alpha\}_{\alpha\in I}$. If the identifcation between differential forms and densitites on $\phi_{\alpha_j}(U_{\alpha_j})$ is understood, then we can also write

$$\int_{(M,[v])} \omega = \sum_{j\in J} \int_{\phi_{\alpha_j}(\operatorname{supp}\rho_j)} (\rho_j\,\omega)_{\alpha_j}.$$

Remark 9.61. Typically the chosen orientation is understood, so one simply writes $\int_M \omega$.

In the following we will need the following useful remark:

Lemma 9.62 (Localization)**.** *Let $\{(U_\alpha,\phi_\alpha)\}_{\alpha\in I}$ be an oriented atlas of $(M,[v])$ and $\{\rho_j\}_{j\in J}$ a partition of unity subordinate to it. Let ω be a top form with support contained in U_{α_k} for some $k\in J$. Then*

$$\int_{(M,[v])} \omega = \int_{\phi_{\alpha_k}(U_{\alpha_k})} \omega_{\alpha_k}.$$

Proof. We have

$$\begin{aligned}\int_{(M,[v])} \omega &= \sum_{j\in J} \int_{\phi_{\alpha_j}(\operatorname{supp}\rho_j)} (\rho_j\,\omega)_{\alpha_j} = \\ &= \sum_{j\in J} \int_{\phi_{\alpha_j}(\operatorname{supp}\rho_j\cap\operatorname{supp}\omega)} (\rho_j\,\omega)_{\alpha_j} = \int_{\phi_{\alpha_j}(U_{\alpha_j}\cap U_{\alpha_k})} (\rho_j\,\omega)_{\alpha_j} = \\ &= \sum_{j\in J} \int_{\phi_{\alpha_k}(U_{\alpha_j}\cap U_{\alpha_k})} (\rho_j\,\omega)_{\alpha_k} = \sum_{j\in J} \int_{\phi_{\alpha_k}(U_{\alpha_k})} (\rho_j\,\omega)_{\alpha_k} = \\ &= \int_{\phi_{\alpha_k}(U_{\alpha_k})} \sum_{j\in J} (\rho_j\,\omega)_{\alpha_k} = \int_{\phi_{\alpha_k}(U_{\alpha_k})} \omega_{\alpha_k}.\end{aligned}$$

□

9.4.2. *Examples.* Open subsets of $\mathbb{R}^n$ are clearly examples of orientable manifolds. More generally, every parallelizable manifold is orientable. In fact, an isomorphism of TM with $M\times\mathbb{R}^n$ induces, after a choice of basis, an isomorphism $\phi\colon \Lambda^n T^*M \to M\times\mathbb{R}$. If f is any nowhere vanishing function (e.g., f constant and different from zero), then $\phi^* f$ is a volume form.

S^n is also orientable. For $n = 1$, we easily see it as the top form $\mathrm{d}\theta$, where θ is the angle parametrizing S^1, is a volume form. For higher dimensional spheres, we may for example observe that there is an atlas with two charts whose intersection is connected (take, e.g., the atlas defined by the stereographic projection). If this atlas is not oriented, i.e., if the determinant of the Jacobian of the transition map is negative,[31] then we change one of the two chart maps by composing it with an orientation-reversing diffeomorphism of $\mathbb{R}^n$ (e.g., the one defined in Remark 9.51). Another way to see that S^n is oriented is by constructing a volume form explicitly. We may start from the standard volume form on $\mathbb{R}^{n+1}$, write it in polar coordinates and contract it with the vector field $\frac{\partial}{\partial r}$, obtaining a volume form v. Equivalently, we may observe that the Euler vector field, whose flow in polar coordinates clearly just rescales the radius, is $r\frac{\partial}{\partial r}$. Therefore, we have $v = \frac{1}{r}\iota_E \mathrm{d}^{n+1}x$. We may also get rid completely of the radius dependence if we rescale v by r^n. Namely, we have the volume form

(9.10)
$$\omega := \frac{1}{r^{n+1}}\iota_E \mathrm{d}^{n+1}x = \frac{\sum_{i=1}^{n+1}(-1)^{i+1}x^i \mathrm{d}x^1 \wedge \cdots \wedge \widehat{\mathrm{d}x^i} \wedge \cdots \wedge \mathrm{d}x^{n+1}}{||x||^{n+1}},$$

where the restriction to S^n is understood.

Another example is the tangent bundle TM of any manifold, also a non orientable one. If $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ is any atlas of M, then the atlas $(\widehat{U}_\alpha, \widehat{\phi}_\alpha)_{\alpha\in I}$ introduced in subsection 6.2 is oriented. In fact, by (6.2), we have $\det \mathrm{d}_x\widehat{\phi}_{\alpha\beta} = (\det \mathrm{d}_x\phi_{\alpha\beta})^2$. Similarly, one shows that T^*M is also orientable.

An example of non orientable manifold is the Möbius band. This is obtained by quotienting $[0, 1] \times \mathbb{R}$ by the equivalence relation $(0, y) \sim (1, -y)$. There is an atlas with two charts: one having domain $(0, 1)\times\mathbb{R}$, the other having domain $([0, 1/4) \cup (3/4, 1]) \times \mathbb{R}$. On $(0, 1) \times \mathbb{R}$ and on $[0, 1/4) \times \mathbb{R}$ the chart maps are defined to be the inclusions into $\mathbb{R} \times \mathbb{R}$. The chart map on $(3/4, 1] \times \mathbb{R}$ is defined by $(x, y) \mapsto (x, -y)$. The intersection of the two chart domains is $((0, 1/4) \cup (3/4, 1)) \times \mathbb{R}$ and one easily sees that this atlas is not oriented. More precisely, one sees that the transition map corresponding to $(0, 1/4) \times \mathbb{R}$ is orientation preserving (actually, the identity), whereas the transition map corresponding to $(3/4, 1) \times \mathbb{R}$ is orientation reversing. To prove that the Möbius strip is not orientable, we show that it does not possess a volume form. Assume on the contrary that we had a volume form

[31] This is actually the case of the stereographic projection with atlas as in Example 4.12.

v. On the first chart it would be represented by $f\mathrm{d}x \wedge \mathrm{d}y$ and on the second by $g\mathrm{d}x \wedge \mathrm{d}y$ where f and g are nowhere vanishing functions. As the charts are connected (notice that $([0,1/4) \cup (3/4,1]) \times \mathbb{R}$ is connected in the atlas topology), they have a definite sign. Suppose e.g. that $f > 0$. Using the transition map corresponding to $(0,1/4) \times \mathbb{R}$ we would conclude that also $g > 0$. However, using the transition map corresponding to $(3/4,1) \times \mathbb{R}$ we would conclude that $g < 0$, which is a contradiction.

A typical way to generate nonorientable manifolds is by a suitable quotient. An involution on a manifold M is a map $\psi\colon M \to M$ satisfying $\psi \circ \psi = \mathrm{Id}$. An involution ψ defines an equivalence relation by $x \sim_\psi y$ if $x = y$ or $x = \psi(y)$. If M is Hausdorff and ψ has no fixed points (i.e., $\psi(x) \neq x$ for all x), then the quotient $M/\sim_\psi$ has a smooth structure for which the canonical projection $\pi\colon M \to M/\sim_\psi$ is a smooth submersion. (Note that this structure is unique up to diffeomorphisms since π is then a surjective submersion.) In fact, we can pick an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ on M with the property that $U_\alpha \cap \psi(U_\alpha) = \emptyset$ for all α.[32] Observe then that π restricted to U_α defines a diffeomorphism with its image $\widetilde{U}_\alpha = \pi(U_\alpha)$. We then define an atlas $\{(\widetilde{U}_\alpha, \widetilde{\phi}_\alpha)\}_{\alpha\in I}$ on $M/\sim_\psi$ by setting $\widetilde{\phi}_\alpha = \phi_\alpha \circ \pi^{-1}$.[33] Note that the transition map $\widetilde{\phi}_{\alpha\beta}$ is equal to $\phi_\beta^{-1} \circ \phi_\alpha$ if $U_\alpha \cap \mathrm{U}_\beta \neq \emptyset$ and to $\phi_\beta^{-1} \circ \psi \circ \phi_\alpha$ otherwise. Also notice that the quotient manifold is also Hausdorff.

Lemma 9.63. *Let M be a connected orientable manifold and ψ an involution with no fixed points. Then $M/\sim_\psi$ is orientable if and only ψ is orientation preserving.*

Proof. If v is a volume form on $M/\sim_\psi$, then π^*v is a volume form on M. If ψ were orientation reversing, then we would have $-[\pi^*v] = \psi^*[\pi^*v] = [\psi^*\pi^*v] = [\pi^*v]$, which is a contradiction.

For the other implication, assume that ψ is orientation preserving. Then, using an oriented atlas on M for the construction of the atlas of $M/\sim_\psi$, we get that the transition maps $\widetilde{\phi}_{\alpha\beta}$ are also orientation preserving. Hence, M is orientable. □

[32] For example, observe that for each $x \in M$ we can find disjoint open neighborhoods N_x and $O_{\psi(x)}$ of x and $\psi(x)$, respectively. It follows that $W_x := N_x \cap \psi^{-1}(O_{\psi(x)})$ is an open neighborhood of x with the property that $W_x \cap \psi(W_x) = \emptyset$ (in fact W_x is contained in N_x and $\psi(W_x)$ is contained in $O_{\psi(x)}$). As a consequence, if we start with any atlas $\{(V_j, \psi_j)\}_{j\in J}$, we may define an equivalent atlas with the desired property by setting $I = J \times M$ and $U_{(j,x)} := V_j \cap W_x$.

[33] This construction is a particular instance of a general construction to define a smooth structure on a quotient by a properly discontinuous, free action of a discrete group.

As an application of this construction we consider the real projective space $\mathbb{RP}^n$, i.e., the space of lines through the origin in $\mathbb{R}^{n+1}$. We may realize $\mathbb{RP}^n$ as the quotient of $\mathbb{R}^{n+1}$ by the equivalence relation $x \sim \lambda x$ for all $\lambda \neq 0$. Alternatively, we may first quotient just by using positive λs, thus getting the sphere S^n, and then quotienting by $\lambda = \pm 1$. That is, we get $\mathbb{RP}^n = S^n / \sim_\psi$, where $\psi\colon S^n \to S^n$ is the antipodal map sending x to $-x$. By the above construction this shows that $\mathbb{RP}^n$ is a smooth manifold. Using the volume form of equation (9.10), we see that ψ is orientation preserving if and only if n is odd. Therefore, $\mathbb{RP}^n$ is orientable if and only if n is odd.

9.4.3. *Restriction and integration.* Recall that differential forms may be pulled back. If $F\colon M \to N$ is a smooth map and ω is an n-form on N, with $n = \dim M$, then $F^*\omega$ is a top form on M which can be integrated, provided we have an orientation and a partition of unity on M.

A particular case is the inclusion of an oriented submanifold $\iota\colon S \hookrightarrow M$. If ω is an s-form on M, $s = \dim S$, one then simply writes $\int_S \omega$ to denote the integral of the pullback of ω by ι over S with its given orientation.

One usually also defines $\int_S \omega = 0$ when the degree of ω is different form the dimension of S.

Remark 9.64 (Integration of 1-forms)**.** If $\omega = \sum_{i=1}^n \omega_i \mathrm{d}x^i$ is a 1-form on an open subset U of $\mathbb{R}^n$ and $\gamma\colon I \to U$ is a differentiable curve, then $\gamma^*\omega(t) = \sum_{i=1}^n \omega_i(\gamma(t))\dot\gamma^i \mathrm{d}t$. Using the standard orientation of $I \subset \mathbb{R}$, we identify $\mathrm{d}t$ with the Lebesgue meaure dt. Hence we get the familiar formula

$$\int_\gamma \omega = \int_I \gamma^*\omega = \int_I \sum_{i=1}^n \omega_i(\gamma(t))\dot\gamma^i dt.$$

9.4.4. *Computing the divergence.* We now show how Cartan's formula may be used to simplify the derivations in subsubsection 8.2.6 on an oriented manifold and to recover the local formula (8.14) for the divergence of a vector field.

First observe that using an orientation $[v]$ we can easily compute the divergence of a vector field by Cartan's formula.

Lemma 9.65. *Let* $[v]$ *be an orientation. Then*

$$\mathsf{L}_X \phi_{[v]} = \phi_{[v]} \mathsf{L}_X$$

for all vector fields X*.*

Proof. Let $v \in [v]$ and $\omega = fv$ a top form. Then $\mathsf{L}_X\omega = X(f)v + f\mathsf{L}_X v$. Since v is a volume form, there is a uniquely defined function g such that $\mathsf{L}_X v = gv$. Hence $\phi_{[v]}\mathsf{L}_X\omega = (X(f) + fg)|v|$. On the other hand, $\phi_{[v]}\omega = f|v|$. Hence $\mathsf{L}_X\phi_{[v]}\omega = X(f)|v| + f\mathsf{L}_X|v|$. Thus, we have just to show that $\mathsf{L}_X|v| = g|v|$. This is clear if we go in a chart where we have $|v_\alpha| = \pm v_\alpha$. □

Once an orientation $[v]$ is fixed, we identify top forms and densities by $\phi_{[v]}$. If σ is a density, by abuse of notation we denote by σ also the corresponding top form (i.e., $(\phi_{[v]})^{-1}\sigma$). Then by the above Lemma and by Remark 9.30 on page 103, we have

$$\mathsf{L}_X\sigma = \mathrm{d}\iota_X\sigma.$$

This immediately implies

$$\mathsf{L}_{fX}\sigma = \mathsf{L}_X(f\sigma) = X(f)\sigma + f\mathsf{L}_X\sigma.$$

Remark 9.66 (The divergence in a chart)**.** If U is an open subset of $\mathbb{R}^n$, we can write every top form ω as $\omega = f\mathrm{d}^n x$ for a uniquely defined function f. Then

$$\iota_X\omega = \sum_{i=1}^n (-1)^{i+1} fX^i \mathrm{d}x^1 \wedge \cdots \wedge \widehat{\mathrm{d}x^i} \wedge \cdots \wedge \mathrm{d}x^n.$$

Hence

$$\mathsf{L}_X\omega = \sum_{i=1}^n \partial_i(fX^i)\,\mathrm{d}^n x.$$

Notice that on U we can identify top forms and densities, using the atlas with one chart. If ω is a positive density, we then have

$$\mathrm{div}_\omega X = \frac{1}{f}\sum_{i=1}^n \partial_i(fX^i).$$

In particular, we can apply this to a density σ on a (possibly non orientable) manifold M. In a chart (U_α, ϕ_α), the representation σ_α of σ defines a top form $\sigma_\alpha \mathrm{d}^n x$. From this we recover equation (8.14).

9.4.5. *Integration of vector fields.* Using a Riemannian metric g on a manifold M, one can define the integral of a vector field X on a curve $\gamma\colon I \to M$ by

$$\int_\gamma X \cdot \mathrm{d}x := \int_I g_{\gamma(t)}(X_{\gamma(t)}, \dot\gamma(t))\, dt.$$

This integral is usually called the work of X along γ (this term comes from the application in mechanics when X is a force field). By Remark 8.82 we can reduce this to the integral of 1-forms. Let $\omega_X :=$

$\Phi_g(X)$ with Φ_g the isomorphism between TM and T^*M defined in Remark 8.82. Then we have

$$\int_\gamma X \cdot \mathrm{d}x = \int_\gamma \omega_X. \tag{9.11}$$

Given orientations one can also integrate vector fields on hypersurfaces (i.e., embedded submanifolds of codimension one) in Riemannian manifolds. This is called the flux of the vector field through the hypersurface.

To do this, we have to introduce some notations. Let (M, g) be a Riemannian manifolds and S a hypersurface. Using g_p we can define the orthogonal complement $(T_pS)^\perp$ of T_pS in T_pM for all $p \in S$. A normalized vector n_p in $(T_pS)^\perp$ is called a normal vector at p. Explicitly, n_p satisfies the equations $g_p(n_p, u) = 0\ \forall u \in T_pS$ and $g_p(n_p, n_p) = 1$. Notice that, since $\dim(T_pS)^\perp = \dim T_pM - \dim T_pS = 1$, there are exactly two normal vectors at each p (related to each other by multiplication by -1). Using orientations we can select one of them. Namely, let $[v]$ be an orientation on M and $[w]$ an orientation on S. If we denote by ι_S the inclusion of S into M, we say that n_p is compatible with orientations if $\iota_S^*\iota_{n_p} v_p = c\, w_p$ with $v \in [v]$, $w \in [w]$ and c a positive constant (notice that this does not depend on the choices of v and w).[34] The map $p \mapsto n_p$ is called the normal vector field to S (notice that it is not a vector field on S nor on M, but a section of the pullback of TM to S). If X is a vector field on M, then we denote by $X \cdot n$ the function on S defined by

$$(X \cdot n)(p) := g_p(X_p, n_p).$$

Finally, the integral $\int_S X \cdot n\, v_{g_S}$, where v_{g_S} denotes the Riemannian density of the restriction of g_S of g to S, is called the flux of X through S.

We now want to relate the flux to the integral of a differential form. Notice that using the orientation of M we may identify the Riemannian density v_g with a top form on M. Using the orientation of S we can then integrate the top form $\iota_X v_g$ on S.

Proposition 9.67. *Let S be an oriented hypersurface of an oriented Riemannian manifold (M, g). Then, for every $X \in \mathfrak{X}(M)$, we have*

$$\int_S X \cdot n\, v_{g_S} = \int_S \iota_X v_g.$$

[34] Conversely, if S is orientable, then a continuous choice of n_ps induces an orientation $[w]$ of S by $w_p := \iota_S^*\iota_{n_p} v_p$ for all $p \in S$. This also shows that S is orientable if and only it possesses a continuous choice of normal vectors.

It is worth noticing that on the left hand side we have the integration of a density on the manifold S (and the orientations are used to determine n) whereas on the right hand side we have the integration of a top form on the oriented manifold S (and the orientation of M is used to identify the Riemannian density with a top form).

Proof. To prove this result we just show that the integrands are equal in every chart images, where we used an adapted atlas.

Let (U, ϕ) be an adapted chart of M around a point $p \in S$. Recall that this means that we have a chart (V, ψ) of S around p with $V = \phi^{-1}(\{(x^1, \dots, x^n) \in \phi(U) \mid x^n = 0\})$ and $\psi = \phi|_V$. We expand

$$g_U = \sum_{i=1}^{n-1} h_{ij}\, \mathrm{d}x^i \mathrm{d}x^j + \sum_{i=1}^{n-1} b_i\, (\mathrm{d}x^i \mathrm{d}x^n + \mathrm{d}x^n \mathrm{d}x^i) + \alpha\, (\mathrm{d}x^n)^2.$$

That is, we write the components of g_U as a block matrix

$$\mathsf{g} = \begin{pmatrix} \mathsf{h} & \mathbf{b} \\ \mathbf{b}^t & \alpha \end{pmatrix}.$$

The first remark is that $(g_S)_V = \sum_{i=1}^{n-1} h_{ij}\, \mathrm{d}x^i \mathrm{d}x^j$. In particular, h is positive definite. Moreover, $(v_{g_S})_V = \sqrt{\det \mathsf{h}}$. Next we compute n. In the chart image we write it as a block vector $\left(\begin{smallmatrix} \mathbf{k} \\ \mu \end{smallmatrix}\right)$. We have

$$\begin{pmatrix} \mathsf{h} & \mathbf{b} \\ \mathbf{b}^t & \alpha \end{pmatrix} \begin{pmatrix} \mathbf{k} \\ \mu \end{pmatrix} = \begin{pmatrix} \mathsf{h}\mathbf{k} + \mu \mathbf{b} \\ \mathbf{b}^t \mathbf{k} + \mu \alpha \end{pmatrix}.$$

Requiring n to be orthogonal to vectors tangent to S then yields $\mathsf{h}\mathbf{k} + \mu \mathbf{b} = 0$ which can be solved to give $\mathbf{k} = -\mu \mathsf{h}^{-1} \mathbf{b}$ and hence

$$\begin{pmatrix} \mathsf{h} & \mathbf{b} \\ \mathbf{b}^t & \alpha \end{pmatrix} \begin{pmatrix} \mathbf{k} \\ \mu \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mu(\alpha - \mathbf{b}^t \mathsf{h}^{-1} \mathbf{b}) \end{pmatrix}.$$

The normalizing condition finally gives $\mu^2(\alpha - \mathbf{b}^t \mathsf{h}^{-1} \mathbf{b}) = 1$, which in particular shows that $\alpha - \mathbf{b}^t \mathsf{h}^{-1} \mathbf{b} > 0$. Finally, we write X_U also as a block vector $\left(\begin{smallmatrix} \mathbf{v} \\ \xi \end{smallmatrix}\right)$ and hence we have

$$(X \cdot n)_V = \xi \mu (\alpha - \mathbf{b}^t \mathsf{h}^{-1} \mathbf{b}) = \frac{\xi}{\mu}.$$

Finally,

$$(X \cdot n\, v_{g_S})_V = \frac{\xi}{\mu} \sqrt{\det \mathsf{h}}.$$

Notice that μ is a function of the components of g_U. To make this explicit we have to extract a square root whose sign choice is determined by the orientations. Namely, let v be a representative of $[v]$ such that $v_U = \epsilon\, \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n$ with $\epsilon = \pm 1$ and let w be a representative of $[w]$ such that $w_V = \widetilde{\epsilon}\, \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^{n-1}$ with $\widetilde{\epsilon} = \pm 1$. Since $(\iota_S^* \iota_{n_p} v_p)_V =$

$(-1)^{n-1}\mu\epsilon\,\mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^{n-1}$, we get that the sign of μ is $(-1)^{n-1}\epsilon\widetilde{\epsilon}$. Hence

$$\mu = \frac{(-1)^{n-1}\epsilon\widetilde{\epsilon}}{\sqrt{\alpha - \mathbf{b}^t\mathsf{h}^{-1}\mathbf{b}}},$$

so

$$(X \cdot n\, v_{g_S})_V = (-1)^{n-1}\epsilon\,\widetilde{\epsilon}\,\xi\sqrt{\det \mathsf{h}}\,\sqrt{\alpha - \mathbf{b}^t\mathsf{h}^{-1}\mathbf{b}}.$$

We now consider the right hand side of the equation in the Proposition. Using the orientation of M we may identify the Riemannian density v_g with a volume form and we have $(v_g)_U = \epsilon\,\sqrt{\det \mathsf{g}}\,\mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n$. Hence $(\iota_S^*\iota_X v_g)_V = (-1)^{n-1}\epsilon\,\xi\,\sqrt{\det \mathsf{g}}\,\mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^{n-1}$. We finally have to use the orientation on S to regard it as a density and get $(-1)^{n-1}\epsilon\widetilde{\epsilon}\,\xi\,\sqrt{\det \mathsf{g}}$. The proposition is now proved since[35]

$$\det \mathsf{g} = (\alpha - \mathbf{b}^t\mathsf{h}^{-1}\mathbf{b})\det \mathsf{h}.$$

Notice that from the positivity of g and the consequent positivity of h we conclude again that $\alpha - \mathbf{b}^t\mathsf{h}^{-1}\mathbf{b} > 0$. □

9.5. **Manifolds with boundary and Stokes theorem.** We finally return to Stokes theorem, the higher dimensional generalization of the fundamential theorem of analysis. Our goal is to show that if ω is a smooth $(n-1)$-form with compact support on an oriented n-dimensional manifold with boundary M, then

$$\boxed{\int_M \mathrm{d}\omega = \int_{\partial M} \omega}$$

To do this we will have to introduce the notion of manifolds with boundary, define the boundary and show that it is a manifold, discuss orientation and finally prove the formula. As for the formula itself, it is actually enough to understand the local case.

