Differentiable manifolds

Lecture Notes for Geometry 2

Henrik Schlichtkrull

Department of Mathematics
University of Copenhagen

Preface

The purpose of these notes is to introduce and study *differentiable manifolds*. Differentiable manifolds are the central objects in differential geometry, and they generalize to higher dimensions the curves and surfaces known from Geometry 1. Together with the manifolds, important associated objects are introduced, such as tangent spaces and smooth maps. Finally the theory of differentiation and integration is developed on manifolds, leading up to Stokes' theorem, which is the generalization to manifolds of the fundamental theorem of calculus.

These notes continue the notes for Geometry 1, about curves and surfaces. As in those notes, the figures are made with Anders Thorup's spline macros. The notes are adapted to the structure of the course, which stretches over 9 weeks. There are 9 chapters, each of a size that it should be possible to cover in one week. The notes were used for the first time in 2006. The present version has been revised, but further revision is undoubtedly needed. Comments and corrections will be appreciated.

Henrik Schlichtkrull
December, 2008

Contents

Chapter 1

Manifolds in Euclidean space

In Geometry 1 we have dealt with parametrized curves and surfaces in $\mathbb{R}^2$ or $\mathbb{R}^3$. The definitions we have seen for the two notions are analogous to each other, and we shall begin by generalizing them to arbitrary dimensions. As a result we obtain the notion of a parametrized m-dimensional manifold in $\mathbb{R}^n$.

The study of curves and surfaces in Geometry 1 was mainly through parametrizations. However, as it was explained, important examples of curves and surfaces arise more naturally as level sets, for example the circle $\{(x, y) \mid x^2 + y^2 = 1\}$ and the sphere $\{(x, y, z) \mid x^2 + y^2 + z^2 = 1\}$. In order to deal with such sets, we shall define a notion of manifolds, which applies to *subsets* in $\mathbb{R}^n$ without the specification of a particular parametrization. The new notion will take into account the possibility that the given subset of $\mathbb{R}^n$ is not covered by a single parametrization. It is easy to give examples of subsets of $\mathbb{R}^3$ that we conceive as surfaces, but whose natural parametrizations do not cover the entire set (at least if we require the parametrizations to be regular).

For example, we have seen that for the standard spherical coordinates on the sphere there are two singular points, the poles. In order to have a regular parametrization we must exclude these points. A variation of the standard spherical coordinates with interchanged roles of y and z will have singular poles in two other points. The entire sphere can thus be covered by spherical coordinates if we allow two parametrizations covering different, overlapping subsets of the sphere. Note that in contrast, the standard parametrization of the circle by trigonometric coordinates is everywhere regular.

1.1 Parametrized manifolds

In the following m and n are arbitrary non-negative integers with $m \leq n$.

Definition 1.1.1. A *parametrized manifold* in $\mathbb{R}^n$ is a smooth map $\sigma\colon U \to \mathbb{R}^n$, where $U \subset \mathbb{R}^m$ is a non-empty open set. It is called *regular at* $x \in U$ if the $n \times m$ Jacobi matrix $D\sigma(x)$ has rank m (that is, it has linearly independent columns), and it is called *regular* if this is the case at all $x \in U$. An *m-dimensional parametrized manifold* is a parametrized manifold $\sigma\colon U \to \mathbb{R}^n$ with $U \subset \mathbb{R}^m$, which is regular (that is, regularity is implied at all points when we speak of the dimension).

Clearly, a parametrized manifold with $m = 2$ and $n = 3$ is the same as a parametrized surface, and the notion of regularity is identical to the one introduced in Geometry 1. For $m = 1$ there is a slight difference with the notion of parametrized curves, because in Geometry 1 we have required a curve $\gamma: I \to \mathbb{R}^n$ to be defined on an *interval*, whereas here we are just assuming U to be an open set in $\mathbb{R}$. Of course there are open sets in $\mathbb{R}$ which are not intervals, for example the union of two disjoint open intervals. Notice however, that if $\gamma: U \to \mathbb{R}^n$ is a parametrized manifold with $U \subset \mathbb{R}$, then for each $t_0 \in U$ there exists an open interval I around t_0 in U, and the restriction of γ to that interval is a parametrized curve in the old sense. In future, when we speak of a parametrized curve, we will just assume that it is defined on an open set in $\mathbb{R}$.

Perhaps the case $m = 0$ needs some explanation. By definition $\mathbb{R}^0$ is the trivial vector space $\{0\}$, and a map $\sigma: \mathbb{R}^0 \to \mathbb{R}^n$ has just one value $p = \sigma(0)$. By definition the map $0 \mapsto p$ is smooth and regular, and thus a 0-dimensional parametrized manifold in $\mathbb{R}^n$ is a point $p \in \mathbb{R}^n$.

Example 1.1.1 Let $\sigma(u, v) = (\cos u, \sin u, \cos v, \sin v) \in \mathbb{R}^4$. Then

$$D\sigma(u, v) = \begin{pmatrix} -\sin u & 0 \\ \cos u & 0 \\ 0 & -\sin v \\ 0 & \cos v \end{pmatrix}$$

has rank 2, so that σ is a 2-dimensional manifold in $\mathbb{R}^4$.

Example 1.1.2 The graph of a smooth function $h: U \to \mathbb{R}^{n-m}$, where $U \subset \mathbb{R}^m$ is open, is an m-dimensional parametrized manifold in $\mathbb{R}^n$. Let $\sigma(x) = (x, h(x)) \in \mathbb{R}^n$, then $D\sigma(x)$ is an $n \times m$ matrix, of which the first m rows comprise a unit matrix. It follows that $D\sigma(x)$ has rank m for all x, so that σ is regular.

Many basic results about surfaces allow generalization, often with proofs analogous to the 2-dimensional case. Below is an example. By definition, a *reparametrization* of a parametrized manifold $\sigma: U \to \mathbb{R}^n$ is a parametrized manifold of the form $\tau = \sigma \circ \phi$ where $\phi: W \to U$ is a diffeomorphism of open sets.

Theorem 1.1. *Let $\sigma: U \to \mathbb{R}^n$ be a parametrized manifold with $U \subset \mathbb{R}^m$, and assume it is regular at $p \in U$. Then there exists a neighborhood of p in U, such that the restriction of σ to that neighborhood allows a reparametrization which is the graph of a smooth function, where $n - m$ among the variables $x_1, \dots, x_n$ are considered as functions of the remaining m variables.*

Proof. The proof, which is an application of the inverse function theorem for functions of m variables, is entirely similar to the proof of the corresponding result for surfaces (Theorem 2.11 of Geometry 1). □

1.2 Embedded parametrizations

We introduce a property of parametrizations, which essentially means that there are no self intersections. Basically this means that the parametrization is injective, but we shall see that injectivity alone is not sufficient to ensure the behavior we want, and we shall supplement injectivity with another condition.

Definition 1.2.1. Let $A \subset \mathbb{R}^m$ and $B \subset \mathbb{R}^n$. A map $f: A \to B$ which is continuous, bijective and has a continuous inverse is called a *homeomorphism*.

The sets A and B are metric spaces, with the same distance functions as the surrounding Euclidean spaces, and the continuity of f and f^{-1} is assumed to be with respect to these metrics.

Definition 1.2.2. A regular parametrized manifold $\sigma: U \to \mathbb{R}^n$ which is a homeomorphism $U \to \sigma(U)$, is called an *embedded parametrized manifold*.

In particular this definition applies to curves and surfaces, and thus we can speak of embedded parametrized curves and embedded parametrized surfaces.

In addition to being smooth and regular, the condition on σ is thus that it is injective and that the inverse map $\sigma(x) \mapsto x$ is continuous $\sigma(U) \to U$. Since the latter condition of continuity is important in the following, we shall elaborate a bit on it.

Definition 1.2.3. Let $A \subset \mathbb{R}^n$. A subset $B \subset A$ is said to be *relatively open* if it has the form $B = A \cap W$ for some open set $W \subset \mathbb{R}^n$.

For example, the interval $B = [0; 1[$ is relatively open in $A = [0, \infty[$, since it has the form $A \cap W$ with $W =]-1, 1[$. As another example, let $A = \{(x, 0)\}$ be the x-axis in $\mathbb{R}^2$. A subset $B \subset A$ is relatively open if and only if it has the form $U \times \{0\}$ where $U \subset \mathbb{R}$ is open (of course, no subsets of the axis are open in $\mathbb{R}^2$, except the empty set). If A is already open in $\mathbb{R}^n$, then the relatively open subsets are just the open subsets.

It is easily seen that $B \subset A$ is relatively open if and only if it is open in the metric space of A equipped with the distance function of $\mathbb{R}^n$.

The continuity of $\sigma(x) \mapsto x$ from $\sigma(U)$ to U means by definition that every open subset $V \subset U$ has an open preimage in $\sigma(U)$. The preimage of V by this map is $\sigma(V)$, hence the condition is that $V \subset U$ open implies $\sigma(V) \subset \sigma(U)$ open. By the preceding remark and definition this is equivalent to require that for each open $V \subset U$ there exists an open set $W \subset \mathbb{R}^n$ such that

$$\sigma(V) = \sigma(U) \cap W. \tag{1.1}$$

The importance of this condition is illustrated in Example 1.2.2 below.

Notice that a reparametrization $\tau = \sigma \circ \phi$ of an embedded parametrized manifold is again embedded. Here $\phi: W \to U$ is a diffeomorphism of open sets, and it is clear that τ is a homeomorphism onto its image if and only if σ is a homeomorphism onto its image.