As a warm up we start with the two-dimensional case, which basically presents all the features of the general case. Let $\mathbb{H}^2 := \{(x,y) \in \mathbb{R}^2 \mid y \geq 0\}$ be the upper half plane. Let $\omega = \omega_x \mathrm{d}x + \omega_y \mathrm{d}y$ be a smooth 1-form on $\mathbb{H}^2$: by this we mean that ω is the restriction to $\mathbb{H}^2$ of a smooth form defined on an open neighborhood of $\mathbb{H}^2$ in $\mathbb{R}^2$. We want

[35] This follows from the general rule

$$\det\begin{pmatrix} \mathsf{A} & \mathsf{B} \\ \mathsf{C} & \mathsf{D} \end{pmatrix} = \det(\mathsf{D} - \mathsf{C}\mathsf{A}^{-1}\mathsf{B})\det \mathsf{A},$$

where $\left(\begin{smallmatrix} \mathsf{A} & \mathsf{B} \\ \mathsf{C} & \mathsf{D} \end{smallmatrix}\right)$ is a block matrix with A invertible. This formula follows immediately from the decomposition

$$\begin{pmatrix} \mathsf{A} & \mathsf{B} \\ \mathsf{C} & \mathsf{D} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \mathsf{C}\mathsf{A}^{-1} & 1 \end{pmatrix}\begin{pmatrix} 1 & \mathsf{B} \\ 0 & \mathsf{D} - \mathsf{C}\mathsf{A}^{-1}\mathsf{B} \end{pmatrix}\begin{pmatrix} \mathsf{A} & 0 \\ 0 & 1 \end{pmatrix}.$$

to consider the integral of $\mathrm{d}\omega$ on $\mathbb{H}^2$. To be sure of convergence we assume that ω, and hence $\mathrm{d}\omega$, has compact support. First observe that $\mathrm{d}\omega = (\partial_x\omega_y - \partial_y\omega_x)\,\mathrm{d}x\wedge\mathrm{d}y$. Thus,

$$\int_{\mathbb{H}^2}\mathrm{d}\omega = \int_{\mathbb{H}^2}(\partial_x\omega_y - \partial_y\omega_x)\,dx\,dy.$$

By Fubini's theorem we have

$$\int_{\mathbb{H}^2}\partial_x\omega_y\,dx\,dy = \int_0^{+\infty}\left(\int_{-\infty}^{+\infty}\partial_x\omega_y\,dx\right)\,dy.$$

By the fundamental theorem of analysis we have $\int_{-\infty}^{+\infty}\partial_x\omega_y\,dx = 0$ since ω has compact support. Again by Fubini's theorem we have

$$\int_{\mathbb{H}^2}\partial_y\omega_x\,dx\,dy = \int_{-\infty}^{+\infty}\left(\int_0^{+\infty}\partial_y\omega_x\,dy\right)\,dx.$$

The fundamental theorem of analysis and the fact that ω has compact support now imply that $\int_0^{+\infty}\partial_y\omega_x\,dy = -\omega_x|_{y=0}$. If we denote by $\partial\mathbb{H}^2 = \{(x,y)\in\mathbb{R}^2 \mid y=0\}$ the boundary of $\mathbb{H}^2$ (i.e., the x-axis), we finally get

$$\int_{\mathbb{H}^2}\mathrm{d}\omega = \int_{\partial\mathbb{H}^2}\omega.$$

This result can easily be generalized to the upper half spaces

$$\mathbb{H}^n := \{(x^1,\dots,x^n)\in\mathbb{R}^n \mid x^n \geq 0\}.$$

Again by a smooth differential form on $\mathbb{H}^n$ we mean the restriction to $\mathbb{H}^n$ of a smooth differential form defined on an open neighborhood of $\mathbb{H}^n$ in $\mathbb{R}^n$. To integrate $\mathrm{d}\omega$ on $\mathbb{H}^n$ (or, if you prefer to work with manifolds and use the theory developed so far, on the interior of $\mathbb{H}^n$) we pick the standard orientation d^nx. In order to avoid signs in the Stokes theorem, on the boundary

$$\partial\mathbb{H}^n = \{(x^1,\dots,x^n)\in\mathbb{R}^n \mid x^n = 0\}$$

we take the orientation induced by the outward pointing vector field $-\partial_n$, i.e.,

$$[i^*\iota_{-\partial_n}\mathrm{d}^nx] = (-1)^n[\mathrm{d}x^1\wedge\cdots\wedge\mathrm{d}x^{n-1}],$$

where i denotes the inclusion map of $\partial\mathbb{H}^n$ into $\mathbb{H}^n$.[36] We then have a first version of Stokes theorem:

[36]If we had been working with lower half spaces, defined by the condition $x^n \leq 0$, we would have chosen the boundary orientation also by the outward pointing vector field, which in this case would have been ∂_n, without an extra sign.

Lemma 9.68. *Let ω be a smooth $(n-1)$-form on $\mathbb{H}^n$ with compact support. Then, using the orientations defined above, we have*

$$\int_{\mathbb{H}^n} \mathrm{d}\omega = \int_{\partial\mathbb{H}^n} \omega.$$

Proof. We write $\omega = \sum_{j=1}^n (-1)^{j-1}\,\omega^j\,\mathrm{d}x^1\wedge\cdots\wedge\widehat{\mathrm{d}x^j}\wedge\cdots\wedge\mathrm{d}x^n$. (Notice that the components ω^j are related by a sign to the components $\omega_{i_1\dots i_{n-1}}$ of the usual notation.) Then $\mathrm{d}\omega = \sum_{j=1}^n \partial_j\omega^j\,\mathrm{d}^n x$. Hence, using the standard orientation,

$$\int_{\mathbb{H}^n} \mathrm{d}\omega = \sum_{j=1}^n \int_{\mathbb{H}^n} \partial_j\omega^j\, d^n x.$$

By Fubini's theorem we can integrate the jth term first along the jth axis. Since ω has compact support, for $j<n$ we get

$$\int_{-\infty}^{+\infty} \partial_j\omega^j\, dx^j = 0,$$

whereas for $j=n$ we get

$$\int_0^{+\infty} \partial_n\omega^n\, dx^n = -\omega^n|_{x_n=0}.$$

Thus,

$$\int_{\mathbb{H}^n} \mathrm{d}\omega = -\int_{\partial\mathbb{H}^n} \omega^n\, d^{n-1}x.$$

On the other hand $i^*\omega = (-1)^{n-1}\,\omega^n|_{x_n=0}\,\mathrm{d}x^1\wedge\cdots\wedge\mathrm{d}x^{n-1}$. Using the orientation of $\partial\mathbb{H}^n$ defined above, we finally see that also

$$\int_{\partial\mathbb{H}^n} \omega = -\int_{\partial\mathbb{H}^n} \omega^n\, d^{n-1}x,$$

which concludes the proof. □

By the same argument one proves the following

Lemma 9.69. *If ω is a smooth $(n-1)$-form on $\mathbb{R}^n$ with compact support, then*

$$\int_{\mathbb{R}^n} \mathrm{d}\omega = 0.$$

9.5.1. *Manifolds with boundary.* To extend this result to manifolds we first have to extend the notion of manifolds allowing charts to take values in $\mathbb{H}^n$ (with the topology induced from $\mathbb{R}^n$). This results in the notion of manifold with boundary. This is the standard, though very unfortunate, terminology, since a manifold with boundary is *not* a manifold with extra structure (a boundary), but a generalization: it is a manifold instead that is a special case of a manifold with boundary. As we will see below a manifold is a manifold with boundary whose boundary is empty.

Since we want to consider charts with image in $\mathbb{H}^n$, we also have to understand the properties of the corresponding transition maps. We focus only on the smooth case for simplicity. On $\mathbb{H}^n$ we always use the topology induced from the standard topology on $\mathbb{R}^n$.

Definition 9.70. A map $U \to \mathbb{H}^m$, where U is an open subset of $\mathbb{H}^n$, is called smooth if it is the restriction to U of a smooth map defined on an open neighborhood of U in $\mathbb{R}^n$. A diffeomorphism $U \to V$, where U and V are open subsets of $\mathbb{H}^n$, is a smooth invertible map whose inverse is smooth.

We call $\partial\mathbb{H}^n := \{(x^1,\dots,x^n) \in \mathbb{R}^n \mid x^n = 0\}$ the boundary of $\mathbb{H}^n$ and

$$\mathring{\mathbb{H}}^n = \{(x^1,\dots,x^n) \in \mathbb{R}^n \mid x^n > 0\}$$

the interior of $\mathbb{H}^n$. If U is a subset of $\mathbb{H}^n$ we define its interior and boundary as

$$\mathring{U} := U \cap \mathring{\mathbb{H}}^n \text{ and } \partial U := U \cap \partial\mathbb{H}^n.$$

Elements of these sets are called interior and boundary points of U, respectively.

Lemma 9.71. *Let $F\colon U \to V$ be a diffeomorphism of open subsets of $\mathbb{H}^n$. Then F maps interior points to interior points and boundary points to boundary points. Moreover,*

$$F|_{\mathring{U}}\colon \mathring{U} \to \mathring{V} \text{ and } F|_{\partial U}\colon \partial U \to \partial V$$

are diffeomorphisms. If v is a vector in $\mathbb{R}^n$ and $p \in \partial U$, then the sign of the last component of v is equal to the sign of the last component of d_pFv. Finally, if F is orientation preserving, then also the restrictions $F|_{\mathring{U}}$ and $F|_{\partial U}$ are so.

Proof. Recall that, by definition, F is the restriction of a diffeomorphism $\widetilde{F}\colon \widetilde{U} \to \mathbb{R}^n$ with $\widetilde{U}$ an open neighborhood of U in $\mathbb{R}^n$.

Fix a point q in U. Let W be an open neighborhood of q in $\mathbb{R}^n$. Since $\widetilde{F}^{-1}$ is in particular continuous, $\widetilde{F}(W)$ is open in $\mathbb{R}^n$. Now assume that

$F(q) \in \partial V$. Then every open neighborhood of $F(q)$ in $\mathbb{R}^n$ contains points not belonging to V. In particular, this is true for $\widetilde{F}(W)$, so W does also contain points not belonging to U. Since this is true for every open neighborhood W of q in $\mathbb{R}^n$, this implies that also q is a boundary point.

Next observe that the restriction of a diffeomorphism to an open subset is still a diffeomorphism, so $F|_{\mathring{U}} \colon \mathring{U} \to \mathring{V}$ is a diffeomorphism. This restriction is of course orientation preserving if F is.

Let us now write F in components $F^1, \dots, F^n$. Since F maps the boundary to the boundary, we have that $F^n(x^1, \dots, x^{n-1}, 0) = 0$. This implies that $\partial_i F^n|_{x^n=0} = 0$, $i = 1, \dots, n-1$. The Jacobian of F on the boundary can then be written as a block matrix

$$\begin{pmatrix} (\partial_i F^j)_{i,j=1,\dots,n-1} & (\partial_n F^j)_{j=1,\dots,n-1} \\ \mathbf{0}^t & \partial_n F^n \end{pmatrix}.$$

Moreover, since F maps interior points to interior points we have that $\partial_n F^n|_{x^n=0} > 0$. Since $\mathrm{d}F$ is invertible, it then follows that also $(\partial_i F^j)_{i,j=1,\dots,n-1}|_{x^n=0}$ is invertible. But this is the Jacobian of the restriction of F to ∂U.

If $v = (v^1, \dots, v^n)$ is a vector and $p \in \partial U$, then we have $(\mathrm{d}_p F v)^n = \partial_n F^n(p)\, v^n$. Since $\partial_n F^n(p) > 0$, the signs of the last components agree.

Finally, notice that, if F is orientation preserving, then $\det \mathrm{d}F > 0$. From the above block form and from the remark that $\partial_n F^n|_{x^n=0} > 0$ it follows that also $\det((\partial_i F^j)_{i,j=1,\dots,n-1}|_{x^n=0})$ is positive, so F restricted to the boundary is also orientation preserving. □

We now extend the notion of a chart on a set M as a pair (U, ϕ) where U is a subset of M and ϕ is an injective map from U to $\mathbb{H}^n$ for some n. We extend the notion of transition map, atlas, open atlas. We say that an atlas is smooth if all transition maps are smooth as in Definition 9.70. Two smooth atlases are defined to be equivalent if their union is a smooth atlas.

Definition 9.72. An n-dimensional smooth manifold with boundary is an equivalence class of smooth atlases whose charts take values in $\mathbb{H}^n$.

Let M be a manifold with boundary and q a point in M. If q is sent by a chart map to an interior point of $\mathbb{H}^n$, then by Lemma 9.71 it will be sent to an interior point by every chart map; on the other hand, if q is sent to a boundary point by a chart map, then it will be sent to a boundary point by every chart map. This means that we have the notion of interior and boundary points of a manifold with boundary: interior points are mapped to interior points of $\mathbb{H}^n$ by chart maps and boundary points to boundary points. We denote by $\mathring{M}$ the set of interior points of M and by ∂M the set of boundary points of M. Lemma 9.71 implies that $\mathring{M}$ and ∂M get a manifold structure, with $\dim \mathring{M} = \dim M = \dim \partial M + 1$, just by restricting atlases of M. The manifold ∂M is called the boundary of M. A compact manifold with boundary M with $\partial M = \emptyset$ is also called a closed manifold. Notice that this terminology is a bit confusing: closed here has a different meaning than in point-set topology.

Example 9.73. Any open subset U of $\mathbb{H}^n$ is a smooth manifold with boundary with the equivalence class of the atlas consisting of a single chart and the inclusion map to $\mathbb{H}^n$ as the chart map. We have $\mathring{U} = U \cap \mathring{\mathbb{H}}^n$ and $\partial U = U \cap \partial\mathbb{H}^n$.

Example 9.74. Every smooth manifold M is also a smooth manifold with boundary. Namely, to any atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ of M we associate an atlas $(U_\alpha, \widetilde{\phi}_\alpha)$ with $\widetilde{\phi}_\alpha := F \circ \phi_\alpha$ and F a diffeomorphism $\mathbb{R}^n \to \mathring{\mathbb{H}}^n$ (e.g., $F\colon (x^1, \dots, x^n) \mapsto (x^1, \dots, x^{n-1}, \mathrm{e}^{x^n})$). We have $\mathring{M} = M$ and $\partial M = \emptyset$, so we have a one-to-one correspondence between manifolds and manifolds with boundary whose boundary is empty.

Example 9.75. An interval I in $\mathbb{R}$ is an example of a one-dimensional manifold with boundary. If $I = [a, b]$, then $\partial I = \{a\} \cup \{b\}$; if $I = [a, b)$, then $\partial I = \{a\}$; if $I = (a, b]$, then $\partial I = \{b\}$; if $I = (a, b)$, then $\partial I = \emptyset$. In all cases, $\mathring{I} = (a, b)$.

Example 9.76. If M is a manifold and N is a manifold with boundary, then $M \times N$ is a manifold with boundary by using a product atlas. The interior of $M \times N$ is $M \times \mathring{N}$, and $\partial(M \times N) = M \times \partial N$.

Example 9.77. Let M be an oriented manifold and S an oriented hypersurface in M. Pick a Riemannian structure on M (recall that this is always possible if M is Hausdorff and second countable) and let n denote the normal vector field to S (see subsection 9.4.5 for notations). An adapted oriented atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ for M yields an oriented atlas $(V_\alpha, \psi_\alpha)_{\alpha\in I}$ with $V_\alpha := \phi_\alpha^{-1}(\{(x^1, \dots, x^n) \in \phi_\alpha(U_\alpha) \mid x^n = 0\})$ and $\psi_\alpha := T_\alpha \circ \phi_\alpha|_{V_\alpha}$ with $T_\alpha\colon (x^1, \dots, x^{n-1}, x^n) \mapsto (x^1, \dots, x^{n-1}, -x^n)$ if the nth component of the normal vector field in this chart is negative and the identity map otherwise (we assume for simplicity that the V_αs are connected). (Notice that some V_α may be empty). Pick a collection $\epsilon_\alpha \in \mathbb{R}_{>0} \cup \{+\infty\}$ and let $W_\alpha := \phi_\alpha^{-1}(\{(x^1, \dots, x^n) \in \phi_\alpha(U_\alpha) \mid 0 \le x^n < \epsilon_\alpha\})$. Then $S_\epsilon := \bigcup_\alpha W_\alpha$ is a manifold with boundary whose interior is an open subset of M and whose boundary is S.

Example 9.78. Let W be an open subset of $\mathbb{R}^{n+1}$ and $F\colon W \to \mathbb{R}$ a smooth map. Assume that for all $c \in M := F^{-1}(0)$ the map $\mathrm{d}_c F$ is surjective. Then, for all $\epsilon \in \mathbb{R}_{>0} \cup \{+\infty\}$, $M_\epsilon := F^{-1}([0,\epsilon))$ is a smooth manifold with boundary with $\partial M_\epsilon = M$ and interior the open subset $F^{-1}((0,\epsilon))$ of $\mathbb{R}^{n+1}$.

As a particular case we have the

Example 9.79. The closed n-dimensional ball

$$B^n := \left\{ (x^1,\dots,x^n) \in \mathbb{R}^n \mid \sum_{i=1}^n (x^i)^2 \le 1 \right\}$$

is a smooth manifold with boundary. Its interior is the open n-dimensional ball and its boundary is the $(n-1)$-dimensional sphere S^{n-1}. (Notice that the closed ball is not a closed manifold.)

What we have defined for manifolds can be extended to manifolds with boundary. We start with maps. If M and N are manifolds with boundary and $F\colon M \to N$ a set-theoretic map, we can represent it in charts by composing with chart maps. We say that F is smooth if all its representations are smooth according to Definition 9.70. Lemma 9.71 imediately implies

Lemma 9.80. *Let $F\colon M \to N$ be a diffeomorphism of manifolds with boundary. Then F maps interior points to interior points and boundary points to boundary points. Moreover, $F|_{\mathring{M}}\colon \mathring{M} \to \mathring{N}$ and $F|_{\partial M}\colon \partial M \to \partial N$ are diffeomorphisms.*

Vector bundles are defined exactly as in Section 8. A smooth section of a vector bundle $\pi\colon E \to M$ is defined again as a smooth map $\sigma\colon M \to E$ such that $\pi \circ \sigma = \mathrm{Id}_M$.

In particular we are interested in the tangent bundle TM of a smooth manifold with boundary M. It is defined again as the union of the tangent spaces T_pM, $p \in M$, the tangent space T_pM at p being defined exactly as in subsection 6.1. Notice that for $p \in \partial M$ the tangent space $T_p\partial M$ is a subspace of codimension one of T_pM. A vector field is defined as a section of TM and a k-form as a section of $\Lambda^k T^*M$. Notice that in an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ a vector field X and k-form ω are represented by collections X_α and ω_α of smooth vector fields and smooth k-forms on open subsets $\phi_\alpha(U_\alpha)$ of $\mathbb{H}^n$, $n = \dim M$. This means that the components of X_α and ω_α are restrictions to $\phi_\alpha(U_\alpha)$ of smooth functions defined on open neighborhoods of $\phi_\alpha(U_\alpha)$ in $\mathbb{R}^n$.

9.5.2. *Stokes theorem.* We are now ready to prove Stokes theorem. What we still have to discuss is only orientations. An orientation of a manifold with boundary is again defined as the choice of an equivalence class of volume forms.

If M is a manifold with boundary, a vector $n_p \in T_pM$, $p \in \partial M$, is called outward pointing if the last component of its representation in any chart is negative (by Lemma 9.71 this notion is chart independent). An orientation $[v]$ of M then induces an orientation $[w]$ of ∂M with the property that for any $p \in \partial M$ we have that $\iota^* \iota_{n_p} v_p = c\, w_p$ with $v \in [v]$, $w \in [w]$ and c a positive constant, where ι denotes the inclusion of ∂M into M (notice that it is enough to check this condition for a single p in each connected component of ∂M). This is called the induced orientation of ∂M. In a chart the induced orientation is obtained by contracting the representation of $[v]$ with the outward pointing vector field. That this procedure is consistent is again a consequence of Lemma 9.71.

Theorem 9.81 (Stokes Theorem)**.** *Let M be an oriented, Hausdorff, second-countable, n-dimensional manifold with boundary, $n > 0$, and ω a smooth $(n-1)$-form on M with compact support. Then*

$$\int_M \mathrm{d}\omega = \int_{\partial M} \omega,$$

where on ∂M we use the induced orientation.

Remark 9.82. In particular, if M has no boundary, then $\int_M \mathrm{d}\omega = 0$.

Proof. Let $\{(U_\alpha, \phi_\alpha)\}_{\alpha \in I}$ be an oriented atlas of M corresponding to the given orientation and let $\{\rho_j\}_{j\in J}$ be a partition of unity subordinate to it. First observe that

$$\mathrm{d}\omega = \mathrm{d}\left(\sum_{j\in J} \rho_j \omega\right) = \sum_{j\in J} \mathrm{d}(\rho_j\omega).$$

Next notice that $\operatorname{supp}(\rho_j\omega) \subset U_{\alpha_j}$. Thus, by the localization Lemma 9.62 we have

$$\int_M \mathrm{d}(\rho_j\omega) = \int_{\phi_{\alpha_j}(U_{\alpha_j})} \mathrm{d}(\rho_j\omega)_{\alpha_j}.$$

If $\phi_{\alpha_j}(U_{\alpha_j})$ is contained in the interior of $\mathbb{H}^n$, then we regard $\mathrm{d}(\rho_j\omega)_{\alpha_j}$ as a compactly supported top form on $\mathbb{R}^n$ (extending it by zero outside ot its support); hence, by Lemma 9.69, we get $\int_{\phi_{\alpha_j}(U_{\alpha_j})} \mathrm{d}(\rho_j\omega)_{\alpha_j} = 0$. Otherwise, we regard $\mathrm{d}(\rho_j\omega)_{\alpha_j}$ as a compactly supported top form on $\mathbb{H}^n$ (again extending it by zero outside of its support); hence, by Lemma 9.68, we get

$$\int_{\phi_{\alpha_j}(U_{\alpha_j})} \mathrm{d}(\rho_j\omega)_{\alpha_j} = \int_{\partial(\phi_{\alpha_j}(U_{\alpha_j}))} (\rho_j\omega)_{\alpha_j}.$$

Noticing that $\partial(\phi_{\alpha_j}(U_{\alpha_j})) = \phi_{\alpha_j}(\partial U_{\alpha_j})$ by definition and that both are oriented by outward pointing vectors, we get, again by the localization Lemma, that

$$\int_M \mathrm{d}(\rho_j\omega) = \int_{\partial M} \rho_j\omega.$$

Summing over j yields the result. □

Remark 9.83. Using equation (9.11), Remark 9.7 and the the definition of the flux right before Proposition 9.67, we can recover the original version of Stokes theorem: If U is an open subset of $\mathbb{R}^3$, X a vector field on U and $\Sigma \subset U$ an orientable surface with boundary $\partial\Sigma$, then

$$\int_{\partial\Sigma} X \cdot \mathrm{d}x = \int_\Sigma \mathrm{curl} X \cdot n\, v,$$

where v denotes the Riemannian density associated to the restriction to Σ of the Euclidean metric on $\mathbb{R}^3$ and n is the normal vector field to Σ that $\partial\Sigma$ encircles with the anticlockwise orientation (equivalently, the vector product of the tangent vector at $p \in \partial\Sigma$ with n_p is outward pointing).

Using Proposition 9.67, Remark 9.30 and the definition of the divergence of a vector field, we immediately[37] get a further consequence:

Theorem 9.84 (Gauss Theorem)**.** *Let M be a smooth manifold with boundary, g a Riemannian metric on M and X a vector field with compact support. Then*

$$\int_{\partial M} X \cdot n\, v_{g_{\partial M}} = \int_M \mathrm{div}_g X\, v_g,$$

where n is the outward pointing normal vector field.

Example 9.85. Consider the m-dimensional ball of radius R,

$$B_R^m = \left\{ (x^1, \dots, x^m) \in \mathbb{R}^m \;\middle|\; \sum_{i=1}^m (x^i)^2 \le R^2 \right\},$$

[37]The immediate proof following these steps actually requires orientability of M, but the theorem is true in general, as we will explain in Section 9.7 (see also, e.g., [4, Sect. 14]).

and its boundary, the $(m-1)$-dimensional sphere of radius R,

$$S_R^{m-1} = \left\{ (x^1, \dots, x^m) \in \mathbb{R}^m \;\middle|\; \sum_{i=1}^m (x^i)^2 = R^2 \right\},$$

We regard them both as endowed with the restriction of the Euclidean metric on $\mathbb{R}^m$. Let $E = \sum_{i=1}^m x^i \partial_i$ denote the Euler vector field. We have $\mathrm{div} E = m$. The outward pointing vector field on S_R^{m-1} is $\frac{x}{R}$. Hence $E \cdot n = R$. The Gauss theorem then yields

$$m \operatorname{Vol}(B_R^m) = R \operatorname{Vol}(S_R^{m-1}).$$

By a simple change of variables, we also have $\operatorname{Vol}(B_{\lambda R}^m) = \lambda^m \operatorname{Vol}(B_R^m)$ for all $\lambda > 0$. Hence, $R\frac{\partial}{\partial R} \operatorname{Vol}(B_R^m) = m \operatorname{Vol}(B_R^m)$. Thus, we get

$$\operatorname{Vol}(S_R^{m-1}) = \frac{\partial}{\partial R} \operatorname{Vol}(B_R^m).$$

If we regard again the sphere as the boundary of the ball and write ∂ for $\frac{\partial}{\partial R}$, we get the more suggestive equation

$$\operatorname{Vol}(\partial B_R^m) = \partial \operatorname{Vol}(B_R^m).$$

9.5.3. *The winding number.* We consider a simple application of Stokes theorem. Consider on $\mathbb{R}^2 \setminus \{0\}$ the 1-form

$$\omega := \frac{x\,\mathrm{d}y - y\,\mathrm{d}x}{x^2 + y^2}.$$

One can easily see, by an explicit computation, that ω is closed. One can also go to polar coordinates, $F\colon \mathbb{R}_{>0} \times S^1 \to \mathbb{R}^2 \setminus \{0\}$, and easily compute $F^*\omega = \mathrm{d}\theta$. Given a differentiable loop $\gamma\colon S^1 \to \mathbb{R}^2 \setminus \{0\}$ we define its winding number around 0 as

$$w(\gamma) := \frac{1}{2\pi} \int_\gamma \omega := \frac{1}{2\pi} \int_{S^1} \gamma^* \omega.$$

Note that this definition has the following immediate generalizations. First, it is enough to assume that γ be piecewise deifferentiable as the integral along γ can be defined as the sum of the integrals over the pieces where γ is differentiable. Second, we can define the winding number around any point $p \in \mathbb{R}^2$ for loops that miss that point just by translating the above expression.

Remark 9.86. Note that working in polar coordinates shows that $w(\gamma)$ is an integer that measures how many times the curve γ winds around 0 in the counterclockwise direction. In fact, fix a point $t_0 \in S^1$ and remove from $\mathbb{R}^2 \setminus \{0\}$ the half line that passes through 0 and $\gamma(t_0)$. In this complement the angle $\theta \in [0, 2\pi]$ is a coordinate. We split the image of γ into the portions that start and end on the removed line but do not cross it. The integral of ω along γ is the sum of the integrals on each of these portions, and each of these integrals is equal to 0, 2π or -2π. Note that one may take this as the definition of the winding number, show that it does not depend on the choice of the initial point t_0, and extend it to continuous loops γ.

We now want to show that the winding number is a "homotopy invariant." Namely, given two loops γ_0 and γ_1 as above, we say that a differentiable map $\Gamma\colon [0,1]\times S^1 \to \mathbb{R}^2\setminus\{0\}$ is a homotopy from γ_0 and γ_1 if, for all $t\in S^1$, $\Gamma(0,t)=\gamma_0(t)$ and $\Gamma(1,t)=\gamma_1(t)$. We also say, in this case, that γ_0 and γ_1 are homotopic.

Lemma 9.87. *If γ_0 and γ_1 are homotopic, then $w(\gamma_0)=w(\gamma_1)$.*

Proof. By Stokes theorem we have $0=\int_{[0,1]\times S^1}\mathrm{d}\Gamma^*\omega=\int_{\gamma_1}\omega-\int_{\gamma_0}\omega$. □

One may show that this result extends to the setting of continuous curves and continuous homotopies. For more details, see, e.g., [2].

9.6. **Singular homology.** Stokes theorem may be formulated on much more general objects than manifolds with boundary. A far reaching generalization are manifolds with corners, i.e., spaces locally modeled on open subsets of $\mathbb{R}^k\times\mathbb{R}^l_{\geq 0}$. We do not present the general theory, but focus on the very important example of simplices. The standard p-simplex is the closed subset[38]

$$\Delta^p := \left\{(x^1,\dots,x^p)\in\mathbb{R}^p \;\middle|\; \sum_{i=1}^p x^i\leq 1,\ x^i\geq 0\ \forall i\right\}$$

of $\mathbb{R}^p$. Notice that the interior of Δ^p is a p-dimensional manifold. The 0-simplex is just a point, the 1-simplex is an interval, the 2-simplex is a triangle, and the 3-simplex is a tetrahedron.

A smooth differential form on Δ^p is by definition the restriction to Δ^p of a smooth differential form defined on an open neighborhood of Δ^p in $\mathbb{R}^p$. We can integrate top forms on Δ^p just identifying them with densities by the standard orientation of $\mathbb{R}^p$.

Let now ω be a smooth $(p-1)$-form on Δ^p. The integral of $\mathrm{d}\omega$ on Δ^p can then be computed, by Fubini's theorem and by the fundamental

[38] The p-simplex is also defined by some authors as

$$\Delta^p := \left\{(x^0,\dots,x^p)\in\mathbb{R}^{p+1} \;\middle|\; \sum_{i=0}^p x^i = 1,\ x^i\geq 0\ \forall i\right\}.$$

theorem of analysis, as a sum of integrals of ω on the faces of Δ^p. Notice that each of these faces is a $(p-1)$-simplex (related to the standard one by a diffeomorphism; see below).

Explictly, we write $\omega = \sum_{j=1}^p \omega^j\,\mathrm{d}x^1 \wedge \dots \widehat{\mathrm{d}x^j} \wedge \dots \wedge \mathrm{d}x^p$. Then $\mathrm{d}\omega = \sum_{j=1}^p (-1)^{j+1}\,\partial_j\omega^j\,\mathrm{d}^p x$ and

$$\int_{\Delta^p} \mathrm{d}\omega = \sum_{j=1}^p (-1)^{j+1} \int_{\Delta^p} \partial_j \omega^j \, d^p x.$$

By Fubini's theorem, we can integrate the jth term first in the jth coordinate; by the fundamental theorem of analysis we then get

$$\int_{\Delta^p} \partial_j \omega^j \, dx^j = \omega^j|_{x^j = 1 - \sum_{\substack{i=1 \\ i\neq j}}^p x^i} - \omega^j|_{x^j=0}.$$

Hence

$$\int_{\Delta^p} \mathrm{d}\omega = \sum_{j=1}^p (-1)^{j+1} \int_{\Delta^p \cap \left\{\sum_{i=1}^p x^i = 1\right\}} \omega^j \, dx^1 \cdots \widehat{dx^j} \cdots dx^p + \tag{9.12}$$
$$+ \sum_{j=1}^p (-1)^j \int_{\Delta^p \cap \{x^j = 0\}} \omega^j \, dx^1 \cdots \widehat{dx^j} \cdots dx^p.$$

We may rewrite this formula in a more readable way if we regard the faces on which we integrate as images of $(p-1)$-simplices; namely, for $i = 0, \dots, p$, we define smooth[39] maps

$$k_i^{p-1} \colon \Delta^{p-1} \to \Delta^p,$$

by

$$k_0^{p-1}(a^1, \dots, a^{p-1}) = \left(1 - \sum_{i=1}^{p-1} a^i, a^1, \dots, a^{p-1}\right)$$

and

$$k_j^{p-1}(a^1, \dots, a^{p-1}) = (a^1, \dots, a^{j-1}, 0, a^j, \dots, a^{p-1}),$$

for $j > 0$.

The jth integral in the second line of (9.12) is just the integral on Δ^{p-1} of the pullback of ω by k_j^{p-1}. In fact,

$$(k_j^{p-1})^*\omega = (k_j^{p-1})^* \sum_{i=1}^p \omega^i \, \mathrm{d}x^1 \wedge \dots \widehat{\mathrm{d}x^i} \wedge \dots \wedge \mathrm{d}x^p =$$
$$= \omega^j(a^1, \dots, a^{j-1}, 0, a^j, \dots, a^{p-1})\, \mathrm{d}^{p-1}a.$$

[39]Again, smooth means that these maps are restrictions of smooth maps defined on open neighborhoods.

We then integrate by the standard orientation on Δ^{p-1} and rename variables: $x^i = a^i$ for $i < j$ and $x^i = a^{i+1}$ for $i > j$.

The jth integral in the first line is on the other hand the integral on Δ^{p-1} of the pullback of $(-1)^{j+1}\omega^j \,\mathrm{d}x^1 \wedge \dots \widehat{\mathrm{d}x^j} \wedge \dots \wedge \mathrm{d}x^p$ by k_0^{p-1}. In fact,

$$\begin{aligned}(k_0^{p-1})^*\omega^j \,\mathrm{d}x^1 \wedge \dots \widehat{\mathrm{d}x^j} \wedge \dots \wedge \mathrm{d}x^p = \\ = -\omega^j(1-\sum\nolimits_i a^i, a^1, \dots, a^{p-1}) \sum\nolimits_i \mathrm{d}a^i \wedge \mathrm{d}a^1 \wedge \dots \wedge \widehat{\mathrm{d}a^{j-1}} \wedge \dots \wedge \mathrm{d}a^{p-1} = \\ = (-1)^{j+1}\omega^j(1-\sum\nolimits_i a^i, a^1, \dots, a^{p-1})\,\mathrm{d}^{p-1}a.\end{aligned}$$

We then integrate by the standard orientation on Δ^{p-1} and rename variables: $x^i = a^i$ for $i < j$ and $x^i = a^{i+1}$ for $i > j$.

Summing up all contributions, we finally get the Stokes theorem for a simplex:

$$\int_{\Delta^p} \mathrm{d}\omega = \sum_{j=0}^{p} (-1)^j \int_{\Delta^{p-1}} (k_j^{p-1})^*\omega \tag{9.13}$$

where the term with $j = 0$ corresponds to the whole sum in the first line of (9.12) and each other term corresponds to a term in the second line.

Things become more interesting if we consider a smooth map

$$\sigma\colon \Delta^p \to M,$$

where M is a smooth manifold (again, we mean that σ is the restriction of a smooth map defined on an open neighborhood of Δ^p). If α is a smooth p-form on M, one defines

$$\int_\sigma \alpha := \int_{\Delta^p} \sigma^*\alpha.$$

If we define $\sigma^j := \sigma \circ k_j^{p-1}\colon \Delta^{p-1} \to M$, then (9.13) implies

$$\int_\sigma \mathrm{d}\omega = \sum_{j=0}^{p} (-1)^j \int_{\sigma^j} \omega.$$

This equation gets an even better form if we introduce the notion of chains.

Definition 9.88. A p-chain with real coefficients in a smooth manifold M is a finite linear combination $\sum_k a_k\sigma_k$, $a_k \in \mathbb{R}$ for all k, of smooth maps $\sigma_k\colon \Delta^p \to M$. If α is a smooth p-form on M, one defines

$$\int_{\sum_k a_k\sigma_k} \omega := \sum_k a_k \int_{\sigma_k} \omega. \tag{9.14}$$

We then have the Stokes theorem for chains

$$\boxed{\int_\sigma \mathrm{d}\omega = \int_{\partial\sigma} \omega}$$

where

$$\partial\sigma := \sum_{j=0}^{p} (-1)^j \sigma^j. \tag{9.15}$$

We let $\Omega_p(M,\mathbb{R})$ denote the vector space of p-chains in M with real coefficients and extend ∂ to it by linearity. We then have, with the terminology of subsection 9.3, that ∂ is an endomorphism of degree -1 of the graded vector space $\Omega_\bullet(M,\mathbb{R})$. By a simple calculation, one can actually verify that ∂ is a boundary operator, see Definition 9.15; i.e.,

$$\partial \circ \partial = 0.$$

For $\sigma \in \Omega_p(M,\mathbb{R})$ and $\omega \in \Omega^p(M)$ we define

$$(\sigma,\omega) := \int_\sigma \omega.$$

This is a bilinear map $\Omega_p(M,\mathbb{R}) \times \Omega^p(M) \to \mathbb{R}$. Stokes theorem for chains now reads

$$(\sigma, \mathrm{d}\omega) = (\partial\sigma, \omega).$$

A chain in the kernel of ∂ is called a cycle, a chain in the image of ∂ is called a boundary. One defines $H_p(M,\mathbb{R})$ as the quotient of p-cycles by p-boundaries. The graded vector space $H_\bullet(M,\mathbb{R})$ is called the singular homology with real coefficients of M.

Notice that a smooth map $F\colon M \to N$ induces a graded linear map $F_*\colon \Omega_\bullet(M,\mathbb{R}) \to \Omega_\bullet(N,\mathbb{R})$, $\sigma \mapsto F \circ \sigma$. We clearly have $\partial F_* = F_*\partial$, which implies that F_* descends to a graded linear map

$$F_*\colon H_\bullet(M,\mathbb{R}) \to H_\bullet(N,\mathbb{R}).$$

If F is a diffeomorphism, then F_* is an isomorphism in singular homology. This shows that the singular homology is also an invariant of manifolds.[40] Another important remark is that

$$(F_*\sigma, \omega) = (\sigma, F^*\omega)$$

for all $\sigma \in \Omega_\bullet(M,\mathbb{R})$ and all $\omega \in \Omega^\bullet(N)$.

[40]Notice that one can also define $\mathcal{C}^l$-chains, for any l, just by requiring the maps to be $\mathcal{C}^l$. One can prove that the $\mathcal{C}^l$-singular homologies are all isomorphic to each other. In particular, one can work with continuos maps. This shows that homeomorphic manifolds have the same singular homology. Also observe that to define continuous maps we just have to assume that M is a topological space. This means that one can define singular homology for topological spaces as well.

Finally, the Stokes theorem for chains implies that the above bilinear map pairing chains to differential forms descends to a bilinear map

$$H_p(M,\mathbb{R}) \times H^p(M) \to \mathbb{R}.$$

The most important result in this context is the de Rham theorem that asserts that this induced pairing is nondegenerate. Notice that this implies that $H^p(M)$ is isomorphic to $H_p(M,\mathbb{R})^*$ for all p.

Remark 9.89. One can define chains with coefficients in any unital ring R: A p-chain σ with coefficients in R is a finite linear combination $\sum_k a_k\sigma_k$, $a_k \in R$ for all k, of maps $\sigma_k\colon \Delta^p \to M$. One denotes by $\Omega_p(M,R)$ the R-module of p-chains in M with coefficients in R and defines ∂ by formula (9.15). One still has that ∂ is a boundary operator on the graded R-module $\Omega_\bullet(M,R)$ and one can define the singular homology $H_\bullet(M,R)$ of M with coefficients in R. Again a map $F\colon M \to N$ induces an R-linear map in homology and a homeomorphism induces an isomorphism. Notice that the pairing (9.14) with differential forms is not defined if we do not specify a homomorphism $R \to \mathbb{R}$ to make sense of the right hand side. In the special case $R = \mathbb{Z}$, one simply writes $\Omega_p(M)$ and $H_p(M)$ instead of $\Omega_p(M,\mathbb{Z})$ and $H_p(M,\mathbb{Z})$. The latter is usually called that pth homology group of M. Using the inclusion homomorphism $\mathbb{Z} \to \mathbb{R}$ one can pair chains with integral coefficients with differential forms. Notice however that, in general, the induced pairing $H_p(M) \times H^p(M) \to \mathbb{R}$ is degenerate.

9.7. **The nonorientable case.** Differential forms, unlike densities, require an orientation to be integrated; they are, however, more flexible as they can be restricted and integrated on (oriented) submanifolds and form a complex. This in particular leads to Stokes theorem. In this section we want to see what can be saved of the theory of differential forms without orientation. We will see that there is a variant of Stokes theorem and that, in particular, Gauss theorem holds also on nonorientable manifolds.

Recall that a top form, in local coordinates, transforms with the determinant of the Jacobian of the transformation maps, whereas a density transforms with its absolute value. To keep track of this twist we introduce the orientation bundle $\mathrm{or}(M)$ of a manifold M: this is the line bundle whose transition functions are given by the signs of the Jacobians. Namely, if we fix an atlas $\{(U_\alpha,\phi_\alpha)\}_{\alpha\in I}$, we set $A_{\alpha\beta}(q) = \operatorname{sgn}\det \mathrm{d}_{\phi_\alpha(q)}\phi_{\alpha\beta}$, in the notations of Section 8.1.4. Note that the transition functions of the orientation bundle are locally constant. We now come to the central

Definition 9.90. A twisted differential form on M is a section of $\Lambda^\bullet T^*M \otimes \mathrm{or}(M)$.

Note that a twisted top form is the same as a density (they both transform with the absolute value of the determinant of the Jacobian).

Lemma 9.91. *Let M be Hausdorff and second countable. Then* $\mathrm{or}(M)$ *is trivial if and only if M is orientable.*

Proof. If M is orientable, then $\mathrm{or}(M)$ is trivial by Proposition 8.28.

Vice versa, if $\mathrm{or}(M)$ is trivial, we have global sections. A choice of global section allows identifying top forms and twisted top forms, i.e., densities. If M is Hausdorff and second countable, we may pick a positive density which in turns defines a volume form. □

Notice that choosing a global section of $\mathrm{or}(M)$, in the orientable case, yields an identification between differential forms and twisted differential forms.

On twisted differential forms we may define the twisted de Rham differential just by letting it operate on the differential form factor. This makes sense, since the transition functions of the orientation bundle are locally constant. More explictly, a twisted differential form ω is represented in the atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ by differential forms ω_α on $\phi_\alpha(U_\alpha)$ transforming as

$$\omega_\alpha = \operatorname{sgn} \mathrm{d}\phi_{\alpha\beta}\ \phi_{\alpha\beta}^*\omega_\beta.$$

Since $\operatorname{sgn} \mathrm{d}\phi_{\alpha\beta}$ is locally constant, we have

$$\mathrm{d}\omega_\alpha = \operatorname{sgn} \mathrm{d}\phi_{\alpha\beta}\ \phi_{\alpha\beta}^*\mathrm{d}\omega_\beta,$$

which shows that the collection of the $\mathrm{d}\omega_\alpha$s defines a twisted differential form. One can define the twisted de Rham cohomology, which coincides with the usual one if M orientable. Also note that we can analogously define contractions and Lie derivatives by vector fields and the whole Cartan's calculus extends to the twisted setting. In particular, if we regard a density σ as a twisted top form, we have $\mathsf{L}_X\sigma = \mathrm{d}\iota_X\sigma$, a crucial fact for the proof of Gauss theorem.

Twisted top forms, being the same as densities, can be integrated on the whole manifold (if the integral converges). To define integration on (appropriate) submanifolds, we have to understand how to restrict twisted forms. Note that, if $\iota\colon S \to M$ is a submanifold and u a section of $\mathrm{or}(M)$, the restriction of u to S is a section of the pullback bundle $\iota^*\mathrm{or}(M)$.

Definition 9.92. A submanifold $\iota\colon S \to M$ is called co-orientable if $\iota^*\mathrm{or}(M)$ is isomorphic to $\mathrm{or}(S)$. A choice of isomorphism is called a co-orientation and S with this choice is called co-oriented.

Note then that, if ω is a twisted differential form on M, the restriction of ω to S, $\iota^*\omega$, is a twisted differential form on S. In particular, if the form degree of ω is equal to the dimension of S, then $\iota^*\omega$ is a density on S and can be integrated. We have thus defined a pairing between twisted k-form and k-dimensional co-oriented submanifolds.

In a local adapted atlas, we have tangential coordinates $x^1, \dots, x^k$ and transversal coordinates $y^1, \dots, y^r$, with $k = \dim S$ and $r = \dim M - k$. We assume that S is determined by setting the ys to zero. When we pass to a different adapted chart, we write $\phi_{\alpha\beta} = (\psi_{\alpha\beta}, \chi_{\alpha\beta})$, where $\psi_{\alpha\beta}$ consists of the first k components and $\chi_{\alpha\beta}$ of the last r components. On S we have $\chi_{\alpha\beta} = 0$. As a consequence, for each point (x, y) on $\phi_\alpha(S \cap U_\alpha \cap U_\beta)$, the differential of $\phi_{\alpha\beta}$ is the block matrix $\begin{pmatrix} \mathrm{d}_x\psi & \mathrm{d}_y\psi \\ 0 & \mathrm{d}_y\chi \end{pmatrix}$. It follows that the transition functions of $\iota^*\mathrm{or}(M)$ are $\operatorname{sgn} \mathrm{d}_x\psi_{\alpha\beta} \operatorname{sgn} \mathrm{d}_y\chi_{\alpha\beta}$. Since the $\psi_{\alpha\beta}$s are the transition maps for S in the induced atlas, we see that S is co-orientable if and only if the line bundle on S with transition functions $\operatorname{sgn} \mathrm{d}_y\chi_{\alpha\beta}$, at $y = 0$, called the co-orientation bundle, is trivial.

This immediately implies that a submanifold of an orientable manifold is co-orientable if and only if it is orientable.

It also shows that in the definition of the flow of a vector field through a hypersurface S, see Section 9.4.5, what is needed is actually just a co-orientation of S: this is equivalent to choosing one normal vector field. The proof of Proposition 9.67 can easily be adapted to prove its following more general version:

Proposition 9.93. *Let S be a co-oriented hypersurface of a Riemannian manifold (M, g). Then, for every $X \in \mathfrak{X}(M)$, we have*

$$\int_S X \cdot n\, v_{g_S} = \int_S \iota_X v_g,$$

where the co-orientation of S is used on the left hand side to define n and on the right hand side to define the restriction of the twisted form to S.