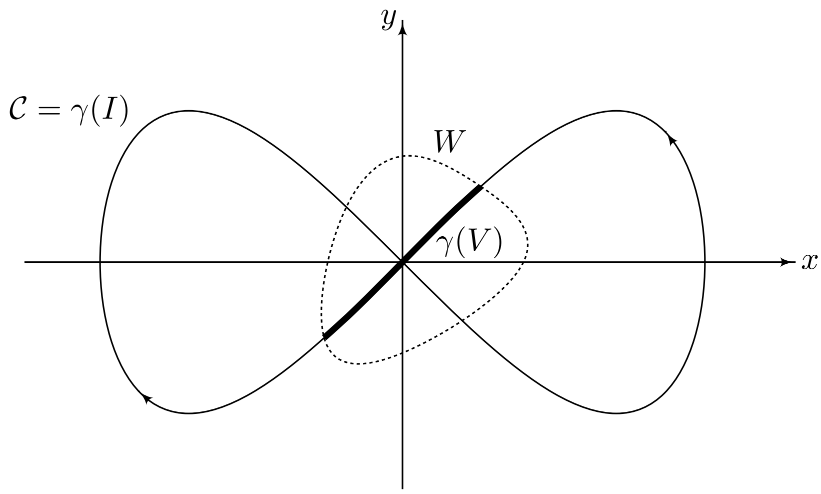
Example 1.2.1 The graph of a smooth function $h\colon U \to \mathbb{R}^{n-m}$, where $U \subset \mathbb{R}^m$ is open, is an embedded parametrized manifold in $\mathbb{R}^n$. It is regular by Example 1.1.2, and it is clearly injective. The inverse map $\sigma(x) \to x$ is the restriction to $\sigma(U)$ of the projection

$$\mathbb{R}^n \ni x \mapsto (x_1, \dots, x_m) \in \mathbb{R}^m$$

on the first m coordinates. Hence this inverse map is continuous. The open set W in (1.1) can be chosen as $W = V \times \mathbb{R}^{n-m}$.

Example 1.2.2 Consider the parametrized curve $\gamma(t) = (\cos t, \cos t \sin t)$ in $\mathbb{R}^2$. It is easily seen to be regular, and it has a self-intersection in $(0,0)$, which equals $\gamma(k\frac{\pi}{2})$ for all odd integers k (see the figure below).

The interval $I =]-\frac{\pi}{2}, 3\frac{\pi}{2}[$ contains only one of the values $k\frac{\pi}{2}$, and the restriction of γ to I is an injective regular curve. The image $\gamma(I)$ is the full set $\mathcal{C}$ in the figure below.

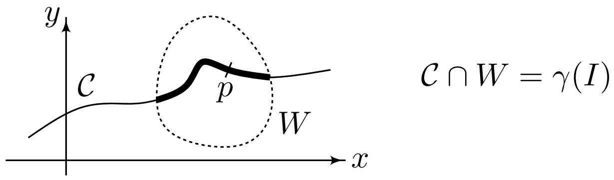


The restriction $\gamma|_I$ is *not* a homeomorphism from I to $\mathcal{C}$. The problem occurs in the point $(0,0) = \gamma(\frac{\pi}{2})$. Consider an open interval $V =]\frac{\pi}{2} - \epsilon, \frac{\pi}{2} + \epsilon[$ where $0 < \epsilon < \pi$. The image $\gamma(V)$ is shown in the figure, and it does not have the form $\mathcal{C} \cap W$ for any open set $W \subset \mathbb{R}^2$, because W necessarily contains points from the other branch through $(0,0)$. Hence $\gamma|_I$ is not an embedded parametrized curve.

It is exactly the purpose of the homeomorphism requirement to exclude the possibility of a 'hidden' self-intersection, as in Example 1.2.2. Based on the example one can easily construct similar examples in higher dimension.

1.3 Curves

As mentioned in the introduction, we shall define a concept of manifolds which applies to subsets of $\mathbb{R}^n$ rather than to parametrizations. In order to understand the definition properly, we begin by the case of curves in $\mathbb{R}^2$. The idea is that a subset of $\mathbb{R}^2$ is a curve, if in a neighborhood of each of its points it is the image of an embedded parametrized curve.



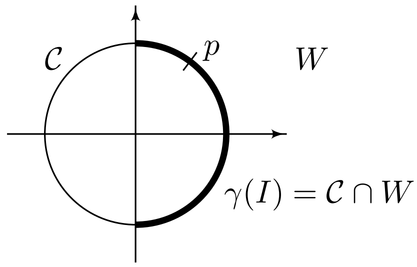
Definition 1.3. A *curve* in $\mathbb{R}^2$ is a non-empty set $\mathcal{C} \subset \mathbb{R}^2$ satisfying the following for each $p \in \mathcal{C}$. There exists an open neighborhood $W \subset \mathbb{R}^2$ of p, an open set $I \subset \mathbb{R}$, and an embedded parametrized curve $\gamma \colon I \to \mathbb{R}^2$ with image

$$\gamma(I) = \mathcal{C} \cap W. \tag{1.2}$$

Example 1.3.1. The image $\mathcal{C} = \gamma(I)$ of an embedded parametrized curve is a curve. In the condition above we can take $W = \mathbb{R}^2$.

Example 1.3.2. The circle $\mathcal{C} = S^1 = \{(x, y) \mid x^2 + y^2 = 1\}$ is a curve. In order to verify the condition in Definition 1.3, let $p \in \mathcal{C}$ be given. For simplicity we assume that $p = (x_0, y_0)$ with $x_0 > 0$.

Let $W \subset \mathbb{R}^2$ be the right half plane $\{(x, y) \mid x > 0\}$, then W is an open neighborhood of p, and the parametrized curve $\gamma(t) = (\cos t, \sin t)$ with $t \in I =]-\frac{\pi}{2}, \frac{\pi}{2}[$ is regular and satisfies (1.2). It is an embedded curve since the inverse map $\gamma(t) \mapsto t$ is given by $(x, y) \mapsto \tan^{-1}(y/x)$, which is continuous.



Example 1.3.3. An 8-shaped set like the one in Example 1.2.2 is *not* a curve in $\mathbb{R}^2$. In that example we showed that the parametrization by $(\cos t, \cos t \sin t)$ was not embedded, but of course this does not rule out that some other parametrization could satisfy the requirement in Definition 1.3. That this is not the case can be seen from Lemma 1.3 below.

It is of importance to exclude sets like this, because there is not a well defined tangent line in the point p of self-intersection. If a parametrization is given, we can distinguish the passages through p, and thus determine a tangent line for each branch. However, without a chosen parametrization both branches have to be taken into account, and then there is not a unique tangent line in p.

The definition of a curve allows the following useful reformulation.

Lemma 1.3. *Let $\mathcal{C} \subset \mathbb{R}^2$ be non-empty. Then $\mathcal{C}$ is a curve if and only if it satisfies the following condition for each $p \in \mathcal{C}$:*

There exists an open neighborhood $W \subset \mathbb{R}^2$ of p, such that $\mathcal{C} \cap W$ is the graph of a smooth function h, where one of the variables x_1, x_2 is considered a function of the other variable.

Proof. Assume that $\mathcal{C}$ is a curve and let $p \in \mathcal{C}$. Let $\gamma: I \to \mathbb{R}^2$ be an embedded parametrized curve satisfying (1.2) and with $\gamma(t_0) = p$. By Theorem 1.1, in the special case $m = 1$, we find that there exists a neighborhood V of t_0 in I such that $\gamma|_V$ allows a reparametrization as a graph. It follows from (1.1) and (1.2) that there exists an open set $W' \subset \mathbb{R}^2$ such that $\gamma(V) = \gamma(I) \cap W' = \mathcal{C} \cap W \cap W'$. The set $W \cap W'$ has all the properties desired of W in the lemma.

Conversely, assume that the condition in the lemma holds, for a given point p say with

$$\mathcal{C} \cap W = \{(t, h(t)) \mid t \in I\},$$

where $I \subset \mathbb{R}$ is open and $h: I \to \mathbb{R}$ is smooth. The curve $t \mapsto (t, h(t))$ has image $\mathcal{C} \cap W$, and according to Example 1.2.1 it is an embedded parametrized curve. Hence the condition in Definition 1.3 holds, and $\mathcal{C}$ is a curve. $\square$

The most common examples of plane curves are constructed by means of the following general theorem, which frees us from finding explicit embedded parametrizations that satisfy (1.2). For example, the proof in Example 1.3.2, that the circle is a curve, could have been simplified by means of this theorem.

Recall that a point $p \in \Omega$, where $\Omega \subset \mathbb{R}^n$ is open, is called *critical* for a differentiable function $f: \Omega \to \mathbb{R}$ if

$$f'_{x_1}(p) = \cdots = f'_{x_n}(p) = 0.$$

Theorem 1.3. *Let $f: \Omega \to \mathbb{R}$ be a smooth function, where $\Omega \subset \mathbb{R}^2$ is open, and let $c \in \mathbb{R}$. If it is not empty, the set*

$$\mathcal{C} = \{p \in \Omega \mid f(p) = c, \ p \text{ is not critical}\}$$

is a curve in $\mathbb{R}^2$.

Proof. By continuity of the partial derivatives, the set of non-critical points in Ω is an open subset. If we replace Ω by this set, the set $\mathcal{C}$ can be expressed as a level curve $\{p \in \Omega \mid f(p) = c\}$, to which we can apply the implicit function theorem (see Geometry 1, Corollary 1.5). It then follows from Lemma 1.3 that $\mathcal{C}$ is a curve. $\square$

Example 1.3.4. The set $\mathcal{C} = \{(x, y) \mid x^2 + y^2 = c\}$ is a curve in $\mathbb{R}^2$ for each $c > 0$, since it contains no critical points for $f(x, y) = x^2 + y^2$.

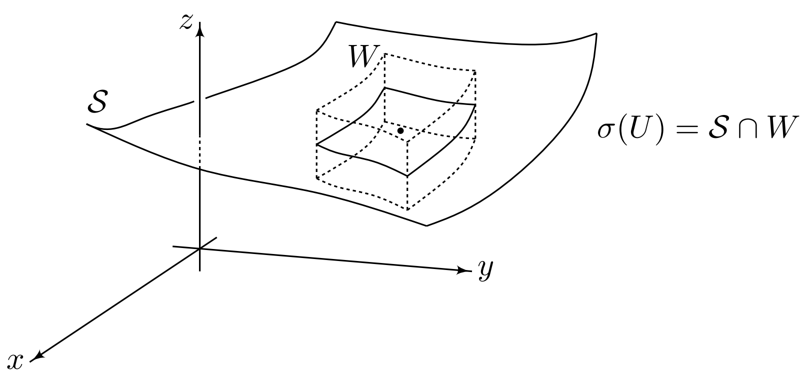
Example 1.3.5. Let $\mathcal{C} = \{(x, y) \mid x^4 - x^2 + y^2 = 0\}$. It is easily seen that this is exactly the image $\gamma(I)$ in Example 1.2.2. The point $(0, 0)$ is the only critical point in $\mathcal{C}$ for the function $f(x, y) = x^4 - x^2 + y^2$, and hence it follows from Theorem 1.3 that $\mathcal{C} \setminus \{(0, 0)\}$ is a curve in $\mathbb{R}^2$. As mentioned in Example 1.3.3, the set $\mathcal{C}$ itself is not a curve, but this conclusion cannot be drawn from from Theorem 1.3.