Let us now come to the case of a manifold with boundary M. By inspection in the proof of Lemma 9.71, we see that ∂M is always co-orientable. We will always assume its standard co-orientation, the one given by taking the last component of an outward pointing vector as a section of the co-orientation bundle. Repeating the same steps as in the proof of the non-twisted Stokes theorem, we get its twisted version:

Theorem 9.94. *Let M be a Hausdorff, second-countable, n-dimensional manifold with boundary, $n > 0$, and ω a smooth twisted $(n-1)$-form on M with compact support. Then*

$$\int_M \mathrm{d}\omega = \int_{\partial M} \omega.$$

From Proposition 9.93, Theorem 9.94 and the Cartan formula for the Lie derivative of a twisted top form, we get the proof to Gauss theorem, see page 129, without having to assume orinetability of the manifold.

Example 9.95. Consider the Möbius band M again, presented as $[0,1] \times \mathbb{R}$ with $(0,y)$ identified with $(1,-y)$ for all $y \in \mathbb{R}$. The curve $\gamma_0\colon [0,1] \to M$, $t \mapsto (t,0)$ is not co-orientable. On the other hand, for each $h \neq 0$, we have the co-orientable curve $\gamma_h\colon [0,1] \to M$ given by

$$\gamma(t) = \begin{cases} (2t, h) & t \in [0, 1/2] \\ (2t-1, -h) & t \in [1/2, 1] \end{cases}$$

Take $h > 0$ and consider the region M_h enclosed by the curve (the one containing γ_0). This is a nonorientable manifold with boundary. As an example for the Gauss theorem, consider the Euclidean metric $dx^2 + dy^2$ and the vector field $X = y\frac{\partial}{\partial y}$. The flow of X through ∂M_h is $2h$. On the other hand, the divergence of y is 1, and the Euclidean area of M_h is also $2h$.

9.8. **Digression: Symplectic manifolds.** A symplectic form ω on a manifold M is by definition a closed, nondegenerate 2-form. Nondegenerate means that for all $p \in M$ we have

$$\omega_p(v, w) = 0\, \forall v \in T_pM \iff w = 0.$$

Equivalently, the determinant of the matrix representing ω in a chart is nonzero at every point. A pair (M, ω), where M is a smooth manifold and ω a symplectic form on M, is called a symplectic manifold.

Remark 9.96. Notice that a 2-form ω on M establishes a morphism $\omega^\sharp\colon TM \to T^*M$ of vector bundles by sending $v \in T_pM$ to the linear form $w \mapsto \omega_p(v, w)$. This morphism is an isomorphism if and only if ω is nondegenerate. Hence, if ω is symplectic, to every function H we can associated a unique vector field X_H such that

$$\iota_{X_H}\omega = -\mathrm{d}H.$$

The vector field X_H is called the Hamiltonian vector field of H (the minus sign on the r.h.s. is purely conventional and not used by all authors). On the other hand, a vector field X is called Hamiltonian if it is the Hamiltonian vector field of a function (which is uniquely defined up to the addition of a locally constant function); any of these function is called a Hamiltonian function for X. Since ω is closed, by Cartan's formula we get

$$\mathsf{L}_{X_H}\omega = 0.$$

Equivalently, ω is invariant under the flow of X_H, which is called the Hamiltonian flow of H.

Example 9.97. Symplectic geometry arises in mechanics. Consider, e.g., Newton's equation for one particle in $\mathbb{R}^3$,

$$m\ddot{x}^i = F^i(x, \dot{x}, t),$$

where F, the force, is a given function and m, the mass, is a given positive number. As usual, one replaces this system of three second-order ODEs by an equivalent system of six first-order ODEs by introducing the momentum $p_i = m\dot{x}^i$:

$$\begin{aligned} \dot{p}_i &= F^i(x, p/m, t) \\ \dot{x}^i &= p_i/m. \end{aligned}$$

A system is called conservative if F does not depend on time and velocities and is minus the gradient of a function U, the potential, of the coordinates: $F^i = -\partial_i U$. In a conservative system, Newton's equations in the p, x variables turn out to be the ODE of the Hamiltonian vector field of the function $H = \sum_i \frac{(p_i)^2}{2m} + U$ with respect to the symplectic form $\omega = \sum_i \mathrm{d}p_i \mathrm{d}q^i$.

More generally,

Example 9.98. Let N be an open subset of $\mathbb{R}^{2n}$ with coordinates $q^1, \dots, q^n$, $p_1, \dots, p_n$. Then

$$\omega = \sum_{i=1}^{n} \mathrm{d}p_i \wedge \mathrm{d}q^i \tag{9.16}$$

is a symplectic form on N.

Example 9.99. The cotangent bundle T^*M of any manifold possesses a natural symplectic form defined as follows. First one defines the Liouville 1-form θ (a.k.a. the Poincaré 1-form or the tautological 1-form). The easiest way to define it is by choosing an atlas $\{(U_\alpha, \phi_\alpha)\}_{\alpha\in I}$ for M and the corresponding trivializing atlas for TM. We denote by q^i_α the coordinates in the charts U_α and by p_i^α the coordinates on the fibers. They satisfy the transition rules

$$q_\beta^i = \phi_{\alpha\beta}^i(q_\alpha),$$
$$p_i^\alpha = \sum_{j=1}^n \frac{\partial \phi_{\alpha\beta}^j}{\partial q_\alpha^i}(q_\alpha)\, p_j^\beta.$$

This implies that the collection of 1-forms θ_α,

$$\theta_\alpha := \sum_{i=1}^n p_i^\alpha \,\mathrm{d}q_\alpha^i,$$

defines a 1-form θ on M since $\theta_\alpha = \widetilde{\phi}_{\alpha\beta}^*\theta_\beta$ for all α, β. Its differential $\omega = \mathrm{d}\theta$ is nondegenerate, since in charts it reads $\sum_{i=1}^n \mathrm{d}p_i^\alpha\,\mathrm{d}q_\alpha^i$, and hence is a symplectic form on T^*M. There is also a coordinate-independent definition of θ. Namely, denote by (q,p), $q \in M$ and $p \in (T_qM)^*$, the points in T^*M and let $\pi\colon T^*M \to M$, $\pi(q,p) = q$ be the projection map. For $v \in T_{(q,p)}T^*M$, define $\theta_{(q,p)}v := p(\mathrm{d}_{(q,p)}\pi\, v)$.

Remark 9.100. Darboux's Theorem, which we will prove later (Theorem 9.107), asserts that every symplectic manifold possesses an atlas such that the symplectic form in each chart is as in (9.16).

From now on, let (N,ω) be a symplectic manifold. We want to draw a few consequences. The first remark follows from linear algebra: a vector space admits a skew-symmetric nondegenerate bilinear form only if it has even dimension. This implies that $\dim N = 2m$, for some integer m. The nondegeneracy of ω implies that the top form $\rho = \omega^m/(m!)$, called the Liouville volume form, is nowhere vanishing. This in turn implies that a symplectic manifold is always orientable (and actually oriented by ρ). The integral $\int_M \rho$ is called the symplectic volume (we put on M the orientation $[\rho]$, so this number is strictly positive, possibly infinite.) This has an interesting corollary: If N is compact, then ω is not exact. To prove this assume on the contrary that ω is exact. This implies that ρ is also exact and by Stokes theorem that $\int_M \rho = 0$; but this is impossible.

Observe that, since ω is invariant under the flow of a Hamiltonian vector field, then so is ρ. This means that the ρ-divergence of a Hamiltonian vector field is always zero.

If H and F are two functions on N, then one can easily see that

$$X_H(F) = -X_F(H) \tag{9.17}$$

as both are equal to $-\iota_{X_H}\iota_{X_F}\omega$. This has two important consequences. The first is Noether's theorem. We need first the

Definition 9.101. A Hamiltonian system is a pair $((N, \omega), H)$ where (N, ω) is a symplectic manifold and H is a function on N. A constant of motion for the Hamiltonian system $((N, \omega), H)$ is a function that is constant on the orbits of X_H. An infinitesimal symmetry for the Hamiltonian system $((N, \omega), H)$ is a Hamiltonian vector field X on N such that $X(H) = 0$.

Theorem 9.102 (Noether's Theorem). *A Hamiltonian vector field is a symmetry for the Hamiltonian system $((N, \omega), H)$ if and only if any of its Hamiltonian functions is a constant of motion.*

Proof. Let F be a Hamiltonian function for the vector field at hand, which we denote by X_F. Being a symmetry means $X_F(H) = 0$. On the other hand, F is a constant of motion if and only if $X_H(F) = 0$. The Theorem then follows from (9.17). □

The second consequence of (9.17) is that the bracket

$$\{H, F\} := X_H(F),$$

called the Poisson bracket on (N, ω), is skew-symmetric.

Lemma 9.103. *The Poisson bracket $\{\ ,\ \}$ is a Lie bracket on $C^\infty(M)$ that in addition satisfies the Leibniz rule $\{f, gh\} = \{f, g\}h + g\{f, h\}$ for all $f, g, h, \in C^\infty(M)$. Moreover, the map $f \mapsto X_f$ is a Lie algebra morphism from $C^\infty(M)$ to $\mathfrak{X}(M)$.*

Proof. The Leibniz rule follows immediately from the Leibniz rule for vector fields and from the definition of the Poisson bracket. Differentiating the defining relation $\{f, g\} = \iota_{X_f}\mathrm{d}g$ and using Cartan calculus, we get, $\mathrm{d}\{f, g\} = \mathsf{L}_{X_f}\mathrm{d}g$. From the definition of the Hamiltonian vector field for g, and again using Cartan calculus, we get $\mathrm{d}\{f, g\} = -\mathsf{L}_{X_f}\iota_{X_g}\omega = [\iota_{X_g}, \mathsf{L}_{X_f}]\omega = \iota_{[X_g, X_f]}\omega$, where we have also used that ω is X_f-invariant. This shows that the Hamiltonian vector field of $\{f, g\}$ is $[X_f, X_g]$:

$$X_{\{f,g\}} = [X_f, X_g].$$

We then have

$$\begin{aligned}\{f, \{g, h\}\} = -\{\{g, h\}, f\} = -[X_g, X_h](f) = \\ = X_h(X_g f) - X_g(X_h f) = \{h, \{g, f\}\} - \{g, \{h, f\}\},\end{aligned}$$

which is the Jacobi identity. The previous identity also shows that $f \mapsto X_f$ is a Lie morphism. □

We can use the Poisson bracket to rephrase Noether's theorem: A function F is a constant of motion for H if and only $\{F, H\} = 0$. Skew-symmetry immediately implies that F is a constant of motion for H if and only if H is a constant of motion for F. In addition we have the

Proposition 9.104. *If F and G are constants of motion for H, then so are FG and $\{F, G\}$. Hence the constants of motion for a given Hamiltonian system form a Poisson subalgebra.*

Proof. By Noether's theorem we have $\{H, F\} = \{H, G\} = 0$. The Leibniz and the Jacobi identities imply $\{H, FG\} = 0 = \{H, \{F, G\}\}$. □

Remark 9.105. Note that, more generally, any Lie bracket on a commutative algebra that also satisfies the Leibniz rule is called a Poisson bracket.

9.8.1. *Normal form.* Symplectic manifold locally look all alike. This is the content of Darboux's theorem. We start with a very useful technical Lemma.

Lemma 9.106 (Moser's trick)**.** *Let ω_0 and ω_1 be symplectic forms on an open subset U of $\mathbb{R}^n$ that coincide at some point q. Then there are open neighborhoods V_0 and V_1 of q and a diffeomorphism $\phi\colon V_0 \to V_1$ such that $\phi(q) = q$ and*

$$\omega_0|_{V_0} = \phi^*\omega_1|_{V_1}.$$

Proof. Consider the convex combination $\omega_t := (1-t)\omega_0 + t\omega_1$, $t \in [0, 1]$. Observe that $\frac{\partial}{\partial t}\omega_t = \omega_1 - \omega_0$. Let U' be an open ball around q inside U. Since ω_0 and ω_1 are closed, by the Poincaré Lemma we then have a 1-form θ on U' such that there $\frac{\partial}{\partial t}\omega_t = \mathrm{d}\theta$. Note that θ is defined up to the differential of a function, and we may always choose it such that θ vanishes at q.[41]

Next we choose a neighborhood U'' of q in U' where ω_t is nondegenerate for all $t \in [0, 1]$. To see that this is possible, consider the map $[0, 1] \times U' \to \mathbb{R}$ that assigns to (t, x) the determinant of ω_t at x. Let C be the preimage of 0, which is closed since this map is continuous. Its complement contains $[0, 1] \times \{q\}$ (since at q we have, for all t, $\omega_t = \omega_0 = \omega_1$ which is nondegenerate), so in particular it contains an open subset of the form $[0, 1] \times U''$.

On U'' then we have, for each t, a unique vector field X_t such that $\iota_{X_t}\omega_t = \mathrm{d}\theta$, which implies $\mathsf{L}_{X_t}\omega_t = 0$. Note that by the assumption on θ also X_t vanishes at q. Finally we denote by $\Phi_t(y) \in U''$ the solution to

[41] Write $\theta = \sum_i \theta_i \mathrm{d}x^i$. Set $f(x) = -\sum_i \theta_i(q) x^i$. Then $\theta + \mathrm{d}f$ vanishes at q.

the ODE $\dot{x} = X_t(x)$ with initial condition y. As $\Phi_t(q) = q$ for all q, we may find a neighborhood U''' of q in U'' such that $\Phi_t(y)$ is defined for all $y \in U'''$ and for all $t \in [0,1]$. One can then show that $\omega_t = (\Phi_t^*)^{-1}\omega_0$ for all t. In fact, define $\widetilde{\omega}(t) := \Phi_t^*\omega_t$. We have

$$\frac{\partial}{\partial t}\widetilde{\omega}(t) = \lim_{\epsilon\to 0}\frac{\Phi_{t+\epsilon}^*\omega_{t+\epsilon} - \Phi_t^*\omega_t}{\epsilon}.$$

The main remark is now that $\Phi_{t+\epsilon}(y) = \Phi_\epsilon^{X_t}(\Phi_t(y)) + O(\epsilon^2)$. Hence

$$\frac{\partial}{\partial t}\widetilde{\omega}(t) = \Phi_t^*\lim_{\epsilon\to 0}\frac{(\Phi_\epsilon^{X_t})^*\omega_t - \omega_t}{\epsilon} = \Phi_t^* L_{X_t}\omega_t = 0.$$

Since $\widetilde{\omega}(0) = \omega_0$, we get that $\widetilde{\omega}(t) = \omega_0$ for all t. To complete the proof we set $V_0 = U'''$, $V_1 = \Phi_1(U''')$ and $\phi = \Phi_1$. □

One application of Moser's trick is

Theorem 9.107 (Darboux's Theorem)**.** *Let* (M,ω) *be a* $2n$*-dimensional symplectic manifold. Then every point* m *of* M *is contained in a chart* (V,τ) *such that* $\tau_*\omega$ *has the form of equation* (9.16)*.*

Proof. Let (W,ψ) be a chart containing m. The representation of ω in this chart at the point $\psi(m)$ is a nondegenerate skew-symmetric matrix which can hence be put in the form $\left(\begin{smallmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0\end{smallmatrix}\right)$ by a linear transformation A. We set $\psi' := A\circ\psi$, $U = \psi'(W)$, $\omega_1 = \psi'_*\omega|_W$ and ω_0 the restriction to U of the 2-form of equation (9.16). By Moser's trick, with the same notations, we conclude the proof by setting $V = (\psi')^{-1}(V_1)$ and $\tau = \phi^{-1}\circ\psi'$. □

9.8.2. *The degenerate case.* In the study of Hamiltonian systems, the nondegeneracy of the symplectic form is used to guarantee existence and uniqueness of Hamiltonian vector fields. Most of the other properties, however, only rely on the fact that the symplectic form is closed. We highlight in this section what survives if we drop the nondegeneracy condition. This is relevant to study "subsystems" defined by submanifolds of a symplectic manifold, as nondegeneracy is in general not preserved by restriction.

Definition 9.108. Let ω be a closed 2-form on a manifold M. We say that a vector field X is vertical if $\iota_X\omega = 0$. A function f is called invariant if $X(f) = 0$ for every vertical vector field X. A function f is called Hamiltonian if there is a vector field X_f such that $\iota_{X_f}\omega = -\mathrm{d}f$; such a vector field is called a Hamiltonian vector field for f.

Note that any two Hamiltonian vector fields for the same Hamiltonian function differ by a vertical vector field. Thus, if f is Hamiltonian and g invariant, we may define the action of f on g by

$$f\{g\} := X_f(g).$$

Notice that we have the Leibniz rule $f\{gh\} = f\{g\}h + gf\{h\}$ for every Hamiltonian function f and all invariant functions g and h.

Also note that a Hamiltonian function is automatically invariant, so in particular we have an action of Hamiltonian functions on Hamiltonian functions that we denote by a bracket: If f and g are Hamiltonian, we set

$$\{f, g\} := f\{g\}.$$

Note that the product fg of two Hamiltonian functions is also Hamiltonian (e.g., with $X_{fg} = fX_g + gX_f$). By inspecting the proof of Lemma 9.103, we see that also $\{f, g\}$ is Hamiltonian and that $\{\ ,\ \}$ is a Poisson bracket on the algebra of Hamiltonian functions.

Note that the Lie bracket of two vertical vector fields X and Y is also vertical: in fact,

$$\iota_{[X,Y]}\omega = [\iota_X, \mathsf{L}_Y]\omega = \iota_X \mathrm{d}\iota_Y\omega - \mathsf{L}_Y\iota_X\omega = 0.$$

(More generally, this computation shows that the Lie bracket of a vertical vector field X with a vector field Y that preserves the symplectic form, e.g., a Hamiltonian vector field, is vertical.)

This points at some form of involutivity. To make this more precise, let us introduce the kernel Δ of the bundle map $\omega^\sharp$ introduced in Remark 9.96. In other words, the kernel Δ_q at $q \in M$ consists of all tangent vectors v at q such that $\omega_q(v, w) = 0$ for all $w \in T_qM$. A vector field X is then vertical if and only if $X_q \in \Delta_q\ \forall q \in M$. If Δ is a regular distribution, it is then involutive by the above formula.

Definition 9.109. A 2-form ω on M is called presymplectic if it has constant rank, i.e., if $\dim \omega^\sharp(T_qM)$ is the same for all $q \in M$. A manifold endowed with a presymplectic form is called a presymplectic manifold.

Lemma 9.110. *The kernel of a presymplectic form is an involutive distribution.*

Proof. By the previous discussion, what is left to show is just the smoothness of Δ or, equivalently, that Δ is a subbundle of TM. This follows from the implicit function theorem if we regard Δ as the preimage of M under $\omega^\sharp$. We may see this in charts. There $\omega^\sharp_\alpha$ yields a map $u_\alpha\colon \phi_\alpha(U_\alpha) \times \mathbb{R}^n \to \mathbb{R}^n$, $(x, v) \mapsto (\omega^\sharp_\alpha)_x(v)$, and we have $\Delta_\alpha = u_\alpha^{-1}(0)$ which is a submanifold, as u_α has constant rank by assumption. □

By Frobenius' theorem Δ is then also integrable. If the leaf space N admits a smooth structure for which the canonical projection $\pi\colon M \to N$ is smooth, we then see that the vertical vector fields for the presymplectic form ω are the same as the vertical vector fields for the projection π. Since ω is closed and invariant, by Corollary 9.36 it is basic. Let $\underline{\omega}$ be the uniquely defined 2-form on N with $\omega = \pi^*\underline{\omega}$. Since π^* is injective, $\underline{\omega}$ is closed. It is also nondegenerate, since we have modded out precisely by the kernel. The symplectic manifold $(N, \underline{\omega})$ is called the symplectic reduction of (M, ω).

Next let X be a symplectic vector field, i.e., $\mathsf{L}_X\omega = 0$. Then, for every vertical vector field Y and by Cartan's calculus we get $\iota_{[X,Y]}\omega = -[\mathsf{L}_X, \iota_Y]\omega = 0$. This means that X is projectable. It also follows that $\phi(X)$ is symplectic. If X is Hamiltonian, i.e., $\iota_X\omega = -\mathrm{d}f$, then f is invariant. If we write $f = \pi^*\underline{f}$, then we also get $\iota_{\phi(X)}\underline{\omega} = -\mathrm{d}\underline{f}$.

Finally, suppose we have an invariant function $f = \pi^*\underline{f}$. We then have $\mathrm{d}f = \pi^*\mathrm{d}\underline{f} = -\pi^*(\iota_{X_{\underline{f}}}\underline{\omega})$, where $X_{\underline{f}}$ is the uniquely defined Hamiltonian vector field of $\underline{f}$. Let now X be a projectable vector field with $\phi(X) = X_{\underline{f}}$. We then have $\mathrm{d}f = -\iota_X\omega$, so f is Hamiltonian.

This is the lucky situation. However, even if the leaf space is not smooth, a presymplectic manifold still has nice features:

Proposition 9.111. *On a presymplectic manifold every invariant function is Hamiltonian.*

Proof. By Frobenius theorem, in a chart image the kernel distribution is spanned by the vector fields $\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^k}$. This implies that the presymplectic form reads $\frac{1}{2}\sum_{i,j>k} \omega_{ij}\mathrm{d}x^i \wedge \mathrm{d}x^j$, where at each point the coefficients ω_{ij} are the entries of a nondegenerate skew-symmetric $(n-k)\times(n-k)$-matrix. A function g is on the other hand invariant if it does not depend of the first k coordinates. As a Hamiltonian vector field X for g we may take a solution to the equation $\sum_{i>k} X^i\omega_{ij} = \frac{\partial g}{\partial x^j}$, $j > k$, which exists since $(\omega_{ij})_{i,j>k}$ is nondegenerate. □

10. Lie groups

Definition 10.1. A Lie group is a smooth Hausdorff manifold G with a group structure such that the multiplication

$$\begin{aligned} G \times G &\to G \\ (g,h) &\mapsto gh \end{aligned}$$

and the inversion

$$\begin{aligned} G &\to G \\ g &\mapsto g^{-1} \end{aligned}$$

are smooth maps.[42]

Example 10.2 (The general linear group)**.** Consider the group

$$\mathrm{GL}(n) := \{A \in \mathrm{Mat}(n \times n, \mathbb{R}) \mid \det A \neq 0\}$$

of invertible $n \times n$-matrices with real coefficients. The vector space $\mathrm{Mat}(n\times n, \mathbb{R})$ can be identified with $\mathbb{R}^{n^2}$ with each entry being a coordinate and gets as such a standard manifold structure. The group $\mathrm{GL}(n)$ inherits a manifold structure being an open subset of $\mathrm{Mat}(n \times n, \mathbb{R})$. Since each entry in the product of two matrices is a polynomial in the entry of its factors, matrix multiplication is a smooth map. By Cramer's rule, each entry of the inverse of a matrix is the ratio of the corresponding adjugate matrix over the determinant of the given matrix; as such it is a rational function and, therefore, a smooth function on the complement of matrices with determinant zero. Hence, $\mathrm{GL}(n)$ is a Lie group. Similarly, the group $\mathrm{GL}(n, \mathbb{C})$ is an open subset of $\mathrm{Mat}(n \times n, \mathbb{C})$, which can be identified with $\mathbb{R}^{2n^2}$ with standard manifold structure, and as such it is a Lie group.

Example 10.3 (Matrix Lie Groups)**.** The classical matrix groups $\mathrm{SL}(n)$, $\mathrm{O}(n)$, $\mathrm{SO}(n)$, $\mathrm{SL}(n, \mathbb{C})$, $\mathrm{U}(n)$ and $\mathrm{SU}(n)$ arise as subsets of $\mathrm{Mat}(n \times n, \mathbb{R})$ or $\mathrm{Mat}(n \times n, \mathbb{C})$ defined by constraints satisfying the conditions of the implicit function theorem:

$$\begin{aligned}
\mathrm{SL}(n) &= \{A \in \mathrm{Mat}(n \times n, \mathbb{R}) \mid \det A = 1\},\\
\mathrm{O}(n) &= \{A \in \mathrm{Mat}(n \times n, \mathbb{R}) \mid A^tA - \mathrm{Id} = 0\},\\
\mathrm{SO}(n) &= \{A \in \mathrm{Mat}(n \times n, \mathbb{R}) \mid A^tA - \mathrm{Id} = 0 \text{ and } \det A = 1\},\\
\mathrm{SL}(n, \mathbb{C}) &= \{A \in \mathrm{Mat}(n \times n, \mathbb{C}) \mid \det A = 1\},\\
\mathrm{U}(n) &= \{A \in \mathrm{Mat}(n \times n, \mathbb{C}) \mid A^\dagger A - \mathrm{Id} = 0\},\\
\mathrm{SU}(n) &= \{A \in \mathrm{Mat}(n \times n, \mathbb{C}) \mid A^\dagger A - \mathrm{Id} = 0 \text{ and } \det A = 1\}.
\end{aligned}$$

Multipication and inversion are smooth as in Example 10.2. Hence they are Lie groups.

10.1. **The Lie algebra of a Lie group.** Lie groups have an "infinitesimal version" which is a Lie algebra (recall Definition 7.17 on page 39

[42]In several textbooks, Lie groups are equivalently defined by requiring that the combined map

$$\begin{array}{ccc} G \times G & \to & G \\ (g,h) & \mapsto & gh^{-1} \end{array}$$

be smooth.

and the subsequent examples). Indeed we will show that the tangent space at the identity e of a Lie group G,

$$\mathfrak{g} := T_eG,$$

has a natural Lie algebra structure.

Remark 10.4. It is customary to denote a Lie group by a capital roman leter and its tangent space at the identity, viewed as a Lie algebra, by the corresponding letter in lowercase gothic.

First notice that, if G is a Lie group, then left multiplication by g

$$\begin{array}{rccc} l_g\colon & G & \to & G \\ & h & \mapsto & gh \end{array}$$

is a smooth map (actually, a diffeomorphism) for each $g \in G$. Notice that the differential of l_g at h yields a linear map

$$\mathrm{d}_h l_g\colon T_hG \to T_{gh}G.$$

A vector field X on G is called left invariant if

$$\boxed{X(gh) = \mathrm{d}_h l_g X(h)} \tag{10.1}$$

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

$$(l_g)_*X = X$$

for all $g \in G$. If X and Y are left-invariant vector fields, then

$$(l_g)_*[X,Y] = [(l_g)_*X, (l_g)_*Y] = [X,Y],$$

so also $[X,Y]$ is left invariant. Hence, the $\mathbb{R}$-vector space $\mathfrak{X}(G)^G$ of left-invariant vector fields is a Lie algebra.

Notice that by specializing (10.1) at $h = e$, where e is the identity element, we get

$$X(g) = \mathrm{d}_e l_g X(e).$$

This shows that a left-invariant vector field is completely determined by its value at the identity. More precisely, we have an isomorphism of $\mathbb{R}$-vector spaces

$$\begin{array}{ccc} \mathfrak{g} & \to & \mathfrak{X}(G)^G \\ \xi & \mapsto & X_\xi \end{array}$$

with $X_\xi(g) = \mathrm{d}_e l_g \xi$, with inverse

$$\begin{array}{ccc} \mathfrak{X}(G)^G & \to & \mathfrak{g} \\ X & \mapsto & X(e) \end{array}$$

Definition 10.5. The Lie algebra of a Lie group is its tangent space at the identity with the Lie bracket induced by its identification with the vector space of left-invariant vector fields. Namely,

$$X_{[\xi,\eta]} = [X_\xi, X_\eta] \tag{10.2}$$

for all $\xi, \eta \in \mathfrak{g}$.

We end this section by the following observation.

Lemma 10.6. *Every Lie group is parallelizable (and hence orientable).*

Proof. The map

$$\begin{array}{ccc} G \times \mathfrak{g} & \to & TG \\ (g,\xi) & \mapsto & (g, \mathrm{d}_e l_g \xi) \end{array}$$

is an isomorphism of vector bundles. □

This yields a new proof of the fact that S^1 and S^3 are parallelizable, see Lemmata 8.25 and 8.31. In fact, S^1 is diffeomorphic to $SO(2)$ by the map

$$\begin{array}{ccc} S^1 & \to & SO(2) \\ \theta & \mapsto & \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \end{array}$$

On the other hand:

Lemma 10.7. *S^3 is diffeomorphic to $SU(2)$.*

Proof. Consider a complex 2×2 matrix $\mathsf{A} = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$ with determinant 1. Its inverse is then $\left(\begin{smallmatrix} \delta & -\beta \\ -\gamma & \alpha \end{smallmatrix}\right)$. Equating it to its adjoint $\left(\begin{smallmatrix} \bar\alpha & \bar\gamma \\ \bar\beta & \bar\delta \end{smallmatrix}\right)$, to make sure that A is in $SU(2)$, yields $\bar\delta = \alpha$ and $\gamma = -\bar\beta$. As a consequence, our matrix A has the form $\left(\begin{smallmatrix} \alpha & \beta \\ -\bar\beta & \bar\alpha \end{smallmatrix}\right)$ with $|\alpha|^2+|\beta|^2 = 1$. We have hence proved that

$$SU(2) = \left\{ \begin{pmatrix} \alpha & \beta \\ -\bar\beta & \bar\alpha \end{pmatrix} \in \mathrm{Mat}(2\times 2, \mathbb{C}) \mid |\alpha|^2 + |\beta|^2 = 1 \right\}.$$

If we identify $\mathbb{C}^2$ with $\mathbb{R}^4$, we see that the equation $|\alpha|^2 + |\beta|^2 = 1$, for $(\alpha,\beta) \in \mathbb{C}^2$, defines S^3. □

10.1.1. *The Lie algebra of matrix Lie groups.* Consider first the case $G = \mathrm{GL}(n)$. First observe that, since $\mathrm{GL}(n)$ is an open subset of $\mathrm{Mat}(n\times n, \mathbb{R})$ with its standard manifold structure, we have

$$\mathfrak{gl}(n) := T_e GL(n) = \mathrm{Mat}(n\times n, \mathbb{R}),$$

where e now denotes the identity matrix. Since the group multiplication is a linear map, its differential is exactly the same linear map; so the left-invariant vector field X_ξ, as a map from $\mathrm{GL}(n)$ to $\mathrm{Mat}(n\times n,\mathbb{R})$, corresponding to a matrix $\xi \in \mathfrak{gl}(n)$ is simply given by

$$X_\xi(A) = A\xi, \quad A \in \mathrm{GL}(n), \tag{10.3}$$

where on the left hand side we just use matrix multiplication. Left-invariance is just given by $X_\xi(AB) = AX_\xi(B)\ \forall A, B \in \mathrm{GL}(n)$. As a derivation, X_ξ can be written as

$$X_\xi(A) = \sum_{j,k=1}^{n} A_{ij}\xi_{jk}\frac{\partial}{\partial A_{ik}},$$

so $[X_\xi, X_\eta] = X_{[\xi,\eta]}$ with $[\xi, \eta] = \xi\eta - \eta\xi$ the usual commutator in $\mathrm{Mat}(n \times n, \mathbb{R})$. Hence we have proved that the Lie algebra of $\mathrm{GL}(n)$ is the vector space of $n \times n$-matrices with Lie bracket given by the commutator.

Similarly, the Lie algebra of $\mathrm{GL}(n, \mathbb{C})$ is $\mathrm{Mat}(n \times n, \mathbb{C})$, regarded as an $\mathbb{R}$-vector space, with Lie bracket given by the commutator.

If G is a submanifold of $\mathrm{GL}(n)$ (or $\mathrm{GL}(n, \mathbb{C})$) given by constraints satisfying the conditions of the implicit function theorem, then its tangent space as the identity may be computed as in Remark 6.10. In particular, it is a subspace of $\mathrm{Mat}(n\times n, \mathbb{R})$ (or $\mathrm{Mat}(n\times n, \mathbb{C})$) and the Lie bracket is still given by the commutator.

For the matrix Lie groups introduced in Example 10.3 we have to linearize the determinant and the quadratic functions A^tA and $A^\dagger A$ at the identity. Notice that

$$\begin{aligned}
\det(e + t\xi + O(t^2)) &= 1 + t\,\mathrm{Tr}\,\xi + O(t^2),\\
(e + t\xi + O(t^2))^t(e + t\xi + O(t^2)) &= e + t(\xi^t + \xi) + O(t^2),\\
(e + t\xi + O(t^2))^\dagger(e + t\xi + O(t^2)) &= e + t(\xi^\dagger + \xi) + O(t^2).
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathfrak{sl}(n) &= \{\xi \in \mathrm{Mat}(n \times n, \mathbb{R}) \mid\ \mathrm{Tr}\,\xi = 0\},\\
\mathfrak{o}(n) &= \{\xi \in \mathrm{Mat}(n \times n, \mathbb{R}) \mid \xi^t + \xi = 0\},\\
\mathfrak{so}(n) &= \{A \in \mathrm{Mat}(n \times n, \mathbb{R}) \mid \xi^t + \xi = 0 \text{ and } \mathrm{Tr}\,\xi = 0\},\\
\mathfrak{sl}(n, \mathbb{C}) &= \{\xi \in \mathrm{Mat}(n \times n, \mathbb{C}) \mid\ \mathrm{Tr}\,\xi = 0\},\\
\mathfrak{u}(n) &= \{\xi \in \mathrm{Mat}(n \times n, \mathbb{C}) \mid \xi^\dagger + \xi = 0\},\\
\mathfrak{su}(n) &= \{\xi \in \mathrm{Mat}(n \times n, \mathbb{C}) \mid \xi^\dagger + \xi = 0 \text{ and } \mathrm{Tr}\,\xi = 0\}.
\end{aligned}$$

Notice that, since the trace of a skew-symmetric matrix vanishes automatically, we have

$$\mathfrak{so}(n) = \mathfrak{o}(n)$$

for all n.

10.2. **The exponential map.** A map $\phi\colon G \to G$ is called left invariant if $\phi(gh) = g\phi(h)\ \forall g, h \in G$.

Proposition 10.8. *The flow Φ_t^X of a left-invariant vector field X is left invariant.*

Proof. Fix $g \in G$ and define $\psi_t(h) := g^{-1}\Phi_t^X(gh)$. We have $\psi_0 = \mathrm{Id}$ and

$$\begin{aligned}\frac{\partial}{\partial t}\psi_t(h) &= \mathrm{d}_{\Phi_t^X(gh)} l_{g^{-1}} \frac{\partial}{\partial t}\Phi_t^X(gh) = \mathrm{d}_{\Phi_t^X(gh)} l_{g^{-1}} X(\Phi_t^X(gh)) = \\ &= \mathrm{d}_{g\psi_t(h)} l_{g^{-1}} X(g\psi_t(h)) = \mathrm{d}_{g\psi_t(h)} l_{g^{-1}} \mathrm{d}_{\psi_t(h)} l_g X(\psi_t(h)) = X(\psi_t(h)).\end{aligned}$$

By the uniqueness of solutions to ODEs, we get $\psi_t = \Phi_t^X$, which completes the proof. □

As a consequence we have

$$\Phi_t^X(g) = g\Phi_t^X(e), \tag{10.4}$$

so it is enough to understand the flow starting at e.

Lemma 10.9. *A left-invariant vector field is complete: i.e., its flow is defined for all time.*

Proof. By the existence and uniqueness theorem, there is an $\epsilon > 0$ such that $\Phi_t^X(e)$ is defined for all $t \in (-\epsilon, \epsilon)$. By (10.4) we conclude that $\Phi_t^X(g)$ is defined for all $t \in (-\epsilon, \epsilon)$ and for all $g \in G$. Next observe that for every $t \in \mathbb{R}$ we can always find an $n \in \mathbb{N}$ such that $\frac{|t|}{n} < \epsilon$. As a consequence, $\phi_{X,\frac{t}{n}}(g)$ is defined for all $g \in G$. By left invariance of the flow, we get that $\Phi_t^X(e) = \Phi_{\frac{t}{n}}^X \circ \cdots \circ \Phi_{\frac{t}{n}}^X(e) = \left(\Phi_{\frac{t}{n}}^X(e)\right)^n$, which shows that $\Phi_t^X(e)$ is defined for all $t \in \mathbb{R}$. Finally, by (10.4) again, we conclude that $\Phi_t^X(g)$ is defined for all $t \in \mathbb{R}$ and for all $g \in G$. □

Definition 10.10. For $\xi \in G$ we define

$$\exp \xi := \phi_1^{X_\xi}(e).$$

The smooth map $\exp\colon \mathfrak{g} \to G$ is called the exponential map.

Notice that the exponential map is in general neither injective nor surjective.

Lemma 10.11.

$$\exp(t\xi) := \Phi_t^{X_\xi}(e), \quad \forall t \in \mathbb{R},\ \forall \xi \in \mathfrak{g}.$$

Proof. This follows from a general property of flows. Namely, let X be a vector field on some manifold M and let $x(t)$ be a solution of the ODE $\dot{x} = X(x)$. Then $x_\lambda(t) := x(\lambda t)$ solves the ODE $\dot{x}_\lambda = \lambda X(x)$ and has the same initial value at $t = 0$. It follows that $\Phi_t^{\lambda X} = \Phi_{\lambda t}^X$ for all λ and t for which the flow is defined. In our case, $\Phi_1^{tX_\xi}(e) = \Phi_t^{X_\xi}(e)$ for all t. The thesis now follows from linearity: $tX_\xi = X_{t\xi}$. □

Proposition 10.12.

$$\exp((t+s)\xi) = \exp(t\xi)\exp(s\xi), \quad \forall t, s \in \mathbb{R},\ \forall \xi \in \mathfrak{g}.$$

This explains the name of exponential map. Notice however that in general $\exp(\xi + \eta) \neq \exp\xi \exp\eta$.

Proof. We have

$$\begin{aligned}\exp((t+s)\xi) = \Phi_{t+s}^{X_\xi}(e) = \Phi_t^{X_\xi}(\Phi_s^{X_\xi}(e)) =& \\ = \Phi_t^{X_\xi}(e)\Phi_s^{X_\xi}(e) &= \exp(t\xi)\exp(s\xi).\end{aligned}$$

□

We can also recover the Lie bracket in $\mathfrak{g}$ from the exponential map:

Lemma 10.13. *For all $\xi, \eta \in \mathfrak{g}$ we have*

$$\exp(s\eta)\exp(t\xi)\exp(-s\eta)\exp(-t\xi) = \exp(st[\eta,\xi] + O(t^2, s^2)). \tag{10.5}$$

In particular,

$$\left.\frac{\partial^2}{\partial s \partial t}\right|_{s=t=0} \exp(s\eta)\exp(t\xi)\exp(-s\eta)\exp(-t\xi) = [\eta, \xi].$$

Proof. Let $\Phi_{s,t} := \exp(s\eta)\exp(t\xi)\exp(-s\eta)\exp(-t\xi)$.

Since $\Phi_{s,0} = e$ for all s, we have $\left.\frac{\partial}{\partial s}\right|_{s=t=0} \Phi_{s,t} = 0$. Similarly, we see that $\left.\frac{\partial}{\partial t}\right|_{s=t=0} \Phi_{s,t} = 0$

By Lemma 10.11, we have

$$\Phi_{s,t} = \Phi_s^{X_\eta}(e)\Phi_t^{X_\xi}(e)\Phi_{-s}^{X_\eta}(e)\Phi_{-t}^{X_\xi}(e)$$

By iterating (10.4), we then have

$$\Phi_s^{X_\eta}(e)\Phi_t^{X_\xi}(e)\Phi_{-s}^{X_\eta}(e)\Phi_{-t}^{X_\xi}(e) = \Phi_{-t}^{X_\xi}(\Phi_{-s}^{X_\eta}(\Phi_t^{X_\xi}(\Phi_s^{X_\eta}(e)))).$$

By Lemma 7.34, we finally get

$$\left.\frac{\partial^2}{\partial s \partial t}\right|_{s=t=0} \Phi_{-t}^{X_\xi}(\Phi_{-s}^{X_\eta}(\Phi_t^{X_\xi}(\Phi_s^{X_\eta}(e)))) = [X_\eta, X_\xi](e) = [X_{[\eta,\xi]}](e) = [\eta, \xi].$$

□

Remark 10.14. There are several equivalent ways to rewrite (10.5). For example,

$$\exp(s\eta)\exp(t\xi) = \exp\left(s\eta + t\xi + \frac{1}{2}st[\eta,\xi] + O(t^2, s^2)\right).$$

or

$$\exp(s\eta)\exp(t\xi)\exp(-s\eta) = \exp(t\xi + st[\eta,\xi] + O(t^2, s^2)),$$

Remark 10.15 (The BCH formula)**.** Since the differential of the exponential map at 0 is the identity map (this follows immediately from Lemma 10.11), the exponential map is a diffeomorphism from a neighborhood of 0 in $\mathfrak{g}$ to a neighborhood of e in G. This means that for ϵ small enough, the will be a unique ζ with $\exp(\epsilon\eta)\exp(\epsilon\xi) = \exp(\zeta)$. the Baker–Campbell–Hausdorff (BCH) formula is the Taylor expansion $\mathrm{BCH}(\eta, \xi)$ of ζ with respect to ϵ. By the above Remark it starts with

$$\mathrm{BCH}(\eta,\xi) = \epsilon\eta + \epsilon\xi + \frac{\epsilon^2}{2}[\eta,\xi] + \cdots.$$

Another useful formula is the following:

Lemma 10.16. *For every $\xi \in \mathfrak{g}$ and for every t, we have*

$$\frac{\partial}{\partial t}\exp(t\xi) = \mathrm{d}_e l_{\exp(t\xi)}\xi. \tag{10.6}$$

Proof.

$$\frac{\partial}{\partial t}\exp(t\xi) = \frac{\partial}{\partial t}\Phi_t^{X_\xi}(e) = X_\xi(\Phi_t^{X_\xi}(e)) = X_\xi(\exp(t\xi)) = \mathrm{d}_e l_{\exp(t\xi)}\xi.$$

□

10.2.1. *The exponential map of matrices.* Consider the group $\mathrm{GL}(n)$. By (10.3), the ODE associated to X_ξ is $\dot{A} = A\xi$, whose solution with initial condition $A(0) = A_0$ is $A(t) = A_0\mathrm{e}^{t\xi}$ with

$$\mathrm{e}^\xi = \sum_{n=0}^\infty \frac{\xi^n}{n!}.$$

The flow is then $\Phi_t^{X_\xi}(A) = A\mathrm{e}^{t\xi}$. We clearly have $\Phi_t^{X_\xi}(AB) = AB\mathrm{e}^{t\xi} = A\Phi_t^{X_\xi}(B)$, which shows left invariance. Moreover, we have $\exp\xi = \mathrm{e}^\xi$ for all $\xi \in \mathrm{Mat}(n\times n, \mathbb{R})$. Similarly, for $G = \mathrm{GL}(n, \mathbb{C})$ we get $\exp\xi = \mathrm{e}^\xi$ for all $\xi \in \mathrm{Mat}(n \times n, \mathbb{C})$

If G is subgroup of $\mathrm{GL}(n)$ or $\mathrm{GL}(n, \mathbb{C})$ given by the implicit function theorem as in subsection 10.1.1, we then have $\exp\xi = \mathrm{e}^\xi$ for all $\xi \in \mathfrak{g}$.

Finally note that Lemma 10.13 has a very simple interpretation in this case. Namely, interpreting η and ξ as matrices,

$$\begin{aligned}\exp(s\eta)\exp(t\xi) &= (1+s\eta+O(s^2))(1+t\xi+O(t^2)) = \\ &= 1+s\eta+t\xi+st\eta\xi+O(t^2,s^2),\\ \exp(-s\eta)\exp(-t\xi) &= (1-s\eta+O(s^2))(1-t\xi+O(t^2)) = \\ &= 1-s\eta-t\xi+st\eta\xi+O(t^2,s^2).\end{aligned}$$

Hence,

$$\exp(s\eta)\exp(t\xi)\exp(-s\eta)\exp(-t\xi) = 1+st[\eta,\xi]+O(t^2,s^2).$$

10.3. **Morphisms.** Let G_1 and G_2 be Lie groups. A Lie group morphism $\Psi\colon G_1 \to G_2$ is a group morphism that is also a smooth map. Let $\psi := \mathrm{d}_e\Psi\colon \mathfrak{g}_1 \to \mathfrak{g}_2$.

Lemma 10.17. *For every $\xi \in \mathfrak{g}_1$ and for every t, we have*

$$\Psi(\exp(t\xi)) = \exp(t\psi(\xi)).$$

Proof. For fixed ξ, define

$$U_t := \Psi(\exp(t\xi)) \qquad \text{and} \qquad V_t(g) := \exp(t\psi(\xi)).$$

We have $U_0 = e = V_0$. We want to show that U and V satisfy the same differential equation, so they must be equal by uniqueness of solutions. To compute their time derivatives, we use (10.6) and the identity $\Psi \circ l_h = l_{\Psi(h)} \circ \psi$ for all $h \in G_1$, which yields

$$\mathrm{d}_h\Psi \mathrm{d}_e l_h = \mathrm{d}_e l_{\Psi(h)} \mathrm{d}_e\Psi.$$

Then both U and V solve the equation

$$\frac{\partial}{\partial t}X_t = \mathrm{d}_e l_{X_t}\psi(\xi).$$

□

Proposition 10.18. *ψ is a morphism of Lie algebras.*

Proof. Apply Ψ to (10.5) and use Lemma 10.17. □

In the language of categories, this shows that there is a functor Lie from the category of Lie groups to the category of Lie algebra. With the previous notations, $\mathsf{Lie}(G) = \mathfrak{g}$ and $\mathsf{Lie}(\Psi) = \psi$.

The interesting question is whether we have some sort of inverse to this. The passage from Lie groups to Lie algebras is sometimes called differentiation and the inverse process integration. We start with the special case when we have an injective morphism $\mathfrak{h} \to \mathfrak{g}$; i.e., $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$. We may expect this to be integrable to a Lie subgroup. This is indeed the case if we define this in an appropriate way.

Definition 10.19. A $\mathsf{Lie\ subgroup}$ of a Lie group G is a Lie group H together with a Lie group morphism $I\colon H \to G$ which is an immersion.

Note that a Lie subgroup is not necessarily closed.

Theorem 10.20 (Lie I)**.** *Let $\iota\colon \mathfrak{h} \to \mathfrak{g}$ be a Lie subalgebra. Then there is a unique connected Lie subgroup H of G with $\mathsf{Lie}(H) = \mathfrak{h}$ and $\mathsf{Lie}(I) = \iota$.*

Proof. Let $\Delta_g := \mathrm{d}_e l_g \mathfrak{h}$. This is clearly an involutive distribution spanned by the left invariant vector fields corresponding to elements of $\mathfrak{h}$. By the Frobenius theorem, there is then a unique maximal integral submanifold $I\colon H \to G$ passing through e. Also note that, by definition, the differential of I at the identity is precisely ι. It remains to show that H is a subgroup.