1.4 Surfaces

We proceed in the same fashion as for curves.

Definition 1.4. A *surface* in $\mathbb{R}^3$ is a non-empty set $\mathcal{S} \subset \mathbb{R}^3$ satisfying the following property for each point $p \in \mathcal{S}$. There exists an open neighborhood $W \subset \mathbb{R}^3$ of p and an embedded parametrized surface $\sigma \colon U \to \mathbb{R}^3$ with image

$$\sigma(U) = \mathcal{S} \cap W. \tag{1.3}$$



Example 1.4.1. The image $\mathcal{S} = \sigma(U)$ of an embedded parametrized surface is a surface in $\mathbb{R}^3$. In the condition above we can take $W = \mathbb{R}^3$.

Example 1.4.2. The sphere $\mathcal{S} = \{(x, y, z) \mid x^2 + y^2 + z^2 = 1\}$ and the cylinder $\mathcal{S} = \{(x, y, z) \mid x^2 + y^2 = 1\}$ are surfaces. Rather than providing for each point $p \in \mathcal{S}$ an explicit embedded parametrization satisfying (1.3), we use the following theorem, which is analogous to Theorem 1.3.

Theorem 1.4. *Let $f: \Omega \to \mathbb{R}$ be a smooth function, where $\Omega \subset \mathbb{R}^3$ is open, and let $c \in \mathbb{R}$. If it is not empty, the set*

$$\mathcal{S} = \{p \in \Omega \mid f(p) = c,\ p \text{ is not critical}\} \tag{1.4}$$

is a surface in $\mathbb{R}^3$.

Proof. The proof, which combines Geometry 1, Corollary 1.6, with Lemma 1.4 below, is entirely similar to that of Theorem 1.3. □

Example 1.4.3. Let us verify for the sphere that it contains no critical points for the function $f(x,y,z) = x^2 + y^2 + z^2$. The partial derivatives are $f'_x = 2x, f'_y = 2y, f'_z = 2z$, and they vanish simultaneously only at $(x,y,z) = (0,0,0)$. This point does not belong to the sphere, hence it is a surface. The verification for the cylinder is similar.

Lemma 1.4. *Let $\mathcal{S} \subset \mathbb{R}^3$ be non-empty. Then $\mathcal{S}$ is a surface if and only if it satisfies the following condition for each $p \in \mathcal{S}$:*

There exist an open neighborhood $W \subset \mathbb{R}^3$ of p, such that $\mathcal{S} \cap W$ is the graph of a smooth function h, where one of the variables x_1, x_2, x_3 is considered a function of the other two variables.

Proof. The proof is entirely similar to that of Lemma 1.3. □

1.5 Chart and atlas

As mentioned in the introduction there exist surfaces, for example the sphere, which require several, in general overlapping, parametrizations. This makes the following concepts relevant.

Definition 1.5. Let $\mathcal{S}$ be a surface in $\mathbb{R}^3$. A *chart* on $\mathcal{S}$ is an injective regular parametrized surface $\sigma: U \to \mathbb{R}^3$ with image $\sigma(U) \subset \mathcal{S}$. A collection of charts $\sigma_i: U_i \to \mathbb{R}^3$ on $\mathcal{S}$ is said to *cover* $\mathcal{S}$ if $\mathcal{S} = \cup_i \sigma_i(U_i)$. In that case the collection is called an *atlas* of $\mathcal{S}$.

Example 1.5.1. The image $\mathcal{S} = \sigma(U)$ of an embedded parametrized surface as in Example 1.4.1 has an atlas consisting just of the chart σ itself.

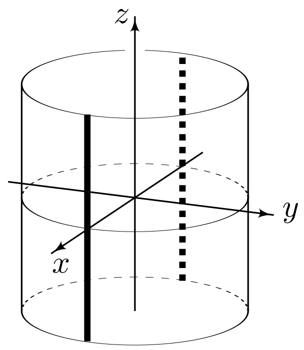
Example 1.5.2. The map

$$\sigma(u,v) = (\cos v, \sin v, u), \quad u, v \in \mathbb{R}$$

is regular and covers the cylinder $\mathcal{S} = \{(x,y,z) \mid x^2 + y^2 = 1\}$, but it is not injective. Let

$$U_1 = \{(u,v) \in \mathbb{R}^2 \mid -\pi < v < \pi\}, \quad U_2 = \{(u,v) \in \mathbb{R}^2 \mid 0 < v < 2\pi\},$$

and let σ_i denote the restriction of σ to U_i for $i = 1, 2$. Then σ_1 and σ_2 are both injective, σ_1 covers $\mathcal{S}$ with the exception of a vertical line on the back where $x = -1$, and σ_2 covers with the exception of a vertical line on the front where $x = 1$. Together they cover the entire set and thus they constitute an atlas.



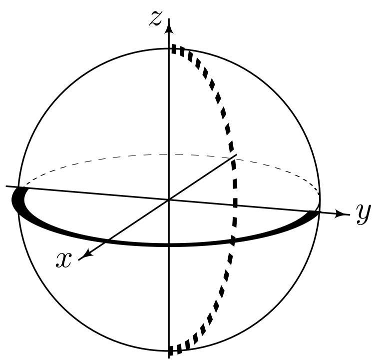
Example 1.5.3. The spherical coordinate map

$$\sigma(u, v) = (\cos u \cos v, \cos u \sin v, \sin u), \quad -\tfrac{\pi}{2} < u < \tfrac{\pi}{2}, -\pi < v < \pi,$$

and its variation

$$\tilde{\sigma}(u, v) = (\cos u \cos v, \sin u, \cos u \sin v), \quad -\tfrac{\pi}{2} < u < \tfrac{\pi}{2}, 0 < v < 2\pi,$$

are charts on the unit sphere. The restrictions on u and v ensure that they are regular and injective. The chart σ covers the sphere except a half circle (a meridian) in the xz-plane, on the back where $x \leq 0$, and the chart $\tilde{\sigma}$ similarly covers with the exception of a half circle in the xy-plane, on the front where $x \geq 0$ (half of the 'equator'). As seen in the following figure the excepted half-circles are disjoint. Hence the two charts together cover the full sphere and they constitute an atlas.



Theorem 1.5. *Let $\mathcal{S}$ be a surface. There exists an atlas of it.*

Proof. For each $p \in \mathcal{S}$ we choose an embedded parametrized surface σ as in Definition 1.4. Since a homeomorphism is injective, this parametrization is a chart on $\mathcal{S}$. The collection of all these charts is an atlas. $\square$

1.6 Manifolds

We now return to the general situation where m and n are arbitrary integers with $0 \leq m \leq n$.

Definition 1.6.1. An m-dimensional *manifold* in $\mathbb{R}^n$ is a non-empty set $\mathcal{S} \subset \mathbb{R}^n$ satisfying the following property for each point $p \in \mathcal{S}$. There exists an open neighborhood $W \subset \mathbb{R}^n$ of p and an m-dimensional embedded (see Definition 1.2.2) parametrized manifold $\sigma: U \to \mathbb{R}^n$ with image $\sigma(U) = \mathcal{S} \cap W$.

The surrounding space $\mathbb{R}^n$ is said to be the *ambient* space of the manifold.

Clearly this generalizes Definitions 1.3 and 1.4, a curve is a 1-dimensional manifold in $\mathbb{R}^2$ and a surface is a 2-dimensional manifold in $\mathbb{R}^3$.

Example 1.6.1 The case $m = 0$. It was explained in Section 1.1 that a 0-dimensional parametrized manifold is a map $\mathbb{R}^0 = \{0\} \to \mathbb{R}^n$, whose image consists of a single point p. An element p in a set $\mathcal{S} \subset \mathbb{R}^n$ is called *isolated* if it is the only point from $\mathcal{S}$ in some neighborhood of p, and the set $\mathcal{S}$ is called *discrete* if all its points are isolated. By going over Definition 1.6.1 for the case $m = 0$ it is seen that a 0-dimensional manifold in $\mathbb{R}^n$ is the same as a discrete subset.

Example 1.6.2 If we identify $\mathbb{R}^m$ with the set $\{(x_1, \dots, x_m, 0 \dots, 0)\} \subset \mathbb{R}^n$, it is an m-dimensional manifold in $\mathbb{R}^n$.

Example 1.6.3 An open set $\Omega \subset \mathbb{R}^n$ is an n-dimensional manifold in $\mathbb{R}^n$. Indeed, we can take $W = \Omega$ and $\sigma =$ the identity map in Definition 1.6.1.

Example 1.6.4 Let $\mathcal{S}' \subset \mathcal{S}$ be a relatively open subset of an m-dimensional manifold in $\mathbb{R}^n$. Then $\mathcal{S}'$ is an m-dimensional manifold in $\mathbb{R}^n$.

The following lemma generalizes Lemmas 1.3 and 1.4.

Lemma 1.6. *Let $\mathcal{S} \subset \mathbb{R}^n$ be non-empty. Then $\mathcal{S}$ is an m-dimensional manifold if and only if it satisfies the following condition for each $p \in \mathcal{S}$:*

There exist an open neighborhood $W \subset \mathbb{R}^n$ of p, such that $\mathcal{S} \cap W$ is the graph of a smooth function h, where $n - m$ of the variables $x_1, \dots, x_n$ are considered as functions of the remaining m variables.