By left invariance, the maximal integral submanifold passing through g is $l_g \circ I$. This means that, for each h in H, gh belongs to the leaf through g. If g is also in H, this leaf is H itself. Thus, H is closed under multplication. If $g = h^{-1}$, then $e = h^{-1}h$ belongs to the leaf through h^{-1}, which shows that $h^{-1} \in H$. □

Example 10.21. Let $G = SO(2) = S^1 \subset \mathbb{C}$. Then $\mathfrak{g} = \mathbb{R}$. Let $\mathfrak{h}$ be also $\mathbb{R}$ and ι be the identity map. Then $H = \mathbb{R}$ and $I(t) = \mathrm{e}^{it}$.

If ι is also surjective, then I is also a submersion. If G is connected, this means that I is a local diffeomorphism. If G is simply connected, then I is a diffeomorphism. This will be important in the proof of the next theorem.

Theorem 10.22 (Lie II)**.** *Let G_1 and G_2 be Lie groups with G_1 simply connected. Let $\psi\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ be a Lie algebra morphism. Then there is a unique Lie group morphism $\Psi\colon G_1 \to G_2$ with $\psi = \mathsf{Lie}(\Psi)$.*

Proof. The product $G_1 \times G_2$ is also a Lie group and the projections $\Pi_i\colon G_1 \times G_2 \to G_i$ are Lie group morphisms. The Lie algebra of $G_1 \times G_2$ is $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ and the projections are $\pi_i = \mathsf{Lie}(\Pi_i)$.

The graph Γ of ψ is a Lie subalgebra of $\mathfrak{g}_1 \oplus \mathfrak{g}_2$. Denoting by ι the inclusion of Γ, we have that $\psi_1 := \pi_1\iota$ is an isomorphism and that $\pi_2\psi_1^{-1} = \psi$.

Bye Lie I there is a Lie subgroup $I\colon H \to G_1 \times G_2$ corresponding to Γ. Since $\mathsf{Lie}(\Pi_1 \circ I) = \psi_1$ and G_1 is simply connected, we have that $\Psi_1 := \Pi_1 \circ I$ is an isomorphism. We then get a group morphism $\Psi\colon G_1 \to G_2$ by $\Psi := \Pi_2 \circ \Psi_1^{-1}$, and we clearly have $\mathsf{Lie}(\Psi) = \psi$. □

To complete the picture we add the following theorem (without proof).

Theorem 10.23 (Lie III)**.** *Let $\mathfrak{g}$ be a finite dimensional Lie algebra over $\mathbb{R}$. Then there is a Lie group G with $\mathsf{Lie}(G) = \mathfrak{g}$. One can choose G to be simply connected and under this assumption it is unique up to isomorphisms (by Lie II).*

The complete proof of this last theorem is due to E. Cartan.

Remark 10.24 (Representations)**.** Let V be a finite dimensional vector space. Then $\mathrm{Aut}(V)$ is a Lie group (isomorphic to $GL(n)$ with $n = \dim V$). A representation of G on V is a Lie group morphism $R\colon G \to \mathrm{Aut}(V)$. The Lie algebra of $\mathrm{Aut}(V)$ is $\mathrm{End}(V)$, so the above construction associates to R a Lie algebra morphism $r\colon \mathfrak{g} \to \mathrm{End}(V)$, i.e., a Lie algebra representation. Lie II implies that Lie algebra representations can be lifted to Lie group representations if the Lie group is simply connected.

Example 10.25. Let G be an $(n \times n)$-matrix Lie group. The defining representation of G is the inclusion of G into $GL(n)$; i.e., one regards each element of G as a matrix acting on $\mathbb{R}^n$ (or $\mathbb{C}^n$ in the complex case). The corresponding Lie algebra representation is the inclusion of $\mathfrak{g}$ into $\mathfrak{gl}(n)$.

Example 10.26 (The adjoint representation)**.** Consider the conjugation $C_g\colon G \to G$, $h \to ghg^{-1}$. Let $\mathrm{Ad}_g := \mathrm{d}_e C_g$. Since $C_g(e) = e$, Ad_g is an automorphism of $\mathfrak{g}$. Differentiating $C_{g_1 g_2} = C_{g_1} \circ C_{g_2}$ shows that Ad is a representation, called the adjoint representation. The induced Lie algebra representation, denoted by ad, turns out to be given by

$$\mathrm{ad}_\eta \xi = [\eta, \xi].$$

In fact, let $\Psi_{s,t} := C_{\exp(s\eta)} \exp(t\xi)$. Then $\frac{\partial}{\partial t}\Big|_{t=0} \Psi_{s,t} = \mathrm{Ad}_{\exp(s\eta)} \xi$, so $\mathrm{ad}_\eta \xi = \frac{\partial^2}{\partial s \partial t}\Big|_{s=t=0} \Psi_{s,t}$. On the other hand, $\Psi_{s,t} = \exp(s\eta)\exp(t\xi)\exp(-s\eta)$. The statement now follows essentially from Lemma 10.13, see Remark 10.14.

10.4. **Actions of Lie groups.** Recall that the action of a group G on a set M is a group homorphism

$$\begin{array}{ccc} G & \to & \mathrm{Iso}(M) \\ g & \mapsto & \Psi_g \end{array}$$

where $\mathrm{Iso}(M)$ is the group of invertible maps $M \to M$. Explicitly, Ψ_g is a map $M \to M$ for all $g \in G$ satisfying $\Psi_e = \mathrm{Id}$ and $\Psi_{gh} = \Psi_g \circ \Psi_h$ for all $g, h \in G$. There is also an associated map

$$\begin{array}{rccc} \Psi\colon & G \times M & \to & M \\ & (g, m) & \mapsto & \Psi_g(m) \end{array}$$

Definition 10.27. An action of a Lie group G on a smooth manifold M is an action as above where Ψ is smooth.

Notice that it follows that Ψ_g is a diffeomorphism of M for all $g \in G$. An example of action is a representation (where the restriction is that the manifold acted upon is a vector space and the maps are linear).

Lemma 10.28. *For all $\xi \in \mathfrak{g}$, the map $\widetilde{\Psi}^\xi_t := \Psi_{\exp(t\xi)}$ is a flow on M.*

Proof. We have $\widetilde{\Psi}^\xi_0 = \Psi_{\exp 0} = \Psi_e = \mathrm{Id}$ and $\widetilde{\Psi}^\xi_{t+s} = \Psi_{\exp((t+s)\xi)} = \Psi_{\exp(t\xi)\exp(s\xi)} = \Psi_{\exp(t\xi)} \circ \Psi_{\exp(s\xi)} = \widetilde{\Psi}^\xi_t \circ \widetilde{\Psi}^\xi_s$. □

Let ψ_ξ denote the vector field on M that generates the flow $\widetilde{\Psi}^\xi_t$. Namely, $\psi_\xi = \frac{\partial}{\partial t}_{|t=0} \widetilde{\Psi}^\xi_t$ and $\widetilde{\Psi}^\xi_t = \Phi^{\psi_\xi}_t$.

Proposition 10.29. *The map $\psi\colon \mathfrak{g} \to \mathfrak{X}(M)$ is a Lie algebra morphism; i.e., ψ is linear and satisfies*

$$[\psi_\xi, \psi_\eta] = \psi_{[\xi,\eta]}, \quad \forall \xi, \eta \in \mathfrak{g}.$$

A Lie algebra morphism $\mathfrak{g} \to \mathfrak{X}(M)$ is called an infinitesimal action of $\mathfrak{g}$ on M.

Proof. Apply Ψ to (10.13) and use Lemma 7.34. □

If G acts freely on M (i.e. $\Psi_g(m) = m \Rightarrow g = e$), then for each $m \in M$ we have an injective immersion $O_m\colon G \to M$, $g \mapsto \Psi_g m$ (the orbit through m). The corresponding integrable distribution is $\Delta_m := \operatorname{span}_{\xi\in\mathfrak{g}} \psi_\xi(m)$. In this case, as a set of generators for the vertical vector fields one may choose $\psi(\mathfrak{g})$. Also note that the leaf space for this distribution is the same as the quotient M/G (i.e., the quotient by the equivalent relation: $m \sim m'$ if and only if $\exists g \in G$ $m = \Psi_g m'$). Distributions that come from Lie group actions are under better control.

Definition 10.30. An action Ψ of a Lie group G on a manifold M is called proper if the map

$$\begin{aligned} G \times M &\to M \times M \\ (g, m) &\mapsto (\Psi_g(m), m) \end{aligned}$$

is proper.[43]

Theorem 10.31. *If the action of G on M is free and proper, then G/M has a manifold structure for which the canonical projection map is a submersion.*

[43]A map is called proper if the preimages of compact sets are compact.

Not that the action of a compact Lie group is automatically proper. Another class of examples is the action of a closed Lie subgroup H of a Lie group G on G itself: $\Psi_h(g) = gh^{-1}$ (conventionally one chooses the action from the right). In this case, the quotient manifold G/H still has a transitive G-action (if we let H act from the right, we may let G act from the left; transitive means that any two points can be related by the action of some element of G) and is called a homogeneous space. If H is normal, then G/H inherits a Lie group structure. The Lie algebra $\mathfrak{h}$ of H is a Lie ideal in $\mathfrak{g}$, and the Lie algebra of G/H is $\mathfrak{g}/\mathfrak{h}$.

10.5. **Left invariant forms.** A differential form ω on G is called left invariant if $l_g^*\omega = \omega$ for all $g \in G$. We denote by $\Omega^k(G)^G$ the space of left invariant k-forms. A left invariant form is completely determined by its value at e. Thus, the map

$$\phi\colon\ \Omega^\bullet(G)^G \ \to \Lambda^\bullet g^*$$

is an isomorphism.

A nonzero element of $\Lambda^{\text{top}} g^*$ then defines a left invariant volume form ω. If we choose the element to be positive, then we have a left invariant measure $\mu(A) := \int_A \omega$. If G is compact we may produce a biinvariant measure $\widetilde{\omega}$ by averaging:

$$\widetilde{\omega} = \int_G r_g^*\omega \,\mathrm{d}\mu(g).$$

We clearly see that $\widetilde{\omega}$ is positive and left invariant. Moreover,

$$r_h^*\widetilde{\omega} = \int_G r_h^* r_g^*\omega \,\mathrm{d}\mu(g) = \int_G r_{hg}^*\omega \,\mathrm{d}\mu(g) = \widetilde{\omega},$$

so it is also right invariant. By rescaling we may also assume $\int_G \widetilde{\omega} = 1$. The corresponding probability measure is known as Haar measure.

The wedge product of two left invariant differential forms is left invariant. Moreover, by the Cartan calculus, the de Rham differential maps $\Omega(G)^G$ to itself. This induces a differential $\delta = \phi \mathrm{d} \phi^{-1}$ on $\Lambda^\bullet g^*$. Applying the formula in Proposition 9.32 to left invariant vector fields, we get

$$\delta\omega(v_0, \dots, v_k) = \sum_{0 \le i < j \le k} (-1)^{i+j} \omega([v_i, v_j], v_0, \dots, \widehat{v_i}, \dots \widehat{v_j}, \dots, v_k),$$

for every $\omega \in \Lambda^k \mathfrak{g}^*$ and every $v_0, \dots, v_k \in \mathfrak{g}$.[44] Note, in particular, that $\delta\omega = 0$ if $\omega \in \Lambda^0 \mathfrak{g} = \mathbb{R}$ and that for $\omega \in \Lambda^1 \mathfrak{g} = \mathfrak{g}$ we have

$$\delta\omega(v_0, v_1) = -\omega([v_0, v_1]).$$

[44] The first term in the formula is not there, for Lemma 9.31 implies that the contraction of a left invariant form with left invariant vector fields is left invariant.

That is, $\delta\colon \mathfrak{g}^* \to \Lambda^2\mathfrak{g}^*$ is minus the transposition of the Lie bracket (as a linear map $\Lambda^2\mathfrak{g} \to \mathfrak{g}$). As δ is a derivation, it is enough to know its action in degree zero and one.

Appendix A. Topology

We recall a few facts about topology.

Definition A.1. A topology on a set S is a collection $\mathcal{O}(S)$ of subsets of S such that

(1) $\emptyset, S \in \mathcal{O}(S)$;
(2) $\forall U, V \in \mathcal{O}(S)$ we have $U \cap V \in \mathcal{O}(S)$;
(3) if $(U_\alpha)_{\alpha\in I}$ is a family indexed by I with $U_\alpha \in \mathcal{O}(S)$ $\forall \alpha \in I$, we have $\cup_{\alpha\in I} U_\alpha \in \mathcal{O}(S)$.

A set with a topology is called a topological space.

Example A.2. The collection of the usual open subsets[45] of $\mathbb{R}^n$ forms a topology on $\mathbb{R}^n$, called its standard topology.

In general, elements of a topology are called open sets and elements of a topological space are called points. A neighborhood of a point is a set containing an open set that contains the given point. An open cover of a topological space S is a collection $\{U_\alpha\}_{\alpha\in I}$ of open sets in S such $\cup_{\alpha\in I} U_\alpha = S$.

Definition A.3. A map $F\colon S \to T$ between topological spaces $(S, \mathcal{O}(S))$ and $(T, \mathcal{O}(T))$ is called continuous if $F^{-1}(U) \in \mathcal{O}(S)$ $\forall U \in \mathcal{O}(T)$. A continous invertible map whose inverse is also continuous is called a homeomorphism. A map that maps open sets to open sets (i.e., in the abve notation, $F(U) \in \mathcal{O}(T)$ $\forall U \in \mathcal{O}(S)$) is called open.

Topologies may often be derived from other topologies.

Example A.4. Let $(S, \mathcal{O}(S))$ be a topological space and let T be a subset. Then

$$\mathcal{O}_S(T) := \{U \subset T \mid \exists V \in \mathcal{O}(S) : U = V \cap T\}$$

is a topology on T. With this topology, called the induced topology or the relative topology, the inclusion map $\iota\colon T \hookrightarrow S$ is continuous.

This is in particular the topology one usually considers on subsets of $\mathbb{R}^n$ with its standard topology.

[45]Recall that a subset U of $\mathbb{R}^n$ is defined to be open if for each $x_0 \in U$ there is an $R > 0$ such that the open ball

$$\{x \in \mathbb{R}^n \mid ||x - x_0|| < R\}$$

is contained in U.

Example A.5. Let $(S, \mathcal{O}(S))$ be a topological space and $\pi\colon S \to T$ be a surjective map. Then

$$\mathcal{O}_{S,\pi}(T) := \{U \subset T \mid \pi^{-1}(U) \in \mathcal{O}(S)\}$$

is a topology on T. With this topology, called the quotient topology, π is continuous.

Notice in particular that π arises when we have a quotient relation on S and define T as the set of equivalence classes.

Remark A.6. Unless stated otherwise, when we speak of $\mathbb{R}^n$, we tacitly assume the standard topology; when speaking of a subset of a topological space or a quotient of a topological space, we tacitly assume the induced topology.

For some consideration on manifolds, we also need the notion of compactness, which we briefly recall. First, recall that a subcover of an open cover $\{U_\alpha\}_{\alpha\in I}$ of S is a subcollection $\{U_\alpha\}_{\alpha\in J}$, $J \subset I$, that is still a cover (i.e., $\cup_{\alpha\in J} U_\alpha = S$). If J is finite, then the subcover is also called finite.

Definition A.7. A topological space K is called compact if every open cover of K possesses a finite subcover. A subset K of a topological space T is called compact if it is compact in the induced topology.

Recall that by the Heine–Borel theorem a subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded (i.e., contained in a ball of finite radius). Compact sets have a lot of important properties, for which we refer to textbooks in topology (or analysis). We only recall those that we are using in these notes.

Lemma A.8. *Let $F\colon S \to T$ be a continuous map of topological spaces. If K is compact in S, then $F(K)$ is compact in T. In particular, if a set is compact in a subset S of a topological space T (with respect to the induced topology), then it is compact also in T.*

Proof. Let $\{U_\alpha\}_{\alpha\in I}$ be an open cover of $F(K)$. Then $\{F^{-1}(U_\alpha)\}_{\alpha\in I}$ is an open cover of K. Since K is compact, there is a finite subcover $\{U_\alpha\}_{\alpha\in J}$, J a finite subset of I. But then $\{U_\alpha\}_{\alpha\in J}$ is a finite subcover of $F(K)$.

For the second stament, just recall that the inclusion map of a subset is continuous with respect to the induced topology. □

Lemma A.9. *A closed subset of a compact set is compact.*

Proof. Let K be compact and $C \subset K$ closed. Then $K \setminus C$ is by definition open. Let $\{U_\alpha\}_{\alpha\in I}$ be an open cover of C. Then $\{U_\alpha\}_{\alpha\in I} \cup$

$K \setminus C$ is an open cover of K and hence possesses a finite subcover. Removing $K \setminus C$ from it (if contained) yields a finite subcover of C. □

Notice that in $\mathbb{R}^n$ a compact set is automatically closed, but this may not be true for a general topology. It is true if we assume the Hausdorff separability condition:

Definition A.10. A topological space S is called Hausdorff if for any two distinct points x and y of S one has an open neighborhood U of x and an open neighborhood V of y with $U \cap V = \emptyset$.

Notice that $\mathbb{R}^n$ is Hausdorff. Also notice that a subset of a Hausdorff space is automatically Hausdorff (with respect to the induced topology).

Lemma A.11. *A compact subset of a Hausdorff space is closed.*

Proof. Let S be a Hausdorff topological space and let $K \subset S$ be closed. We want to prove that $A := S \setminus K$ is open. For this it is enough to prove that for every $a \in A$ there is an open neighborhood U_a of a entirely contained in A.

Given $a \in A$, for every $x \in K$ we have, by the Hausdorff condition, an open neighborhood U_x of x and and open neighborhood V_x of a with $U_x \cap V_x = \emptyset$. Since $\{U_x \cap K\}_{x \in K}$ is an open cover of K in the induced topology and K is compact, we have a finite subcover $\{U_{x_i} \cap K\}_{i \in I}$. We then define $U_a = \cap_{i \in I} V_{x_i}$, which is open since it is a finite intersection of open sets. It also clearly contains a as a is contained in each V_{x_i}. Finally, let $a' \in U_a$. Then by definition $a' \in V_{x_i}$ for all i and hence $a' \notin U_{x_i}$ for all i, which implies $a' \notin K$; so $U_a \subseteq A$.

□

An important notion, also for manifolds, is that of embedding:

Definition A.12. A continuous map between topological spaces is called an embedding if it is a homeomorphism to its image.

Note that equivalently a map is an embedding if it is continuous, open and injective. A useful criterion, which often applies, is the following.

Lemma A.13. *An injective continuous map from a compact space to a Hausdorff space is an embedding.*

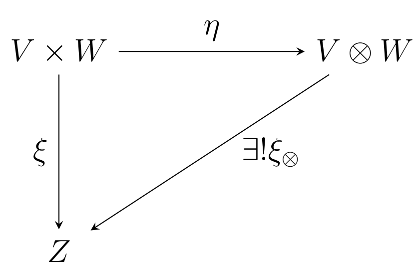
Proof. We have to prove that the map is open or, equivalently passing to complements, that it maps closed subsets to closed subsets. Let $F \colon S \to T$ be the given map with S compact and T Hausdorff. if K is closed in S, then by Lemma A.9 it is also compact. Since F is continuous, by Lemma A.8 $F(K)$ is compact, and by Lemma A.11 it is also closed. Hence, F maps closed sets to closed sets. □

Appendix B. Multilinear algebra

We recall a few basic notions from linear algebra. We consider vector spaces over a ground field $\mathbb{K}$. For the applications in these notes the ground field will be $\mathbb{R}$ and the vector spaces will be finite dimensional.[46]

We begin by recalling that a map $V \times W \to Z$, where V, W and Z are vector spaces, is called bilinear if it is linear with respect to each argument when the other argument is kept fixed. Notice that the set $\mathrm{Bil}(V, W; Z)$ of bilinear maps $V \times W \to Z$ inherits a vector space structure from Z. If $(e_i)_{i\in I}$ is a basis of V and $(f_j)_{j\in J}$ is a basis of W, a blinear map ξ is completely determined by its values $\xi(e_i, f_j)$. This also shows that $\dim \mathrm{Bil}(V, W; Z) = \dim V \dim W \dim Z$. The main idea of the tensor product consists in replacing bilinear maps by linear maps:

Definition B.1. The[47] tensor product of two vector spaces V and W is a pair $(V\otimes W, \eta)$, where $V\otimes W$ is a vector space and $\eta\colon V\times W \to V\otimes W$ is a bilinear map, such that for every vector space Z and every bilinear map $\xi\colon V \times W \to Z$ there is a unique *linear* map $\xi_\otimes\colon V \otimes W \to Z$ such that $\xi = \xi_\otimes \circ \eta$. This property is called the universal property of the tensor product.



Before we show the existence of the tensor product, let us draw some consequences of this definition. First, observe that the association $\xi \mapsto \xi_\otimes$ is linear and has an inverse: to any linear map $\phi\colon V \otimes W \to Z$ we associate the bilinear map $\phi\circ\eta\colon V \times W \to Z$. This shows that we have an isomorphism $\mathrm{Bil}(V, W; Z) \simeq \mathrm{Hom}(V \otimes W, Z)$. In particular,

[46]Unless explicitly stated otherwise, the results in this appendix also hold for infinite dimensional spaces. The proofs are exactly the same if we assume the existence of a basis (which is guaranteed by the axiom of choice). In this case, a sum over an index set is understood to have only finitely many nonvanishing terms.

[47]We are actually defining "a" tensor product, but we will see in Lemma B.2 that all definitions are canonically equivalent.

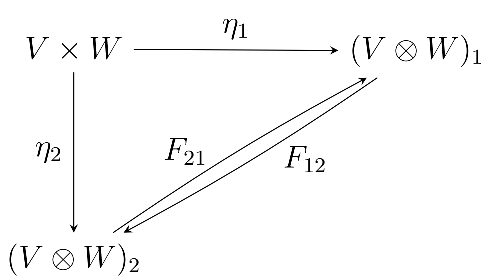
for $Z = \mathbb{K}$ we have $\mathrm{Bil}(V, W; \mathbb{K}) \simeq (V \otimes W)^*$. If V and W are finite dimensional, we finally have

$$V \otimes W \simeq \mathrm{Bil}(V, W; \mathbb{K})^*. \tag{B.1}$$

This is one possible way of constructing the tensor product. The important point is that it does not really matter which construction we use as they are all equivalent:

Lemma B.2. *Suppose $((V \otimes W)_1, \eta_1)$ and $((V \otimes W)_2, \eta_2)$ both satisfy the universal property. Then there is a canonical*[48] *isomorphism $F_{12}\colon (V \otimes W)_1 \to (V \otimes W)_2$ such that $\eta_2 = F_{12}\eta_1$.*

Proof. Since η_2 is a bilinear map, there is a uniquely defined linear map that we denote by F_{12} with the property stated in the Lemma.



We have to prove that it is an isorphism. To do this, we reverse the role of 1 and 2, and get a linear map $F_{21}\colon (V \otimes W)_2 \to (V \otimes W)_2$ such that $\eta_1 = F_{21}\eta_2$. Hence, $\eta_1 = F_{21}F_{12}\eta_1$. This shows that $F_{21}F_{12}$ is the linear map $(V \otimes W)_1 \to (V \otimes W)_1$ corresponding to η_1. By uniqueness we have $F_{21}F_{12} = \mathrm{Id}_1$. Analogously, we prove $F_{12}F_{21} = \mathrm{Id}_2$. □

We now turn to the existence of the tensor product, also for infinite-dimensional vector spaces. As the actual construction does not matter, we may pick one in particular; e.g., using bases.