Proof. The proof is entirely similar to that of Lemma 1.3. □

Theorem 1.6. *Let $f: \Omega \to \mathbb{R}^k$ be a smooth function, where $k \leq n$ and where $\Omega \subset \mathbb{R}^n$ is open, and let $c \in \mathbb{R}^k$. If it is not empty, the set*

$$\mathcal{S} = \{p \in \Omega \mid f(p) = c, \operatorname{rank} Df(p) = k\}$$

is an $n-k$-dimensional manifold in $\mathbb{R}^n$.

Proof. Similar to that of Theorem 1.3 for curves, by means of the implicit function theorem (Geometry 1, Corollary 1.6) and Lemma 1.6. □

A manifold $\mathcal{S}$ in $\mathbb{R}^n$ which is constructed as in Theorem 1.6 as the set of solutions to an equation $f(x) = c$ is often called a *variety*. In particular, if the equation is algebraic, which means that the coordinates of f are polynomials in $x_1, \dots, x_n$, then $\mathcal{S}$ is called an *algebraic variety*.

Example 1.6.5 In analogy with Example 1.4.3 we can verify that the *m-sphere*

$$S^m = \{x \in \mathbb{R}^{m+1} \mid x_1^2 + \cdots + x_{m+1}^2 = 1\}$$

is an m-dimensional manifold in $\mathbb{R}^{m+1}$.

Example 1.6.6 The set

$$\mathcal{S} = S^1 \times S^1 = \{x \in \mathbb{R}^4 \mid x_1^2 + x_2^2 = x_3^2 + x_4^2 = 1\}$$

is a 2-dimensional manifold in $\mathbb{R}^4$. Let

$$f(x_1, x_2, x_3, x_4) = \begin{pmatrix} x_1^2 + x_2^2 \\ x_3^2 + x_4^2 \end{pmatrix},$$

then

$$Df(x) = \begin{pmatrix} 2x_1 & 2x_2 & 0 & 0 \\ 0 & 0 & 2x_3 & 2x_4 \end{pmatrix}$$

and it is easily seen that this matrix has rank 2 for all $x \in \mathcal{S}$.

Definition 1.6.2. Let $\mathcal{S}$ be an m-dimensional manifold in $\mathbb{R}^n$. A *chart* on $\mathcal{S}$ is an m-dimensional injective regular parametrized manifold $\sigma: U \to \mathbb{R}^n$ with image $\sigma(U) \subset \mathcal{S}$, and an *atlas* is a collection of charts $\sigma_i: U_i \to \mathbb{R}^n$ which cover $\mathcal{S}$, that is, $\mathcal{S} = \cup_i \sigma_i(U_i)$.

As in Theorem 1.5 it is seen that every manifold in $\mathbb{R}^n$ possesses an atlas.

1.7 The coordinate map of a chart

In Definition 1.6.2 we require that σ is injective, but we do not require that the inverse map is continuous, as in Definition 1.6.1. Surprisingly, it turns out that the inverse map has an even stronger property, it is smooth in a certain sense.

Definition 1.7. Let $\mathcal{S}$ be a manifold in $\mathbb{R}^n$, and let $\sigma: U \to \mathcal{S}$ be a chart. If $p \in \mathcal{S}$ we call $(x_1, \dots, x_m) \in U$ the *coordinates* of p with respect to σ when $p = \sigma(x)$. The map $\sigma^{-1}: \sigma(U) \to U$ is called the *coordinate map* of σ.

Theorem 1.7. *Let $\sigma: U \to \mathbb{R}^n$ be a chart on a manifold $\mathcal{S} \subset \mathbb{R}^n$, and let $p_0 \in \sigma(U)$ be given. The coordinate map σ^{-1} allows a smooth extension, defined on an open neighborhood of p_0 in $\mathbb{R}^n$.*

More precisely, let $q_0 \in U$ with $p_0 = \sigma(q_0)$. Then there exist open neighborhoods $W \subset \mathbb{R}^n$ of p_0 and $V \subset U$ of q_0 such that

$$\sigma(V) = \mathcal{S} \cap W, \tag{1.5}$$

and a smooth map $\varphi: W \to V$ such that

$$\varphi(\sigma(q)) = q \tag{1.6}$$

for all $q \in V$.

Proof. Let $W \subset \mathbb{R}^n$ be an open neighborhood of p_0 in which $\mathcal{S}$ is parametrized as a graph, as in Lemma 1.6. Say the graph is of the form

$$\tilde{\sigma}(x_1, \dots, x_m) = (x, h(x)) \in \mathbb{R}^n,$$

where $\tilde{U} \subset \mathbb{R}^m$ is open and $h: \tilde{U} \to \mathbb{R}^{n-m}$ smooth, then $\mathcal{S} \cap W = \tilde{\sigma}(\tilde{U})$. Let $\pi(x_1, \dots, x_n) = (x_1, \dots, x_m)$, then

$$\tilde{\sigma}(\pi(p)) = p$$

for each point $p = (x, h(x))$ in $\mathcal{S} \cap W$.

Since σ is continuous, the subset $U_1 = \sigma^{-1}(W)$ of U is open. The map $\pi \circ \sigma: U_1 \to \tilde{U}$ is smooth, being composed by smooth maps, and it satisfies

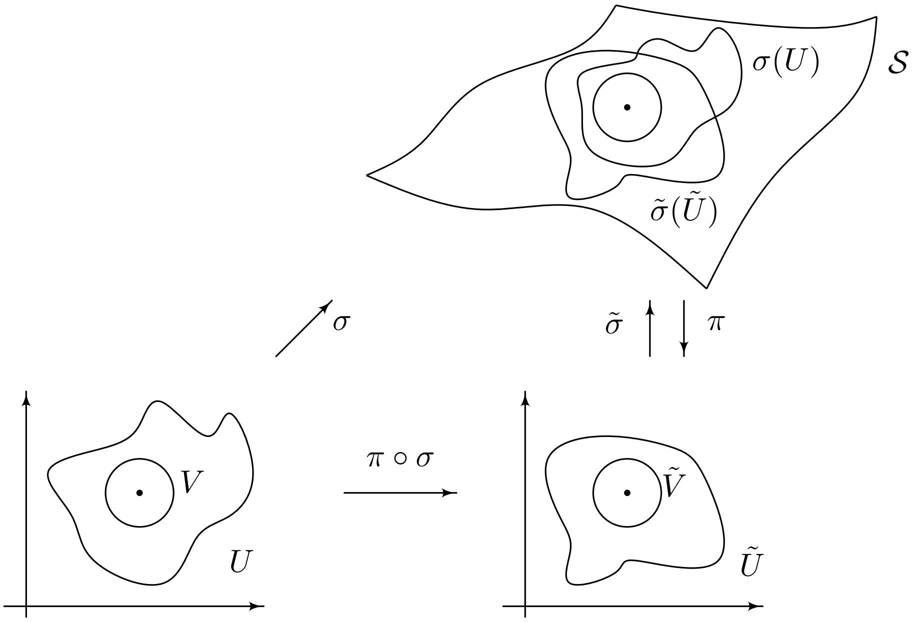
$$\tilde{\sigma} \circ (\pi \circ \sigma) = \sigma. \tag{1.7}$$

By the chain rule for smooth maps we have the matrix product equality

$$D(\tilde{\sigma}) D(\pi \circ \sigma) = D\sigma,$$

and since the $n \times m$ matrix $D\sigma$ on the right has independent columns, the determinant of the $m \times m$ matrix $D(\pi \circ \sigma)$ must be non-zero in each $x \in U_1$ (according to the rule from matrix algebra that $\operatorname{rank}(AB) \le \operatorname{rank}(B)$).

By the inverse function theorem, there exists open sets $V \subset U_1$ and $\tilde{V} \subset \tilde{U}$ around q_0 and $\pi(\sigma(q_0))$, respectively, such that $\pi \circ \sigma$ restricts to a diffeomorphism of V onto $\tilde{V}$. Note that $\sigma(V) = \tilde{\sigma}(\tilde{V})$ by (1.7).



Let $\tilde{W} = W \cap \pi^{-1}(\tilde{V}) \subset \mathbb{R}^n$. This is an open set, and it satisfies

$$\mathcal{S} \cap \tilde{W} = \mathcal{S} \cap W \cap \pi^{-1}(\tilde{V}) = \tilde{\sigma}(\tilde{U}) \cap \pi^{-1}(\tilde{V}) = \tilde{\sigma}(\tilde{V}) = \sigma(V).$$

The map $\varphi = (\pi \circ \sigma)^{-1} \circ \pi\colon \tilde{W} \to V$ is smooth and satisfies (1.6). $\square$

Corollary 1.7. *Let $\sigma\colon U \to \mathcal{S}$ be a chart. Then σ is an embedded parametrized manifold, and the image $\sigma(U)$ is relatively open in $\mathcal{S}$.*

Proof. For each $q_0 \in U$ we choose open sets $V \subset U$ and $W \subset \mathbb{R}^n$, and a map $\varphi\colon W \to V$ as in Theorem 1.7. The inverse of σ is the restriction of the smooth map φ, hence in particular it is continuous. Furthermore, the union of all these sets W is open and intersects $\mathcal{S}$ exactly in $\sigma(U)$. Hence $\sigma(U)$ is relatively open, according to Definition 1.2.3. $\square$

It follows from the corollary that every chart on a manifold satisfies the condition in Definition 1.6.1 of being imbedded with open image. This does not render that condition superfluous, however. The point is that once it is known that $\mathcal{S}$ *is* a manifold, then the condition is automatically fulfilled for *all* charts on $\mathcal{S}$.

1.8 Transition maps

Since the charts in an atlas of a manifold $\mathcal{S}$ in $\mathbb{R}^n$ may overlap with each other, it is important to study the change from one chart to another. The map $\sigma_2^{-1} \circ \sigma_1\colon x \mapsto \tilde{x}$, which maps a set of coordinates x in a chart σ_1 to the coordinates of the image $\sigma_1(x)$ with respect to another chart σ_2, is called the *transition map* between the charts. We will show that such a change of coordinates is just a reparametrization.