Lemma B.3. *The tensor product of any two vector spaces V and W exists.*

Proof. Let $(e_i)_{i \in I}$ be a basis of V and $(f_j)_{j \in J}$ a basis of W. Recall that a basis allows identifying vectors with their coefficients. More precisely, let $\mathrm{Map}(I, \mathbb{K})$ denote the vector space of maps[49] $I \to \mathbb{K}$. To a map $i \mapsto v^i$ we associate the vector $\sum_{i \in I} v^i e_i$. Vice versa to a vector $v \in V$ that we expand as $\sum_{i \in I} v^i e_i$ we associate the map $i \mapsto v^i$. Hence, the choice of a basis establishes an isomorphism $\mathrm{Map}(I, \mathbb{K}) \simeq V$. Also

[48] Canonical means that no choice is required to define it.

[49] In the infinite dimensional case we only consider maps that do not vanish at finitely many points.

notice that to the basis element e_r corresponds the map $i \mapsto \delta^i_r$. By abuse of notation, this map is also denoted by e_r and the maps $(e_i)_{i\in I}$ are clearly a basis of $\mathrm{Map}(I, \mathbb{K})$. This suggests defining

$$V \otimes W = \mathrm{Map}(I \times J, \mathbb{K}).$$

To show that this is the correct choice, we only have to define η and prove the universal property. First observe that the maps

$$e_i \otimes f_j \colon (r, s) \mapsto \delta^r_i \delta^s_j,$$

for $i \in I$ and $j \in J$, form a basis of $\mathrm{Map}(I \times J, \mathbb{K})$.

As η is bilinear, it is enough to define it on basis elements. Following the analogy of a single vector space, we set $\eta(e_i, f_j) = e_i \otimes f_j$. Finally, if ξ is a bilinear map $V \times W \to Z$, we define $\xi_\otimes(e_i \otimes f_j) := \xi(e_i, f_j)$ and we immediately see that $\xi = \xi_\otimes \circ \eta$, as it enough to check this identity on basis vectors. On the other hand, $\xi_\otimes$ is uniquely determined. In fact, the difference ϕ of any two maps $\xi_\otimes$ and $\xi'_\otimes$ corresponding to the same ξ, satisfies $\phi \circ \eta = 0$. Applying this to basis vectors, we get $\phi(e_i \otimes f_j) = 0$ for all i, j, and hence ϕ is the zero map. □

Remark B.4. Since $(e_i \otimes f_j)_{i\in I, j\in J}$ is a basis, every vector z of $V \otimes W$ can be written as

$$z = \sum_{i\in I} \sum_{j\in J} z^{ij}\, e_i \otimes f_j$$

for uniquely determined scalars z^{ij}. Notice that in this representation the components of the vector z have two indices.

It is customary to denote with $v \otimes w$ the value of η on (v, w):

$$v \otimes w := \eta(v, w).$$

Vectors in $V \otimes W$ are usually called tensors.[50] Tensors of the form $v \otimes w$ (i.e., tensors in the image of η) are called pure tensors. With this notation, the universal property reads more clearly as

$$\xi_\otimes(v \otimes w) = \xi(v, w)$$

[50]Vectors owe their name to the fact that they were originally introduced to define actual displacements: vector in Latin means carrier. Tensors owe their name to the fact that they were originally introduced to describe tensions in an elastic material as linear relations, i.e. matrices, between the vectors that describe internal forces and deformations.

for all $v \in V$ and all $w \in W$. The fact that η is a bilinear map is encoded in this new notation by the formulae

$$(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w, \tag{B.2a}$$

$$v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2, \tag{B.2b}$$

$$(\lambda v) \otimes w = v \otimes (\lambda w) = \lambda\, v \otimes w, \tag{B.2c}$$

for all $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$ and $\lambda \in \mathbb{K}$.

Remark B.5. These formulae lead to yet another construction of the tensor product. Namely, one considers the free vector space generated by the elements of $V \times W$, writing $v \otimes w$ instead of $(v, w) \in V \times W$, and imposes the formulae (B.2) (i.e., one quotients by the subspace generated by them).[51] The advantage of this construction is that it does not require introducing bases (so it does not need the axiom of choice). See, e.g., [5, paragraph 2.1] for more details.

Notice that any linear map on $V \otimes W$ is completely determined by its values on all pure tensors $v \otimes w$ as this in particular entails evaluation on the basis vectors $(e_i \otimes f_j)_{i\in I, j\in J}$ (or, more abstractly, since pure tensors are the image of η and a linear map $\xi_\otimes$ is completely determined by the bilinear map $\xi = \xi_\otimes \circ \eta$). This also means that to define a map on $V \otimes W$ we can specify it on all pure tensors $v \otimes w$ and check that it is compatible with (B.2). For example, we have a canonical isomorphism

$$\begin{array}{ccc} V \otimes W & \xrightarrow{\sim} & W \otimes V \\ v \otimes w & \mapsto & w \otimes v \end{array}$$

and a canonical isomorphism

$$\begin{array}{ccc} V \otimes \mathbb{K} & \xrightarrow{\sim} & V \\ v \otimes \lambda & \mapsto & \lambda v \end{array}$$

with inverse $V \to V \otimes \mathbb{K}$, $v \mapsto v \otimes 1$. (Notice that λv is mapped to $(\lambda v) \otimes 1$ which is however the same as $v \otimes \lambda$.) If we have a third vector space Z, then we have a canonical isomorphism

$$\begin{array}{ccc} (V \otimes W) \otimes Z & \xrightarrow{\sim} & V \otimes (W \otimes Z) \\ (v \otimes w) \otimes z & \mapsto & v \otimes (w \otimes z) \end{array}$$

[51] More neatly, one quotients $\operatorname{span}(V \times W)$ by the relations

$$\begin{aligned} ((v_1 + v_2), w) &= (v_1, w) + (v_2, w), \\ (v, (w_1 + w_2)) &= (v, w_1) + (v, w_2), \\ (\lambda v, w) &= (v, \lambda w) = \lambda\,(v, w), \end{aligned}$$

and denotes the equivalence class of (v, w) by $v \otimes w$.

For this reason one usually writes $V \otimes W \otimes Z$ without bracketing. One also says that the tensor product of vector spaces is associative. Another useful map is the canonical inclusion

$$\begin{array}{ccc} V^* \otimes W & \hookrightarrow & \mathrm{Hom}(V,W) \\ \alpha \otimes w & \mapsto & (v \mapsto \alpha(v)w) \end{array}$$

To see that it is injective observe that, if $\alpha(v)w = 0$ for all $v \in V$, then $w = 0$ or $\alpha = 0$, and in either case $\alpha \otimes w = 0$. If V and W are finite dimensional, then this is also an isomorphism since

$$\dim(V^* \otimes W) = \dim V \dim W = \dim \mathrm{Hom}(V,W).$$

If we choose a basis $(e_i)_{i\in I}$ of V, a basis $(f_j)_{J\in J}$ of W, and denote by $(e^i)_{i\in I}$ the dual basis of V^*, then a vector A in $V^*\otimes W$ can be expanded as

$$A = \sum_{i\in I}\sum_{j\in J} A_i^j\, e^i \otimes f_j.$$

The coefficients A_i^j are also the components of the matrix that represents the corresponding linear map on right hand side:

$$e_i \mapsto \sum_{j\in J} A_i^j f_j.$$

Similarly, we have a canonical inclusion

$$\begin{array}{ccc} V^* \otimes W^* & \hookrightarrow & (V \otimes W)^* \\ \alpha \otimes \beta & \mapsto & (v \otimes w \mapsto \alpha(v)\beta(w)) \end{array}$$

which is an isomorphism if V and W are finite dimensional. Moreover, (B.1) shows that, if V is finite dimensional, then $V^* \otimes V^*$ is canonicaly isomorphic to the space $\mathrm{Bil}(V,V;\mathbb{K})$ of bilinear forms on V.

If we have linear maps $\phi\colon V \to V'$ and $\psi\colon W \to W'$, then we canonically have a linear map

$$\begin{array}{cccc} \phi\otimes\psi\colon & V \otimes W & \to & V' \otimes W' \\ & v \otimes w & \mapsto & \phi(v) \otimes \psi(w) \end{array}$$

If we have bases $(e_i)_{i\in I}$ of V, $(f_j)_{j\in J}$ of W, $(e'_{i'})_{i'\in I'}$ of V' and $(f'_{j'})_{j'\in J'}$ of W', we may represent the maps ϕ and ψ by matrices: $\phi(e_i) = \sum_{i'\in I'} \phi_i^{i'} e'_{i'}$ and $\psi(f_j) = \sum_{j'\in J'} \psi_j^{j'} f'_{j'}$. It follows that

$$\phi\otimes\psi(e_i \otimes f_j) = \sum_{i'\in I'}\sum_{j'\in J'} \phi_i^{i'} \psi_j^{j'}\, e'_{i'} \otimes f'_{j'}.$$

B.1. **Tensor powers.** Let V be a vector space. Its kth tensor power is by definition

$$V^{\otimes k} = V \otimes \cdots \otimes V$$

where we have k copies of V on the right hand side. The definition is actually by induction:

$$V^{\otimes 1} := V \quad \text{and} \quad V^{\otimes (k+1)} := V^{\otimes k} \otimes V.$$

As the tensor product of tensor spaces is associative the bracketing is not important. By convention one then also sets

$$V^{\otimes 0} := \mathbb{K}.$$

Observe that

$$\dim V^{\otimes k} = (\dim V)^k.$$

An element of $V^{\otimes k}$ is called a tensor of order k. If we pick a basis $(e_i)_{i \in I}$ on V, then $(e_{i_1} \otimes \cdots \otimes e_{i_k})_{i_1,\dots,i_k \in I}$ is a basis of $V^{\otimes k}$ and a tensor T of order k may be uniquely written as

$$T = \sum_{i_1,\dots,i_k \in I} T^{i_1 \cdots i_k}\, e_{i_1} \otimes \cdots \otimes e_{i_k}.$$

Moreover, we have $V^{\otimes k_1} \otimes V^{\otimes k_2} = V^{\otimes (k_1+k_2)}$ for all k_1, k_2. (We write equal instead of isomorphic, as the isomorphism is canonical.) This corresponds to a bilinear map

$$\begin{array}{cccc} \otimes\colon & V^{\otimes k_1} \times V^{\otimes k_2} & \to & V^{\otimes (k_1+k_2)} \\ & (v_1 \otimes \cdots \otimes v_{k_1}, w_1 \otimes \cdots \otimes w_{k_2}) & \mapsto & v_1 \otimes \cdots \otimes v_{k_1} \otimes w_1 \otimes \cdots \otimes w_{k_2} \end{array}$$

called the tensor product of tensors. It is clearly associative: namely,

$$(T_1 \otimes T_2) \otimes T_3 = T_1 \otimes (T_2 \otimes T_3)$$

for all $T_i \in V^{\otimes k_i}$ and any choice of k_i. Usually one then omits bracketing. One also extends the tensors product to scalars. Namely, if $a \in V^{\otimes 0} = \mathbb{K}$ and $\alpha \in V^{\otimes k}$, one defines $a \otimes \alpha := a\alpha =: \alpha \otimes a$. Notice that $1 \in \mathbb{K}$ is then a unit: $1 \otimes \alpha = \alpha \otimes 1 = \alpha$ for all α.

If we pick a basis, then the components of a tensor product of tensors are just the products of the components of the two factors. The tensor product of tensors then makes

$$T(V) := \bigoplus_{k=0}^{\infty} V^{\otimes k}$$

into an associative algebra called the tensor algebra of V.[52] An element of $T(V)$ is sometimes called a nonhomogenous tensor, but often just a tensor. Elements of a single $V^{\otimes k}$ are also called homogenous tensors.

A linear map $\phi\colon V \to W$ canonically induces linear maps

$$\text{(B.3)} \qquad \begin{array}{rccc} \phi^{\otimes k}\colon & V^{\otimes k} & \to & W^{\otimes k} \\ & v_1 \otimes \cdots \otimes v_k & \mapsto & \phi(v_1) \otimes \cdots \otimes \phi(v_k) \end{array}$$

for all k. Notice that if T_1 and T_2 are in $V^{\otimes k_1}$ and $V^{\otimes k_2}$, then

$$\phi^{\otimes(k_1+k_2)}(T_1 \otimes T_2) = \phi^{\otimes k_1}(T_1) \otimes \phi^{\otimes k_2}(T_2).$$

This construction may be repeated with the dual space V^* of V. More generally, one considers the tensor product

$$T^k_s(V) := V^{\otimes k} \otimes (V^*)^{\otimes s}$$

An element of $T^k_s(V)$ is called a tensor of type (k, s). Tensors of type $(0, s)$ are also called covariant tensors of order s, whereas tensors of type $(k, 0)$ are also known as contravariant tensors of order k.[53] As the notation suggests, by convention we put the linear forms to the right. Hence, if we pick a basis $(e_i)_{i\in I}$ on V, then

$$(e_{i_1} \otimes \cdots \otimes e_{i_k} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s})_{i_1,\dots,i_k,j_1,\dots j_s \in I}$$

is a basis of $T^k_s(V)$, where $(e^j)_{j\in I}$ denotes the dual basis. A tensor T of type (k, s) can then be uniquely written as

$$T = \sum_{i_1,\dots,i_k,j_1,\dots j_s \in I} T^{i_1\cdots i_k}_{j_1\cdots j_s}\, e_{i_1} \otimes \cdots \otimes e_{i_k} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s}.$$

Remark B.6. Particularly important are the tensor spaces $T^1_1(V)$ and $T^0_2(V)$ for V finite dimensional. In this case, $T^1_1(V)$ is canonically identified with the space of endomorphisms of V. In a basis we write $F \in T^1_1(V)$ as

$$F = \sum_{i,j\in I} F^i_j\, e_i \otimes e^j.$$

The coefficients F^i_j are also the entries of the matrix representing the corresponding endomorphism, which we keep denoting by F:

$$F(e_j) = \sum_{i\in I} F^i_j\, e_i.$$

[52] The T in $T(V)$ stands for "tensor" and should not be confused with the T denoting the tangent bundle of a manifold.

[53] This terminology refers to the fact that, if we change basis by some matrix, the components of a vector change by application of the inverse matrix (hence the name *contra*variant), whereas the components of a linear form change by the application of the matrix itself (hence the name *co*variant).

The tensor space $T^0_2(V)$ is instead canonically identified with the space of bilinear forms on V. In a basis we write $B \in T^0_2(V)$ as

$$B = \sum_{i,j\in I} B_{ij}\, e^i \otimes e^j.$$

The coefficients B_{ij} are also the entries of the matrix representing the corresponding bilinear form, which we keep denoting by B:

$$B(e_i, e_j) = B_{ij}.$$

Remark B.7 (Einstein's convention)**.** A useful habit, which we implicitly used above, consists in taking care of the position of the indices. We use lower indices to denote basis vectors (e_i) and upper indices to denote the components v^i in the expansion of a vector

$$v = \sum_i v^i e_i.$$

For the dual basis we use the same letters as for the basis but with upper indices: (e^i). For the components of a linear form we then use lower indices:

$$\omega = \sum_i \omega_i e^i.$$

Consequently a vector in T^k_s will have k upper and s lower indices. This notation allows recognizing at a glance the type of a tensor. A further convention, due to Einstein, tacitly assumes a summation over every repeated index, once in the upper and once in the lower position. For example, with this convention the expansion of a vector v and of a linear form ω read $v = v^i e_i$ and $\omega = \omega_i e^i$. This very useful convention requires some training. As in these notes we only occasionally work with coordinates, we prefer not to make use of it.

A tensor of type (k, s) may be written, by definition, as a linear combination of tensors of the form $T \otimes S$ where T is of type $(k, 0)$ and S is of type $(0, s)$. The tensor product of tensors extends to the general case by

$$\begin{array}{ccc} T^{k_1}_{s_1}(V) \otimes T^{k_2}_{s_2}(V) & \to & T^{k_1+k_2}_{s_1+s_2}(V) \\ (T_1 \otimes S_1) \otimes (T_2 \otimes S_2) & \mapsto & T_1 \otimes T_2 \otimes S_1 \otimes S_2 \end{array}$$

Similarly, an isomorphism $\phi\colon V \to W$ induces canonically isomorphisms[54]

$$\text{(B.4)} \qquad \begin{array}{rccc} \phi^k_s\colon & T^k_s(V) & \to & T^k_s(W) \\ & T \otimes S & \mapsto & \phi^{\otimes k}(T) \otimes ((\phi^*)^{-1})^{\otimes s}(S) \end{array}$$

[54]In this case, ϕ must be an isomorphism because we have to define an associated map $V^* \to W^*$.

for all k, s. Again, if $U_1 \in T^{k_1}_{s_1}(V)$ and $U_2 \in T^{k_2}_{s_2}(V)$, we have

$$\phi^{k_1+k_2}_{s_1+s_2}(U_1 \otimes U_2) = \phi^{k_1}_{s_1}(U_1) \otimes \phi^{k_2}_{s_2}(U_2).$$

Finally, observe that the pairing $V \otimes V^* \to \mathbb{K}$, $(v, \alpha) \mapsto \alpha(v)$ canonically induces linear maps

$$I^m_n \colon T^k_s(V) \to T^{k-1}_{s-1}(W),$$

for all $1 \leq m \leq k$ and $1 \leq n \leq s$, obtained by pairing the mth vector with the nth linear form in the tensor. These linear maps are called contractions.

B.2. **Exterior algebra.** For applications in the theory of manifolds (viz., differential forms), we also need the concept of exterior algebra. For simplicity, we develop it in the case when the ground field $\mathbb{K}$ has characteristic zero (e.g., $\mathbb{K} = \mathbb{R}$).[55] The objects of interest are then the skew-symmetric tensors. In terms of a basis, these are the tensors whose components are skew-symmetric with respect to the exchange of indices.

More invariantly, we proceed as follows. First observe that a permutation σ on k elements defines an endomorphism of $V^{\otimes k}$ given by

$$v_1 \otimes \cdots \otimes v_k \mapsto v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)}.$$

on pure tensors. We denote the so defined endomorphism also by σ. In particular, if (e_i) is a basis and $\alpha = \sum_{i_1,\dots,i_k} \alpha^{i_1\dots i_k} e_{i_1} \otimes \cdots \otimes e_{i_k}$ is a k-tensor, then

$$\text{(B.5)} \quad \sigma\alpha = \sum_{i_1,\dots,i_k} \alpha^{i_1\dots i_k} e_{i_{\sigma(1)}} \otimes \cdots \otimes e_{i_{\sigma(k)}} = \\ = \sum_{i_1,\dots,i_k} \alpha^{i_{\sigma^{-1}(1)}\dots i_{\sigma^{-1}(k)}} e_{i_1} \otimes \cdots \otimes e_{i_k}.$$

Notice that this defines a representation of the symmetric group S_k (i.e., the group of permutations over k elements) on $V^{\otimes k}$: namely,

$$(\sigma_1\sigma_2)\alpha = \sigma_1(\sigma_2\alpha) \quad \text{and} \quad \mathrm{Id}\,\alpha = \alpha$$

for all $\sigma_1, \sigma_2 \in S_k$ and $\alpha \in V^{\otimes k}$ (we denote by Id the identity permutation).

Since we are interested in skew-symmetric tensors, we twist this representation by the sign:[56] a k-tensor is called skew-symmetric if

$$\sigma\alpha = \mathrm{sgn}\sigma\,\alpha$$

[55]Fot the general case, see subsection B.2.2.

[56]A parallel discussion, without this twist, leads to the symmetric algebra.

for all $\sigma \in S_k$. Since the symmetric group is generated by transpositions, we also have that α is skew-symmetric if and only if $\tau\alpha = -\alpha$ for every transposition τ. We denote by $\Lambda^k V$ the vector space of skew-symmetric k-tensors.

If we expand α in a basis, we see that α is skew-symmetric if and only if its components change sign by the exchange of any two indices. More generally, by (B.5), we see that α is skew-symmetric if and only if

$$\alpha^{i_{\sigma(1)}\cdots i_{\sigma(k)}} = \mathrm{sgn}\sigma\, \alpha^{i_1\cdots i_k} \tag{B.6}$$

for all σ and all $i_1, \dots, i_k$.

Notice that the map $\phi^{\otimes k}$ defined in equation (B.3) commutes with the action of the permutation group:

$$\phi^{\otimes k}\sigma = \sigma\phi^{\otimes k}$$

for all $\sigma \in S_k$. This implies that $\phi^{\otimes k}$ maps skew-symmetric tensors to skew-symmetric tensors. The restriction of $\phi^{\otimes k}$ to $\Lambda^k V$ is usually denoted by $\Lambda^k\phi$. In summary, a linear map $\phi\colon V \to W$ canonically induces linear maps

$$\Lambda^k\phi\colon \Lambda^k V \to \Lambda^k W$$

for all k.

The tensor product of two skew-symmetric tensors is in general no longer skew-symmetric. However, one can always skew-symmetrize it and define the wedge product of $\alpha_1 \in \Lambda^{k_1}V$ and $\alpha_2 \in \Lambda^{k_2}V$ by

$$\alpha_1 \wedge \alpha_2 := \mathrm{Alt}^k(\alpha_1 \otimes \alpha_2), \tag{B.7}$$

with $k = k_1 + k_2$, where

$$\mathrm{Alt}^k\alpha := \frac{1}{k!}\sum_{\sigma\in S_k} \mathrm{sgn}\sigma\, \sigma\alpha.$$

Lemma B.8. *The alternating map Alt^k has image equal to $\Lambda^k V$. Moreover, if $\alpha \in \Lambda^k V$, then $\mathrm{Alt}^k\alpha = \alpha$.*

Proof. For $\tau \in S_k$, let us compute

$$\tau\,\mathrm{Alt}^k\alpha = \frac{1}{k!}\sum_{\sigma\in S_k} \mathrm{sgn}\sigma\, \tau\sigma\alpha = \mathrm{sgn}\tau\frac{1}{k!}\sum_{\sigma\in S_k} \mathrm{sgn}(\tau\sigma)\, \tau\sigma\alpha.$$

By the change of variable $\widehat{\sigma} = \tau\sigma$ we then get

$$\tau\,\mathrm{Alt}^k\,\alpha = \mathrm{sgn}\tau\frac{1}{k!}\sum_{\widehat{\sigma}\in S_k} \mathrm{sgn}\widehat{\sigma}\,\widehat{\sigma}\alpha = \mathrm{sgn}\tau\,\mathrm{Alt}^k\,\alpha,$$

for all $\tau \in S_k$, which proves that the image of Alt^k is in $\Lambda^k V$.

We then move to the second statement. From $\sigma\alpha = \operatorname{sgn}\sigma\, \alpha$ we get $\mathrm{Alt}^k(\alpha) := \frac{1}{k!}\sum_{\sigma\in S_k} \alpha = \alpha$. This also proves that the image of Alt^k is the whole of $\Lambda^k V$. □

If ϕ is a linear map as above, then we clearly have

$$\Lambda^k\phi(\alpha_1 \wedge \alpha_2) = (\Lambda^{k_1}\alpha_1) \wedge (\Lambda^{k_2}\alpha_2).$$

We extend the wedge product to the direct sum $\Lambda V := \bigoplus_{k=0}^{\infty} \Lambda^k V$.

Lemma B.9. $(\Lambda V, \wedge)$ *is an associative algebra with unit* $1 \in \Lambda^0 V = \mathbb{K}$.

This algebra is called the exterior algebra of V.

Proof. We compute

$$(\alpha_1 \wedge \alpha_2) \wedge \alpha_3 = \left(\frac{1}{(k_1+k_2)!}\sum_{\sigma\in S_{k_1+k_2}} \operatorname{sgn}\sigma\, \sigma(\alpha_1\otimes\alpha_2)\right) \wedge \alpha_3 =$$

$$= \frac{1}{(k_1+k_2+k_3)!(k_1+k_2)!}\sum_{\substack{\widetilde{\sigma}\in S_{k_1+k_2+k_3}\\ \sigma\in S_{k_1+k_2}}} \operatorname{sgn}\widetilde{\sigma}\operatorname{sgn}\sigma\, \widetilde{\sigma}(\sigma(\alpha_1\otimes\alpha_2)\otimes\alpha_3).$$

Let $\sigma\times\mathrm{Id}_{k_3}$ be the permutation over $k_1+k_2+k_3$ elements that is the identity on the last k_3 element and σ on the first k_1+k_2 elements. Then $\sigma(\alpha_1\otimes\alpha_2)\otimes\alpha_3 = (\sigma\times\mathrm{Id}_{k_3})(\alpha_1\otimes\alpha_2\otimes\alpha_3)$. Notice that $\operatorname{sgn}\sigma = \operatorname{sgn}(\sigma\times\mathrm{Id}_{k_3})$. We then make the change of variable $\widetilde{\sigma}\mapsto\widehat{\sigma} = \widetilde{\sigma}(\sigma\times\mathrm{Id}_{k_3})$ and get

$$(\alpha_1\wedge\alpha_2)\wedge\alpha_3 = \frac{1}{(k_1+k_2+k_3)!(k_1+k_2)!}\sum_{\substack{\widehat{\sigma}\in S_{k_1+k_2+k_3}\\ \sigma\in S_{k_1+k_2}}} \operatorname{sgn}\widehat{\sigma}\, \widehat{\sigma}(\alpha_1\otimes\alpha_2\otimes\alpha_3).$$

If we perform the sum over σ, we finally obtain

$$(\alpha_1 \wedge \alpha_2) \wedge \alpha_3 = \frac{1}{(k_1+k_2+k_3)!}\sum_{\widehat{\sigma}\in S_{k_1+k_2+k_3}} \operatorname{sgn}\widehat{\sigma}\, \widehat{\sigma}(\alpha_1\otimes\alpha_2\otimes\alpha_3).$$

By an analogous computation, one sees that this is also the expression for $\alpha_1 \wedge (\alpha_2 \wedge \alpha_3)$.

We next check that $1 \in \mathbb{K}$ is the unit. Since $1\otimes\alpha = \alpha$ and $\alpha \in \Lambda^k V$ is skew-symmetric, we have

$$1\wedge\alpha = \frac{1}{k!}\sum_{\sigma\in S_k} \operatorname{sgn}\sigma\, \sigma(1\otimes\alpha) = \frac{1}{k!}\sum_{\sigma\in S_k} \operatorname{sgn}\sigma\, \sigma\alpha = \frac{1}{k!}\sum_{\sigma\in S_k} \alpha = \alpha.$$

Similarly, one sees that $\alpha\wedge 1 = \alpha$.

Notice that dividing by the order of the group of permutations in the definition of the wedge product is fundamental for this Lemma to hold. □

Remark B.10. By induction, using the first part of this proof, one can also prove that for $\alpha_i \in \Lambda^{k_i}V$, $i = 1, \dots, r$, we have

$$\alpha_1 \wedge \cdots \wedge \alpha_r = \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}\sigma\, \sigma(\alpha_1 \otimes \cdots \otimes \alpha_r).$$

with $k = \sum_{i=1}^r k_i$.

Lemma B.11. *The wedge product is graded commutative, i.e.,*

$$\alpha_2 \wedge \alpha_1 = (-1)^{k_1 k_2} \alpha_1 \wedge \alpha_2$$

for all $\alpha_1 \in \Lambda^{k_1}V$ and $\alpha_2 \in \Lambda^{k_2}V$. In particular, $\alpha \wedge \alpha = 0$ if $\alpha \in \Lambda^k V$ with k odd.

Proof. Let $\tau \in S_k$, $k = k_1 + k_2$, denote the permutation that exchanges the first k_1 elements with the last k_2 elements. We have $\alpha_2 \otimes \alpha_1 = \tau(\alpha_1 \otimes \alpha_2)$. Then, by (B.7), we have

$$\alpha_2 \wedge \alpha_1 = \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}\sigma\, \sigma\tau(\alpha_1 \otimes \alpha_2).$$

By the change of variables $\sigma \mapsto \widehat{\sigma} = \sigma\tau$, we get

$$\alpha_2 \wedge \alpha_1 = \mathrm{sgn}\tau \frac{1}{k!} \sum_{\widehat{\sigma} \in S_k} \mathrm{sgn}\widehat{\sigma}\, \widehat{\sigma}(\alpha_1 \otimes \alpha_2).$$

This completes the proof since $\mathrm{sgn}\tau = (-1)^{k_1 k_2}$. □

Lemma B.12. *If $(e_i)_{i \in I}$ is a basis of V, then $(e_{j_1} \wedge \cdots \wedge e_{j_k})_{j_1 < \cdots < j_k \in I}$ is a basis of $\Lambda^k V$.*

Proof. We expand $\alpha \in \Lambda^k V \subset V^{\otimes k}$ as $\alpha = \sum_{i_1, \dots, i_k} \alpha^{i_1 \dots i_k} e_{i_1} \otimes \cdots \otimes e_{i_k}$. Since $\alpha = \mathrm{sgn}\sigma\, \sigma\alpha$ for all σ, we can also write $\alpha = \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}\sigma\, \sigma\alpha$. We then get, by Remark B.10,

$$\alpha = \sum_{i_1, \dots, i_k \in I} \alpha^{i_1 \dots i_k}\, e_{i_1} \wedge \cdots \wedge e_{i_k},$$

which shows that $(e_{i_1} \wedge \cdots \wedge e_{i_k})_{i_1, \dots, i_k \in I}$ is a system of generators for $\Lambda^k V$.

These generators are however linearly dependent. By the graded commutativity we have $e_i \wedge e_i = 0$ for all i. This implies that, if an index is repeated, then $e_{i_1} \wedge \cdots \wedge e_{i_k} = 0$, since we can use the graded commutativity to move the two e_i's with the same index next to each other. If all the indices are different from each other, then there is a unique permutation σ such that $i_{\sigma(1)} < i_{\sigma(2)} < \cdots < i_{\sigma(k)}$. We can then write

$$\alpha^{i_1\dots i_k} e_{i_1} \wedge \cdots \wedge e_{i_k} = \alpha^{i_1\dots i_k} \operatorname{sgn}\sigma\, e_{i_{\sigma(1)}} \wedge \cdots \wedge e_{i_{\sigma(k)}} = \\ = \alpha^{i_{\sigma(1)}\dots i_{\sigma(k)}} e_{i_{\sigma(1)}} \wedge \cdots \wedge e_{i_{\sigma(k)}} = \alpha^{j_1\dots j_k} e_{j_1} \wedge \cdots \wedge e_{j_k}$$

where we have used (B.6) and have set $j_r = i_{\sigma(r)}$, $r = 1, \dots, k$. By construction we have $j_1 < j_2 < \cdots < j_k$. Notice that if we fix the ordered j_r's there are $k!$ corresponding unordered i_r's. Hence

$$\alpha = \sum_{j_1 < \cdots < j_k \in I} k!\, \alpha^{j_1\dots j_k}\, e_{j_1} \wedge \cdots \wedge e_{j_k},$$

which shows that $(e_{j_1} \wedge \cdots \wedge e_{j_k})_{j_1 < \cdots < j_k \in I}$ is also a system of generators for $\Lambda^k V$. We want to prove that they are linearly independent. Let $\lambda^{j_1\dots j_k}$ be a collection of scalars for $j_1 < \cdots < j_k$ such that $\sum_{j_1 < \cdots < j_k \in I} \lambda^{j_1\dots j_k}\, e_{j_1} \wedge \cdots \wedge e_{j_k} = 0$. For $i_1, \dots, i_k$ pairwise distinct, define $\alpha^{i_1\dots i_k} = \operatorname{sgn}\sigma \lambda^{i_{\sigma(1)}\dots i_{\sigma(k)}}$ where σ is the unique permutation such that $i_{\sigma(1)} < \cdots < i_{\sigma(k)}$; if an index is repeated, we define $\alpha^{i_1\dots i_k} = 0$. We then have $\sum_{i_1,\dots,i_k \in I} \alpha^{i_1\dots i_k}\, e_{i_1} \wedge \cdots \wedge e_{i_k} = 0$. By Remark B.10, this implies $\sum_{i_1,\dots,i_k \in I} \alpha^{i_1\dots i_k}\, e_{i_1} \otimes \cdots \otimes e_{i_k} = 0$. Hence $\alpha^{i_1\dots i_k} = 0$ for all $i_1, \dots, i_k$, which implies $\lambda^{j_1\dots j_k} = 0$ for all $j_1 < \cdots < j_k$. □

This implies that to define a linear map on $\Lambda^k V$ it is enough to define it on pure elements, i.e., elements of the form $v_1 \wedge \cdots \wedge v_k$, checking that it is multilinear and alternating in the vectors $v_1, \dots, v_n$.

Corollary B.13. *If* $\dim V = n$, *then* $\dim \Lambda^k V = \binom{n}{k}$. *In particular,* $\Lambda^k V = \{0\}$ *if* $k > n$.

Observe that $\Lambda^n V$ is one-dimensional if $n = \dim V$. This means, that if ϕ is an endomorphism of V, then $\Lambda^n \phi$ is the multiplication by a scalar. It turns out that this scalar is the determinant of ϕ:

$$\boxed{\Lambda^n \phi\, \alpha = \det \phi\, \alpha} \tag{B.8}$$

for all $\alpha \in \Lambda^n V$.

Proof. Let $(e_1, \dots, e_n)$ be a basis of V. We have

$$\Lambda^n \phi\, e_1 \wedge \cdots \wedge e_n = \phi(e_1) \wedge \cdots \wedge \phi(e_n) = \sum_{i_1,\dots,i_n} \phi_1^{i_1} \cdots \phi_n^{i_n} e_{i_1} \wedge \cdots \wedge e_{i_n},$$

where (ϕ_i^j) is the matrix representing ϕ in this basis. If any index is repeated, the contribution vanishes. If all indices are pairwise different, we let σ be the permutation with $\sigma(j) = i_j$. Then

$$\Lambda^n\phi\, e_1 \wedge \cdots \wedge e_n = \sum_{\sigma\in S_n} \operatorname{sgn}\sigma\, \phi_1^{\sigma(1)} \cdots \phi_n^{\sigma(n)}\, e_1 \wedge \cdots \wedge e_n,$$

which completes the proof by the Leibniz formula for the determinant. □

B.2.1. *Contractions.* The pairing between a vector space and its dual extends to the exterior algebra. We describe its most important appearance. To fit with the application to differential forms we use the exterior algebra of the dual here.

An element of ΛV^* is called a form and an element of $\Lambda^k V^*$ a k-form. A k-form $a_1 \wedge \cdots \wedge a_k$ with $a_i \in V^*$ for all i is called pure. A linear map defined on $\Lambda^k V^*$ is completely determined by its values on the pure forms (as in particular basis elements are pure forms). On the other hand, a map defined on pure forms extends to a linear map if it is multinear and alternating on the pure forms.

A vector v in V defines a linear map $\iota_v\colon \Lambda^k V^* \to \Lambda^{k-1} V^*$ called contraction, for all k, defined on pure forms by

$$\begin{aligned}\iota_v(a_1 \wedge \cdots \wedge a_k) = a_1(v)\, a_2 \wedge \cdots \wedge a_k - a_2(v)\, a_1 \wedge a_3 \wedge \cdots \wedge a_k + \cdots =\\ = \sum_{r=1}^{k} (-1)^{r-1} a_r(v)\, a_1 \wedge \cdots \wedge \widehat{a_r} \wedge \cdots \wedge a_k,\end{aligned}$$

where the caret $\widehat{\ }$ indicates that the factor a_r is omitted. On $\Lambda^0 V^*$ the contraction ι_v is defined as the zero map.

Lemma B.14. *The contraction has the following important properties. First, for all $\alpha \in \Lambda^k V^*$, $\beta \in \Lambda^l V^*$ and $v \in V$, one has*

$$\iota_v(\alpha \wedge \beta) = \iota_v\alpha \wedge \beta + (-1)^k \alpha \wedge \iota_v\beta.$$

Second, for all $v, w \in V$ and $\alpha \in \Lambda V^*$, one has

$$\iota_v\iota_w\alpha = -\iota_w\iota_v\alpha$$

Proof. It is enough to check the first identity when α and β are pure, and this follows immediately from the definition.

The second identity can also be easily checked on pure forms. In fact, one can use the first identity to show that $I_{v,w} := \iota_v\iota_w + \iota_w\iota_v$ satisfies

$$I_{v,w}(\alpha \wedge \beta) = I_{v,w}\alpha \wedge \beta + \alpha \wedge I_{v,w}\beta$$

for all α and β. By induction one then sees that $I_{v,w}$ is determined by its actions on 1-forms. Since $I_{v,w}$ is clearly zero on $\Lambda^1 V^*$, it is then zero on the whole ΛV^*. □

Let finally ϕ be a linear map $V \to W$. Since the the transpose of a linear map is defined exactly so as to preserve the pairing of a vector with a linear form, $(\phi^* a)(v) = a(\phi v)$ for all $a \in W^*$ and $v \in V$, we have

$$\iota_v \Lambda^k \phi^* \alpha = \Lambda^{k-1} \phi^* \iota_{\Phi v} \alpha, \tag{B.9}$$

for all v in V and all $\alpha \in \Lambda^k V^*$.

B.2.2. *The exterior algebra as a quotient.* In the above description of the exterior algebra, we had several denominators, which is ok if the ground field $\mathbb{K}$ has characteristic zero. For a general ground field, one can use another definition of the exterior algebra (which is canonically isomorphic to the previous one if the field has characteristic zero).

To start with, we recall a basic construction. An algebra A is a vector space endowed with a bilinear map $A \times A \to A$, usually denoted by $(a, b) \mapsto ab$. The algebra is called associative if $(ab)c = a(bc)$ for all $a, b, c \in A$. A two-sided ideal of an algebra A is a subspace I with the property that $ax \in I$ and $xa \in I$ for all $a \in A$ and $\forall x \in I$. The quotient space A/I then inherits an associative algebra structure by

$$[a][b] := [ab], \qquad a \in [a], b \in [b].$$

Notice that the class $[ab]$ does not depend on the choice of representatives a and b since I is a two-sided ideal.

We now apply this construction to the algebra $T(V) := \bigoplus_{k=0}^{\infty} V^{\otimes k}$, where V is a vector space on some ground field $\mathbb{K}$, of any characteristic, and the associative algebra structure is defined by the tensor product of tensors. We let I be the two-sided ideal generated by elements of the form $v \otimes v$ with $v \in V$. More explictly I is the span of elements of the form $a \otimes v \otimes v \otimes b$ with $a, b \in T(V)$ and $v \in V$. The exterior algebra ΛV of V is then defined as the quotient algebra $T(V)/I$. The induced associative product is denoted by $\wedge$ and is called the exterior product:

$$[a] \wedge [b] := [a \otimes b], \qquad a \in [a], b \in [b].$$

Notice that I is a graded ideal, i.e., $I = \bigoplus_{k=2}^{\infty} I_k$ with

$$I_k = I \cap V^{\otimes k} =$$
$$= \operatorname{span}\{a \otimes v \otimes v \otimes b : v \in V,\ a \in V^{\otimes k_1},\ b \in V^{\otimes k_2},\ k_1 + k_2 = k - 2\}.$$

One then defines $\Lambda^k V = V^{\otimes k}/I_k$ and one gets $\Lambda V = \bigoplus_{k=0}^{\infty} \Lambda^k V$. Observe that $\Lambda^0 V = \mathbb{K}$ and $\Lambda^1 V = V$.

The kth tensor power $\phi^{\otimes k}$ of a linear map $\phi\colon V \to W$ clearly sends the kth component of the ideal of $T(V)$ to the kth component of the ideal of $T(W)$, so it descends to the quotients. We denote it by $\Lambda^k\phi\colon \Lambda^k V \to \Lambda^k W$.

One can prove that the so defined exterior algebra has the same properties as the one we have defined above in terms of skew-symmetric tensors. In the case of characteristic zero the two constructions are equivalent. Namely, let $A_k\colon V^{\otimes k} \to V^{\otimes k}$ be the map defined by $A_k\alpha = \frac{1}{k!}\sum_{\sigma\in S_k} \operatorname{sgn}\sigma\,\sigma\alpha$. One can see that $I_k = \ker A_k$ and that the image of A_k is the space of skew-symmetric k-tensors. The canonical isomorphism between $T(V)/I$ and $\bigoplus_k A_k(V^{\otimes k})$ is also compatible with the wegde products.

Finally observe that in the general construction the exterior algebra is a quotient of the tensor algebra, whereas in the special construction with skew-symmetric tensors it is a subspace.

References

[1] R. Bott and L. W. Tu, *Differential Forms in Algebraic Topology,* Springer 1982.
[2] W. Fulton, *Algebraic Topology. A First Course,* Springer 1995.
[3] N. Hitchin, *Differentiable Manifolds,* https://people.maths.ox.ac.uk/hitchin/hitchinnotes/manifolds2012.pdf
[4] J. M. Lee, *Introduction to Smooth Manifolds,* Springer 2002.
[5] F. W. Warner, *Foundations of Differentiable Manifolds and Lie Groups,* Springer 1983.

Index

$\mathcal{C}$-map, 8, 17
$\mathcal{C}$-atlas, 3
$\mathcal{C}$-equivalent, 5
$\mathcal{C}$-function, 8
$\mathcal{C}$-isomorphism, 9
$\mathcal{C}$-manifold, 5
$\mathcal{C}$-morphism, 8
$\mathcal{C}$-submanifold, 11
$\mathcal{C}^k$-diffeomorphism, 9
1-form, 77

action, 156
adapted atlas, 11
adapted chart, 11
antipodal map, 117
atlas, 3

basic form, 106
boundary
  in homology, 134
  of a manifold, 126
bump function, 15

chain, 133
chart, 2
chart map, 2
closed form, 97
closed manifold, 126
co-orientable submanifold, 137
compact, 159
constant of motion, 141
continuous, 158
contraction, 77, 94, 169, 174
coordinate, 6
coordinate function, 6
coordinate map, 2
coordinate neighborhood, 6
coordinatized tangent vector, 21
cotangent bundle, 77
cycle, 134

de Rham cohomology, 97
de Rham complex, 96
de Rham differential, 78, 94
density, 63
  on a vector space, 67
derivation, 33, 37
diffeomorphism, 9
differentiable curve, 24
differentiable manifold, 19
differential, 22
differential form, 92
dimension, 7
distribution, 47
divergence of a vector field, 75
dual bundle, 54

embedding
  of differentiable manifolds, 19
  of topological spaces, 160
Euler vector field, 104
evaluation, 33
exact
  1-form, 78
exact form, 97
exponential map, 150
exterior algebra, 171, 175
exterior product, 175

fiber, 25
flow, 30
flux, 119
frame, 67

Gauss theorem, 129
geodesic, 86
germ, 33
graded algebra, 100
graded commutative algebra, 100
graded derivation, 100
graded endomorphism, 99
graded Lie algebra, 100
graded morphism, 99
graded skew-commutative algebra, 100
graded subspace, 101
graded vector space, 99
gradient, 88

Hamiltonian flow, 139
Hamiltonian system, 141
Hamiltonian vector field, 139
Hausdorff, 160
homeomorphism, 158
homotopy, 131
horizontal form, 106

immersion, 19, 23
implicit function theorem, 7, 25
induced topology, 158
infinitesimal action, 156
infinitesimal symmetry, 141
integrable distribution, 49
integral curve, 28
integral manifold, 49
invariant form, 106
involution, 116
isomorphism, 57

Jacobi identity, 39

Laplacian, 88
left invariant, 147
Lie algebra, 39
Lie bracket, 39
Lie derivative
  of a 1-form, 79, 80
  of a density, 72
  of a differential form, 93
  of a function, 43
  of a tensor field, 82
  of a vector field, 43
Lie group, 145
Lie subalgebra, 39
Lie subgroup, 154
line bundle, 62
line with two origins, 14, 15
Liouville 1-form, 139
Liouville volume form, 140
local sections, 56
locally finite, 18
Lorentzian metric, 91

manifold with boundary, 124, 125
maximal atlas, 5
morphism, 57

neighborhood, 158

open, 158
open atlas, 3
open cover, 158
open sets, 158
orientable manifold, 108
orientation, 109
orientation bundle, 135
orientation-preserving diffeomorphism, 110
orientation-reversing diffeomorphism, 110
oriented, 112
oriented atlas, 112
oriented manifold, 109
outward pointing vector, 122, 128

pairing, 56
parallelizable manifold, 59
partition of unity, 18
Poincaré 1-form, 139
points, 158
Poisson bracket, 141
projectable vector field, 52
pseudo-Riemannian metric, 90
pullback
  of a 1-form, 77
  of a covariant tensor field, 84
  of a density, 69, 70
  of a function, 10
  of a vector bundle, 55
pure tensors, 163
push-forward
  of a 1-form, 77
  of a density, 71
  of a distribution, 49
  of a function, 10, 41
  of a measure, 71
  of a section, 57
  of a tensor field, 82
  of a vector field, 27, 41

quotient topology, 159

rank
  of a distribution, 47
  of a vector bundle, 53
real projective space, 117
relative topology, 158
Riemannian density, 87
Riemannian manifold, 84
Riemannian metric, 84

second countable space, 18
section, 55
simplex, 131
singular homology, 134
smooth manifold, 19
smooth maps, 19
standard $\mathcal{C}$-manifold structure, 6
standard atlas, 3

standard density, 70
standard orientation, 111
standard topology, 158
standard volume form, 111
Stokes theorem
  for chains, 134
  for manifolds with boundary, 128
submanifold, 11
submersion, 19, 23
support, 15
symplectic form, 138
symplectic manifold, 138
symplectic reduction, 145

tangent bundle, 25
tangent lift, 26, 57
tangent space, 21
tangent vector, 21
tangential coordinates, 11
tautological 1-form, 139
tensor, 166
tensor algebra, 167
tensor field, 81
tensor power, 166
tensor product, 161
tensor product of tensors, 166
tensors, 163
top form, 94, 108
topological manifold, 13
topological space, 158
topology, 158
transition map, 3
transversal coordinates, 11
trivial vector bundle, 58
trivializing atlas, 53
twisted differential form, 136
type
  of a tensor, 167
  of a tensor field, 81

vector bundle, 53
vector bundle map, 57
vector field, 26
  complete, 31
vector subbundle, 56
velocity, 24
vertical vector field, 52, 143
volume form, 108
volume-preserving vector field, 73

wedge product, 170
winding number, 130

Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland
E-mail address: cattaneo @math.uzh.ch