Let $\Omega \subset \mathbb{R}^k$ be open and let $f\colon \Omega \to \mathbb{R}^n$ be a smooth map with $f(\Omega) \subset \mathcal{S}$. Let $\sigma\colon U \to \mathbb{R}^n$ be a chart on $\mathcal{S}$, then the map $\sigma^{-1} \circ f$, which is defined on

$$f^{-1}(\sigma(U)) = \{x \in \Omega \mid f(x) \in \sigma(U)\} \subset \mathbb{R}^k,$$

is called the *coordinate expression* for f with respect to σ.

Lemma 1.8. *The set $f^{-1}(\sigma(U))$ is open and the coordinate expression is smooth from this set into U.*

Proof. Since f is continuous into $\mathcal{S}$, and $\sigma(U)$ is open by Corollary 1.7, it follows that the inverse image $f^{-1}(\sigma(U))$ is open. Furthermore, if an element x_0 in this set is given, we let $p_0 = f(x_0)$ and choose V, W and $\varphi\colon W \to V$ as in Theorem 1.7. It follows that $\sigma^{-1} \circ f = \varphi \circ f$ in a neighborhood of x_0. The latter map is smooth, being composed by smooth maps. $\square$

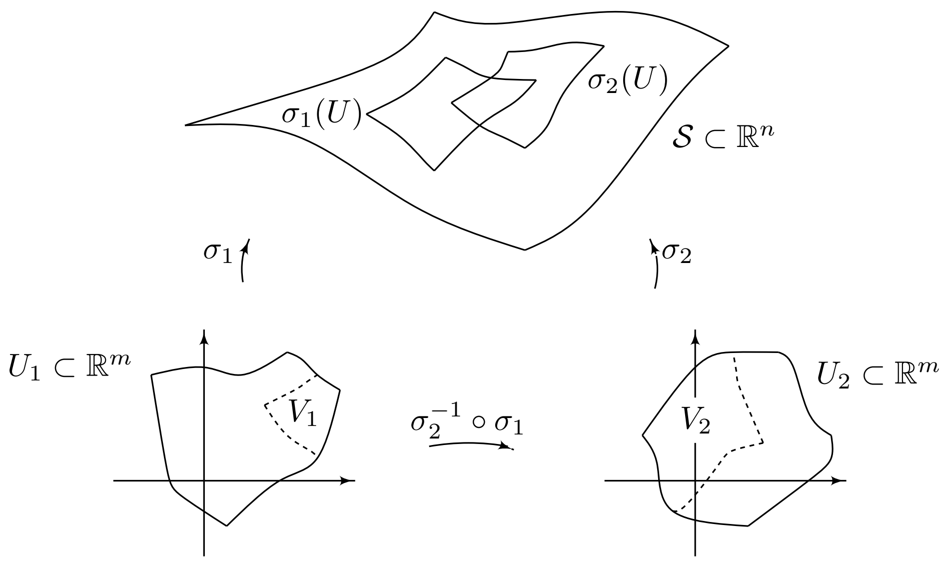
Theorem 1.8. *Let $\mathcal{S}$ be a manifold in $\mathbb{R}^n$, and let $\sigma_1\colon U_1 \to \mathbb{R}^n$ and $\sigma_2\colon U_2 \to \mathbb{R}^n$ be charts on $\mathcal{S}$. Let*

$$V_i = \sigma_i^{-1}(\sigma_1(U_1) \cap \sigma_2(U_2)) \subset U_i$$

for $i = 1, 2$. These are open sets and the transition map

$$\sigma_2^{-1} \circ \sigma_1\colon V_1 \to V_2 \subset \mathbb{R}^m$$

is a diffeomorphism.



Proof. Immediate from Lemma 1.8. $\square$

Chapter 2

Abstract manifolds

The notion of a manifold $\mathcal{S}$ defined in the preceding chapter assumes $\mathcal{S}$ to be a subset of a Euclidean space $\mathbb{R}^n$. However, a more axiomatic and abstract approach to differential geometry is possible, and in many ways preferable. Of course, a manifold in $\mathbb{R}^n$ must satisfy the axioms that we set up for an abstract manifold. Our axioms will be based on properties of charts.

From the point of view of differential geometry the most important property of a manifold is that it allows the concept of a smooth function. We will define this notion and the more general notion of a smooth map between abstract manifolds.

2.1 Topological spaces

Since the aim of differential geometry is to bring the methods of differential calculus into geometry, the most important property that we wish an abstract manifold to have is the possibility of differentiating functions on it. However, before we can speak of differentiable functions, we must be able to speak of continuous functions. In this preliminary section we will briefly introduce the abstract framework for that, the structure of a topological space. Topological spaces is a topic of general topology, here we will just introduce the most essential notions. Although the framework is more general, the concepts we introduce will be familiar to a reader who is acquainted with the theory of metric spaces.

Definition 2.1.1. A *topological space* is a non-empty set X equipped with a distinguished family of subsets, called the *open sets*, with the following properties:

1) the empty set and the set X are both open,
2) the intersection of any finite collection of open sets is again open,
3) the union of any collection (finite or infinite) of open sets is again open.

Example 2.1.1 In the Euclidean spaces $X = \mathbb{R}^k$ there is a standard notion of open sets, and the properties in the above axioms are known to hold. Thus $\mathbb{R}^k$ is a topological space.

Example 2.1.2 Let X be a metric space. Again there is a standard notion of open sets in X, and it is a fundamental result from the theory of metric spaces that the family of all open sets in X has the properties above. In this fashion every metric space is a topological space.

Example 2.1.3 Let X be an arbitrary set. If we equip X with the collection consisting just of the empty set and X, it becomes a topological space. We say in this case that X has the *trivial topology*. In general this topology does not result from a metric, as in Example 2.1.2. The topology on X obtained from the collection of all subsets, is called the *discrete topology*. It results from the discrete metric, by which all non-trivial distances are 1.

The following definitions are generalizations of well-known definitions in the theory of Euclidean spaces, and more generally, metric spaces.

Definition 2.1.2. A *neighborhood* of a point $x \in X$ is a subset $U \subset X$ with the property that it contains an open set containing x. The *interior* of a set $A \subset X$, denoted A°, is the set of all points $x \in A$ for which A is a neighborhood of x.

Being the union of all the open subsets of A, the interior A° is itself an open set, according to Definition 2.1.1.

Definition 2.1.3. Let $A \subset X$ be a subset. It is said to be *closed* if its complement A^c in X is open. In general, the *closure* of A, denoted $\bar{A}$, is the set of all points $x \in X$ for which every neighborhood contains a point from A, and the *boundary* of A is the set difference $\partial A = \bar{A} \setminus A^\circ$, which consists of all points with the property that each neighborhood meets with both A and A^c.

It is easily seen that the closure $\bar{A}$ is the complement of the interior $(A^c)^\circ$ of A^c. Hence it is a closed set. Likewise, the boundary ∂A is closed.

Definition 2.1.4. Let X and Y be topological spaces, and let $f\colon X \to Y$ be a map. Then f is said to be *continuous at a point* $x \in X$ if for every neighborhood V of $f(x)$ in Y there exists a neighborhood U of x in X such that $f(U) \subset V$, and f is said to be *continuous* if it is continuous at every $x \in X$.

Lemma 2.1.1. *The map $f\colon X \to Y$ is continuous if and only if the inverse image of every open set in Y is open in X.*

Proof. The proof is straightforward. □

Every (non-empty) subset A of a metric space X is in a natural way a metric space of its own, when equipped with the restriction of the distance function of X. The open sets in this metric space A are the relatively open sets, that is, the sets $A \cap W$ where W is an open subset of X (see Definition 1.2.3). This observation has the following generalization.

Lemma 2.1.2. *Let X be a topological space, and let $A \subset X$ be non-empty. Then A is a topological space of its own, when equipped with the collection of all subsets of A of the form $A \cap O$, where O is open in X.*

Proof. The conditions in Definition 2.1.1 are easily verified. □

A subset A of a topological space is always assumed to carry the topology of Lemma 2.1.2, unless otherwise is mentioned. It is called the *induced* (or relative) topology, and the open sets are said to be *relatively open.*

If $A \subset X$ is a subset and f is a map $A \to Y$, then f is said to be continuous at $x \in A$, if it is continuous with respect to the induced topology. It is easily seen that if $f\colon X \to Y$ is continuous at $x \in A$, then the restriction $f|_A\colon A \to Y$ is also continuous at x.

Definition 2.1.5. Let X and Y be topological spaces, and let $A \subset X$ and $B \subset Y$. A map $f\colon A \to B$ which is continuous, bijective and has a continuous inverse is called a *homeomorphism* (compare Definition 1.2.1).

Finally, we mention the following important property of a topological space, which is often assumed in order to exclude some rather peculiar topological spaces.

Definition 2.1.6. A topological space X is said to be *Hausdorff* if for every pair of distinct points $x, y \in X$ there exist disjoint neighborhoods of x and y.

Every metric space is Hausdorff, because if x and y are distinct points, then their mutual distance is positive, and the open balls centered at x and y with radius half of this distance will be disjoint by the triangle inequality. On the other hand, equipped with the trivial topology (see example 2.1.3), a set of at least two elements is not a Hausdorff topological space.

2.2 Abstract manifolds

Let M be a Hausdorff topological space, and let $m \geq 0$ be a fixed natural number.

Definition 2.2.1. An m-dimensional *smooth atlas* of M is a collection $(O_i)_{i\in I}$ of open sets O_i in M such that $M = \cup_{i\in I} O_i$, together with a collection $(U_i)_{i\in I}$ of open sets in $\mathbb{R}^m$ and a collection of homeomorphisms, called *charts*, $\sigma_i\colon U_i \to O_i = \sigma_i(U_i)$, with the following property of *smooth transition on overlaps*:

For each pair $i, j \in I$ the map $\sigma_j^{-1} \circ \sigma_i$ is smooth from the open set $\sigma_i^{-1}(O_i \cap O_j) \subset \mathbb{R}^m$ to $\mathbb{R}^m$.

Example 2.2 Let $\mathcal{S} \subset \mathbb{R}^n$ be an m-dimensional manifold in $\mathbb{R}^n$ (see Definition 1.6.1), which we equip with an atlas as in Definition 1.6.2 (as mentioned below the definition, such an atlas exists). It follows from Corollary 1.7 that for each chart σ the image $O = \sigma(U)$ is open in $\mathcal{S}$ and $\sigma\colon U \to O$ is a homeomorphism. Furthermore, it follows from Theorem 1.8 that the transition maps are smooth. Hence this atlas on $\mathcal{S}$ is a smooth atlas according to Definition 2.2.1.

In the preceding example $\mathcal{S}$ was equipped with an atlas as in Definition 1.6.2, but one must keep in mind that there is not a unique atlas associated with a given manifold in $\mathbb{R}^n$. For example, the use of spherical coordinates is just one of many ways to parametrize the sphere. If we use a different atlas on $\mathcal{S}$, it is still the same manifold in $\mathbb{R}^n$. In order to treat this phenomenon abstractly, we introduce an equivalence relation for different atlases on the same space M.

Definition 2.2.2. Two m-dimensional smooth atlases on M are said to be *compatible*, if every chart from one atlas has smooth transition on its overlap with every chart from the other atlas (or equivalently, if their union is again an atlas).

It can be seen that compatibility is an equivalence relation. An equivalence class of smooth atlases is called a *smooth structure*. It follows from Theorem 1.8 that all atlases (Definition 1.6.2) on a given manifold $\mathcal{S}$ in $\mathbb{R}^n$ are compatible. The smooth structure so obtained on $\mathcal{S}$ is called the standard smooth structure.

Definition 2.2.3. An *abstract manifold* (or just a *manifold*) of *dimension* m, is a Hausdorff topological space M, equipped with an m-dimensional smooth atlas. Compatible atlases are regarded as belonging to the same manifold (the precise definition is thus that a manifold is a Hausdorff topological space equipped with a smooth structure). A *chart* on M is a chart from any atlas compatible with the structure.

It is often required of an abstract manifold that it should have a *countable* atlas (see Section 2.9). We do not require this here.

2.3 Examples

Example 2.3.1 Let M be an m-dimensional real vector space. Fix a basis $v_1, \dots, v_m$ for M, then the map

$$\sigma: (x_1, \dots, x_m) \mapsto x_1 v_1 + \dots + x_m v_m$$

is a linear bijection $\mathbb{R}^m \to M$. We equip M with the distance function so that this map is an isometry, then M is a metric space. Furthermore, the collection consisting just of the map σ, is an atlas. Hence M is an m-dimensional abstract manifold.

If another basis $w_1, \dots, w_m$ is chosen, the atlas consisting of the map

$$\tau: (x_1, \dots, x_m) \mapsto x_1 w_1 + \dots + x_m w_m$$

is compatible with the previous atlas. The transition maps $\sigma^{-1} \circ \tau$ and $\tau^{-1} \circ \sigma$ are linear, hence smooth. In other words, the smooth structure of M is independent of the choice of the basis.

Example 2.3.2 This example generalizes Example 1.6.4. Let M be an abstract manifold, and let M' be an open subset. For each chart $\sigma_i\colon U_i \to O_i \subset M$, let $O_i' = M' \cap O_i$, this is an open subset of M', and the collection of all these sets cover M'. Furthermore, $U_i' = \sigma_i^{-1}(O_i')$ is open in $\mathbb{R}^m$, and the restriction σ_i' of σ_i to this set is a homeomorphism onto its image. Clearly the transition maps $(\sigma_j')^{-1} \circ \sigma_i'$ are smooth, and thus M' is an abstract manifold with the atlas consisting of all these restricted charts.

Example 2.3.3 Let $M = \mathbb{R}$ equipped with the standard metric. Let $\sigma(t) = t^3$ for $t \in U = \mathbb{R}$, then σ is a homeomorphism $U \to M$. The collection of this map alone is an atlas on M. The corresponding differential structure on $\mathbb{R}$ is *different* from the standard differential structure, for the transition map $\sigma^{-1} \circ \mathrm{i}$ between σ and the identity is not smooth at $t = 0$.

Example 2.3.4 Let X be an arbitrary set, equipped with the discrete topology. For each point $x \in X$, we define a map $\sigma\colon \mathbb{R}^0 \to X$ by $\sigma(0) = x$. The collection of all these maps is a 0-dimensional smooth atlas on X.

2.4 Projective space

In this section we give an example of an abstract manifold constructed without a surrounding space $\mathbb{R}^n$.

Let $M = \mathbb{R}\mathrm{P}^m$ be the set of 1-dimensional linear subspaces of $\mathbb{R}^{m+1}$. It is called *real projective space*, and can be given the structure of an abstract m-dimensional manifold as follows.

Assume for simplicity that $m = 2$. Let $\pi\colon x \mapsto [x] = \operatorname{span} x$ denote the natural map of $\mathbb{R}^3 \setminus \{0\}$ onto M, and let $S \subset \mathbb{R}^3$ denote the unit sphere. The restriction of π to S is two-to-one, for each $p \in M$ there are precisely two elements $\pm x \in S$ with $\pi(x) = p$. We thus have a model for M as the set of all pairs of antipodal points in S.

We shall equip M as a Hausdorff topological space as follows. A set $A \subset M$ is declared to be open if and only if its preimage $\pi^{-1}(A)$ is open in $\mathbb{R}^3$ (or equivalently, if $\pi^{-1}(A) \cap S$ is open in S). We say that M has the *quotient topology* relative to $\mathbb{R}^3 \setminus \{0\}$. The conditions for a Hausdorff topological space are easily verified. It follows immediately from Lemma 2.1.1 that the map $\pi\colon \mathbb{R}^3 \setminus \{0\} \to M$ is continuous, and that a map $f\colon M \to Y$ is continuous if and only if $f \circ \pi$ is continuous.

For $i = 1, 2, 3$ let O_i denote the subset $\{[x] \mid x_i \neq 0\}$ in M. It is open since $\pi^{-1}(O_i) = \{x \mid x_i \neq 0\}$ is open in $\mathbb{R}^3$. Let $U_i = \mathbb{R}^2$ and let $\sigma_i\colon U_i \to M$ be the map defined by

$$\sigma_1(u) = [(1, u_1, u_2)], \quad \sigma_2(u) = [(u_1, 1, u_2)], \quad \sigma_3(u) = [(u_1, u_2, 1)]$$

for $u \in \mathbb{R}^2$. It is continuous since it is composed by π and a continuous map $\mathbb{R}^2 \to \mathbb{R}^3$. Moreover, σ_i is a bijection of U_i onto O_i, and $M = O_1 \cup O_2 \cup O_3$.

Theorem 2.4.1. *The collection of the three maps $\sigma_i: U_i \to O_i$ forms a smooth atlas on M.*

Proof. It remains to check the following.

1) σ_i^{-1} is continuous $O_i \to \mathbb{R}^2$. For example

$$\sigma_1^{-1}(p) = (\frac{x_2}{x_1}, \frac{x_3}{x_1})$$

when $p = \pi[x]$. The ratios $\frac{x_2}{x_1}$ and $\frac{x_3}{x_1}$ are continuous functions on $\mathbb{R}^3 \setminus \{x_1 = 0\}$, hence $\sigma^{-1} \circ \pi$ is continuous.

2) The overlap between σ_i and σ_j satisfies smooth transition. For example

$$\sigma_1^{-1} \circ \sigma_2(u) = (\frac{1}{u_1}, \frac{u_2}{u_1}),$$

which is smooth $\mathbb{R}^2 \setminus \{u \mid u_1 = 0\} \to \mathbb{R}^2$. □

2.5 Product manifolds

If M and N are metric spaces, the Cartesian product $M \times N$ is again a metric space with the distance function

$$d((m_1, n_1), (m_2, n_2)) = \max(d_M(m_1, m_2), d_N(n_1, n_2)).$$

Likewise, if M and N are Hausdorff topological spaces, then the product $M \times N$ is a Hausdorff topological space in a natural fashion with the so-called *product topology*, in which a subset $R \subset M \times N$ is open if and only if for each point $(p, q) \in R$ there exist open sets P and Q of M and N respectively, such that $(p, q) \in P \times Q \subset R$ (the verification that this is a topological space is quite straightforward).

Example 2.5.1 It is sometimes useful to identify $\mathbb{R}^{m+n}$ with $\mathbb{R}^m \times \mathbb{R}^n$. In this identification, the product topology of the standard topologies on $\mathbb{R}^m$ and $\mathbb{R}^n$ is the standard topology on $\mathbb{R}^{m+n}$.

Let M and N be abstract manifolds of dimensions m and n, respectively. For each chart $\sigma: U \to M$ and each chart $\tau: V \to N$ we define

$$\sigma \times \tau: U \times V \to M \times N \quad \text{by} \quad \sigma \times \tau(x, y) = (\sigma(x), \tau(y)).$$

Theorem 2.5. *The collection of the maps $\sigma \times \tau$ is an $m + n$-dimensional smooth atlas on $M \times N$.*

Proof. The proof is straightforward. □

We call $M \times N$ equipped with the smooth structure given by this atlas for the *product manifold* of M and N. The smooth structure on $M \times N$ depends only on the smooth structures on M and N, not on the chosen atlases.

Notice that if M is a manifold in $\mathbb{R}^k$ and N is a manifold in $\mathbb{R}^l$, then we can regard $M \times N$ as a subset of $\mathbb{R}^{k+l}$ in a natural fashion. It is easily seen that this subset of $\mathbb{R}^{k+l}$ is an $m+n$-dimensional manifold (according to Definition 1.6.2), and that its differential structure is the same as that provided by Theorem 2.5, where the product is regarded as an 'abstract' set.

Example 2.5.2 The product $S^1 \times \mathbb{R}$ is an 'abstract' version of the cylinder. As just remarked, it can be regarded as a subset of $\mathbb{R}^{2+1} = \mathbb{R}^3$, and then it becomes the usual cylinder.

The product $S^1 \times S^1$, which is an 'abstract' version of the torus, is naturally regarded as a manifold in $\mathbb{R}^4$. The usual torus, which is a surface in $\mathbb{R}^3$, is not identical with this set, but there is a natural bijective correspondence.

2.6 Smooth maps in Euclidean spaces

We shall now define the important notion of a smooth map between manifolds. We first study the case of manifolds in $\mathbb{R}^n$.

Notice that the standard definition of differentiability in a point p of a map $f\colon \mathbb{R}^n \to \mathbb{R}^l$ requires f to be defined in an *open* neighborhood of p in $\mathbb{R}^n$. This definition does not make sense for a map defined on an m-dimensional manifold in $\mathbb{R}^n$, because in general a manifold is not an open subset of $\mathbb{R}^n$.

Definition 2.6.1. Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^l$ be arbitrary subsets. A map $f\colon X \to Y$ is said to be *smooth at* $p \in X$, if there exists an open set $W \subset \mathbb{R}^n$ around p and a smooth map $F\colon W \to \mathbb{R}^l$ which coincides with f on $W \cap X$. The map f is called *smooth* if it is smooth at every $p \in X$.

If f is a bijection of X onto Y, and if both f and f^{-1} are smooth, then f is called a diffeomorphism.

A smooth map F as above is called a *local smooth extension* of f. In order to show that a map defined on a subset of $\mathbb{R}^n$ is smooth, one thus has to find such a local smooth extension near every point in the domain of definition. It is easily seen that a smooth function is continuous according to Definition 2.1.4. We observe that the new notion of smoothness agrees with the standard definition when X is open in $\mathbb{R}^n$. We also observe that the smoothness of f does not depend on which subset of $\mathbb{R}^l$ is considered as the target set Y.

Definition 2.6.2. Let $\mathcal{S} \subset \mathbb{R}^n$ and $\tilde{\mathcal{S}} \subset \mathbb{R}^l$ be manifolds. A map $f\colon \mathcal{S} \to \tilde{\mathcal{S}}$ is called *smooth* if it is smooth according to Definition 2.6.1 with $X = \mathcal{S}$ and $Y = \tilde{\mathcal{S}}$.

In particular, the above definition can be applied with $\tilde{\mathcal{S}} = \mathbb{R}$. A smooth map $f\colon \mathcal{S} \to \mathbb{R}$ is said to be a *smooth function*, and the set of these is denoted $C^\infty(\mathcal{S})$. It is easily seen that $C^\infty(\mathcal{S})$ is a vector space when equipped with the standard addition and scalar multiplication of functions. Since a relatively open set $\Omega \subset \mathcal{S}$ is a manifold of its own (see Example 1.6.4), the space $C^\infty(\Omega)$

is defined for all such sets. It is easily seen that that restriction $f \mapsto f|_\Omega$ maps $C^\infty(\mathcal{S}) \to C^\infty(\Omega)$.

Example 2.6.1 The functions $x \mapsto x_i$ where $i = 1, \dots, n$ are smooth functions on $\mathbb{R}^n$. Hence they restrict to smooth functions on every manifold $\mathcal{S} \subset \mathbb{R}^n$.

Example 2.6.2 Let $\mathcal{S} \subset \mathbb{R}^2$ be the circle $\{x \mid x_1^2 + x_2^2 = 1\}$, and let $\Omega = \mathcal{S} \setminus \{(-1, 0)\}$. The function $f\colon \Omega \to \mathbb{R}$ defined by $f(x_1, x_2) = \frac{x_2}{1+x_1}$ is a smooth function, since it is the restriction of the smooth function $F\colon W \to \mathbb{R}$ defined by the same expression for $x \in W = \{x \in \mathbb{R}^2 \mid x_1 \neq -1\}$.

Example 2.6.3 Let $\mathcal{S} \subset \mathbb{R}^n$ be an m-dimensional manifold, and let $\sigma\colon U \to \mathcal{S}$ be a chart. It follows from Theorem 1.7 that σ^{-1} is smooth $\sigma(U) \to \mathbb{R}^m$.

Lemma 2.6. *Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$. If $f\colon X \to \mathbb{R}^m$ is smooth and maps into Y, and if in addition $g\colon Y \to \mathbb{R}^l$ is smooth, then so is $g \circ f\colon X \to \mathbb{R}^l$.*

Proof. Let $p \in X$ be given, and let $F\colon W \to \mathbb{R}^m$ be a local smooth extension of f around p. Likewise let $G\colon V \to \mathbb{R}^l$ be a local smooth extension of g around $f(p) \in Y$. The set $W' = F^{-1}(V)$ is an open neighborhood of p, and $G \circ (F|_{W'})$ is a local smooth extension of $g \circ f$ at p. $\square$

The definition of smoothness that we have given for manifolds in $\mathbb{R}^n$ uses the ambient space $\mathbb{R}^n$. In order to prepare for the generalization to abstract manifolds, we shall now give an alternative description.

Theorem 2.6. *Let $f\colon \mathcal{S} \to \mathbb{R}^l$ be a map. If f is smooth, then $f \circ \sigma$ is smooth for each chart σ on $\mathcal{S}$. Conversely, if $f \circ \sigma$ is smooth for each chart in some atlas of $\mathcal{S}$, then f is smooth.*

Proof. The first statement is immediate from Lemma 2.6. For the converse, assume $f \circ \sigma$ is smooth and apply Lemma 2.6 and Example 2.6.3 to $f|_{\sigma(U)} = (f \circ \sigma) \circ \sigma^{-1}$. It follows that $f|_{\sigma(U)}$ is smooth. If this is the case for each chart in an atlas, then f is smooth around all points $p \in \mathcal{S}$. $\square$

2.7 Smooth maps between abstract manifolds

Inspired by Theorem 2.6, we can now generalize to abstract manifolds.

Definition 2.7.1. Let M be an abstract manifold of dimension m. A map $f\colon M \to \mathbb{R}^l$ is called *smooth* if for every chart (σ, U) in a smooth atlas of M, the map $f \circ \sigma$ is smooth $U \to \mathbb{R}^l$.

The set U is open in $\mathbb{R}^m$, and the smoothness of $f \circ \sigma$ is in the ordinary sense for functions defined on an open set. It is easily seen that the requirement in the definition is unchanged if the atlas is replaced by a compatible one (see Definition 2.2.2), so that the notion only depends on the smooth structure of M. It follows from Theorem 2.6 that the notion is the same as before for manifolds in $\mathbb{R}^n$.

Notice that a smooth map $f\colon M \to \mathbb{R}^l$ is continuous, since in a neighborhood of each point $p \in M$ it can be written as $(f \circ \sigma) \circ \sigma^{-1}$ for a chart σ.

Notice also that if $\Omega \subset M$ is open, then Ω is an abstract manifold of its own (see Example 2.3.2), and hence it makes sense to speak of smooth maps $f\colon \Omega \to \mathbb{R}^l$. The set of all smooth functions $f\colon \Omega \to \mathbb{R}$ is denoted $C^\infty(\Omega)$. It is easily seen that this is a vector space when equipped with the standard addition and scalar multiplication of functions.

Example 2.7.1 Let $\sigma\colon U \to M$ be a chart on an abstract manifold M. It follows from the assumption of smooth transition on overlaps that σ^{-1} is smooth $\sigma(U) \to \mathbb{R}^m$.

We have defined what it means for a map *from* a manifold to be smooth, and we shall now define what smoothness means for a map *into* a manifold.

As before we begin by considering manifolds in Euclidean space. Let $\mathcal{S}$ and $\tilde{\mathcal{S}}$ be manifolds in $\mathbb{R}^n$ and $\mathbb{R}^l$, respectively, and let $f\colon \mathcal{S} \to \tilde{\mathcal{S}}$. It was defined in Definition 2.6.2 what it means for f to be smooth. We will give an alternative description.

Let $\sigma\colon U \to \mathcal{S}$ and $\tilde{\sigma}\colon \tilde{U} \to \tilde{\mathcal{S}}$ be charts on $\mathcal{S}$ and $\tilde{\mathcal{S}}$, respectively, where $U \subset \mathbb{R}^m$ and $\tilde{U} \subset \mathbb{R}^k$ are open sets. For a map $f\colon \mathcal{S} \to \tilde{\mathcal{S}}$, we call the map

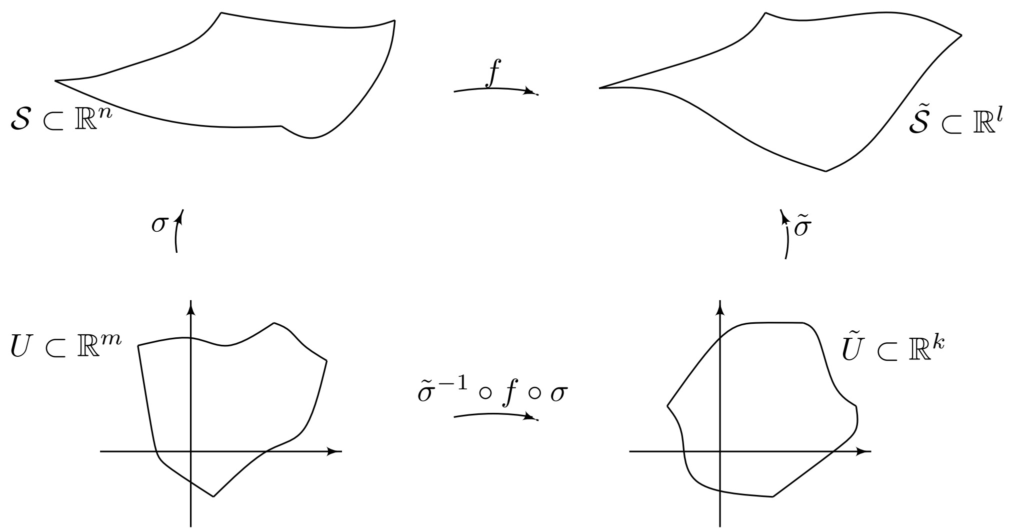
$$\tilde{\sigma}^{-1} \circ f \circ \sigma\colon\ x \mapsto \tilde{\sigma}^{-1}(f(\sigma(x))), \tag{2.1}$$

the *coordinate expression* for f with respect to the charts.

The coordinate expression (2.1) is defined for all $x \in U$ for which $f(\sigma(x)) \in \tilde{\sigma}(\tilde{U})$, that is, it is defined on the set

$$\sigma^{-1}(f^{-1}(\tilde{\sigma}(\tilde{U}))) \subset U, \tag{2.2}$$

and it maps into $\tilde{U}$.



It follows from Corollary 1.7 that $\tilde\sigma(\tilde U)$ is open in $\tilde{\mathcal S}$. Hence if f is continuous, then $f^{-1}(\tilde\sigma(\tilde U))$ is open in $\mathcal S$ by Lemma 2.1.1. Since σ is continuous, the set (2.2), on which $\tilde\sigma^{-1}\circ f\circ\sigma$ is defined, is then open.

Theorem 2.7. *Let $f\colon\mathcal S\to\tilde{\mathcal S}$ be a map. If f is smooth (according to Definition 2.6.2) then it is continuous and $\tilde\sigma^{-1}\circ f\circ\sigma$ is smooth, for all charts σ and $\tilde\sigma$ on $\mathcal S$ and $\tilde{\mathcal S}$, respectively.*

Conversely, assume that for each $p\in\mathcal S$ there exists a chart $\sigma\colon U\to\mathcal S$ around p, and a chart $\tilde\sigma\colon\tilde U\to\tilde{\mathcal S}$ around $f(p)$, such that $f(\sigma(U))\subset\tilde\sigma(\tilde U)$ and such that the coordinate expression $\tilde\sigma^{-1}\circ f\circ\sigma$ is smooth, then f is smooth.

Proof. Assume that f is smooth. It was remarked below Definition 2.6.1 that then it is continuous. Hence (2.2) is open. It follows from Theorem 2.6 that $f\circ\sigma$ is smooth for all charts σ on $\mathcal S$. Hence its restriction to the set (2.2) is also smooth, and it follows from Lemma 2.6 and Example 2.6.3 that the composed map $\tilde\sigma^{-1}\circ f\circ\sigma$ is smooth.

For the converse let $p\in\mathcal S$ be arbitrary and let σ and $\tilde\sigma$ be as stated, such that $\tilde\sigma^{-1}\circ f\circ\sigma$ is smooth. The identity

$$f\circ\sigma=\tilde\sigma\circ(\tilde\sigma^{-1}\circ f\circ\sigma),$$

shows that $f\circ\sigma$ is smooth. Since the charts σ for all p comprise an atlas for $\mathcal S$ this implies that f is smooth, according to Theorem 2.6. □

By using the formulation of smoothness in Theorem 2.7, we can now generalize the notion. Let M and $\tilde M$ be abstract manifolds, and let $f\colon M\to\tilde M$ be a continuous map. Assume $\sigma\colon U\to\sigma(U)\subset M$ and $\tilde\sigma\colon\tilde U\to\tilde\sigma(\tilde U)\subset\tilde M$ are charts on the two manifolds, then as before

$$\sigma^{-1}(f^{-1}(\tilde\sigma(\tilde U)))\subset U$$

is an open subset of U, because f is continuous. Again we call the map

$$\tilde\sigma^{-1}\circ f\circ\sigma,$$

which is defined on this set, the coordinate expression for f with respect to the given charts.

Definition 2.7.2. Let $f\colon M\to\tilde M$ be a map between abstract manifolds. Then f is called *smooth* if for each $p\in\mathcal S$ there exists a chart $\sigma\colon U\to M$ around p, and a chart $\tilde\sigma\colon\tilde U\to\tilde M$ around $f(p)$, such that $f(\sigma(U))\subset\tilde\sigma(\tilde U)$ and such that the coordinate expression $\tilde\sigma^{-1}\circ f\circ\sigma$ is smooth.

A bijective map $f\colon M\to\tilde M$, is called a *diffeomorphism* if f and f^{-1} are both smooth.

Notice that a smooth map $M\to\tilde M$ is continuous. This follows immediately from the definition above, by writing $f=\tilde\sigma\circ(\tilde\sigma^{-1}\circ f\circ\sigma)\circ\sigma^{-1}$ in a neighborhood of each point.

Again it should be checked that the notions are independent of the atlases from which the charts are chosen, as long as each atlas is replaced by a compatible one. The verification of this fact is straightforward.

It follows from Theorem 2.7 that the notion of smoothness is the same as before if $M = \mathcal{S} \subset \mathbb{R}^n$ and $\tilde{M} = \tilde{\mathcal{S}} \subset \mathbb{R}^l$. Likewise, there is no conflict with Definition 2.7.1 in case $\tilde{M} = \mathbb{R}^l$, where in Definition 2.7.2 we can use the identity map for $\tilde{\sigma}: \mathbb{R}^l \to \mathbb{R}^l$.

It is easily seen that the composition of two smooth maps between abstract manifolds is again smooth.

Example 2.7.2 Let M, N be finite dimensional vector spaces of dimension m and n, respectively. These are abstract manifolds, according to Example 2.3.1. Let $f: M \to N$ be a linear map. If we choose a basis for each space and define the corresponding charts as in Example 2.3.1, then the coordinate expression for f is a linear map from $\mathbb{R}^m$ to $\mathbb{R}^n$ (given by the matrix that represents f), hence smooth. It follows from Definition 2.7.2 that f is smooth. If f is bijective, its inverse is also linear, and hence in that case f is a diffeomorphism.

Example 2.7.3 Let $\sigma: U \to M$ be a chart on an abstract m-dimensional manifold M. It follows from the assumption of smooth transition on overlaps that σ is smooth $U \to M$, if we regard U as an m-dimensional manifold (with the identity chart). By combining this observation with Example 2.7.1, we see that σ is a diffeomorphism of U onto its image $\sigma(U)$ (which is open in M by Definition 2.2.1).

Conversely, every diffeomorphism g of a non-empty open subset $V \subset \mathbb{R}^m$ onto an open subset in M is a chart on M. Indeed, by the definition of a chart given at the end of Section 2.2, this means that g should overlap smoothly with all charts σ in an atlas of M, that is $g^{-1} \circ \sigma$ and $\sigma^{-1} \circ g$ should both be smooth (on the sets where they are defined). This follows from the preceding observation about compositions of smooth maps.

2.8 Lie groups

Definition 2.8.1. A *Lie group* is a group G, which is also a manifold, such that the group operations

$$(x, y) \mapsto xy, \qquad x \mapsto x^{-1},$$

are smooth maps from $G \times G$, respectively G, into G.

Example 2.8.1 Every finite dimensional real vector space V is a group, with the addition of vectors as the operation and with neutral element 0. The map $(x, y) \mapsto x + y$ is linear $V \times V \to V$, hence it is smooth (see Example 2.7.2). Likewise $x \mapsto -x$ is smooth and hence V is a Lie group.

Example 2.8.2 The set $\mathbb{R}^\times$ of non-zero real numbers is a 1-dimensional Lie group, with multiplication as the operation and with neutral element 1. Likewise the set $\mathbb{C}^\times$ of non-zero complex numbers is a 2-dimensional Lie group, with complex multiplication as the operation. The smooth structure is determined by the chart $(x_1, x_2) \mapsto x_1 + ix_2$. The product

$$xy = (x_1 + ix_2)(y_1 + iy_2) = x_1y_1 - x_2y_2 + i(x_1y_2 + x_2y_1)$$

is a smooth function of the entries, and so is the inverse

$$x^{-1} = (x_1 + ix_2)^{-1} = \frac{x_1 - ix_2}{x_1^2 + x_2^2}.$$

Example 2.8.3 Let $G = \mathrm{SO}(2)$, the group of all 2×2 real matrices which are orthogonal, that is, they satisfy the relation $AA^t = I$, and which have determinant 1. The set G is in one-to-one correspondence with the unit circle in $\mathbb{R}^2$ by the map

$$(x_1, x_2) \mapsto \begin{pmatrix} x_1 & x_2 \\ -x_2 & x_1 \end{pmatrix}.$$

If we give G the smooth structure so that this map is a diffeomorphism, then it becomes a 1-dimensional Lie group, called the *circle group*. The multiplication of matrices is given by a smooth expression in x_1 and x_2, and so is the inversion $x \mapsto x^{-1}$, which only amounts to a change of sign on x_2.

Example 2.8.4 Let $G = \mathrm{GL}(n, \mathbb{R})$, the set of all invertible $n \times n$ matrices. It is a group, with matrix multiplication as the operation. It is a manifold in the following fashion. The set $\mathrm{M}(n, \mathbb{R})$ of all real $n \times n$ matrices is in bijective correspondence with $\mathbb{R}^{n^2}$ and is therefore a manifold of dimension n^2. The subset $G = \{A \in \mathrm{M}(n, \mathbb{R}) \mid \det A \neq 0\}$ is an open subset, because the determinant function is continuous. Hence G is a manifold.

Furthermore, the matrix multiplication $\mathrm{M}(n, \mathbb{R}) \times \mathrm{M}(n, \mathbb{R}) \to \mathrm{M}(n, \mathbb{R})$ is given by smooth expressions in the entries (involving products and sums), hence it is a smooth map. It follows that the restriction to $G \times G$ is also smooth.

Finally, the map $x \mapsto x^{-1}$ is smooth $G \to G$, because according to Cramer's formula the entries of the inverse x^{-1} are given by rational functions in the entries of x (with the determinant in the denominator). It follows that $G = \mathrm{GL}(n, \mathbb{R})$ is a Lie group.

Example 2.8.5 Let G be an arbitrary group, equipped with the *discrete topology* (see Example 2.3.4). It is a 0-dimensional Lie group.

Theorem 2.8. *Let $G \subset \mathrm{GL}(n, \mathbb{R})$ be a subgroup which is also a manifold in $\mathbb{R}^{n^2}$. Then G is a Lie group.*

Proof. It has to be shown that the multiplication is smooth $G \times G \to G$. According to Definition 2.6.1 we need to find local smooth extensions of the multiplication map and the inversion map. This is provided by Example 2.8.4, since $\mathrm{GL}(n,\mathbb{R})$ is open in $\mathbb{R}^{n^2}$. $\square$

Example 2.8.6 The group $\mathrm{SL}(2,\mathbb{R})$ of 2×2-matrices of determinant 1 is a 3-dimensional manifold in $\mathbb{R}^4$, since Theorem 1.6 can be applied with f equal to the determinant function $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$. Hence it is a Lie group.

2.9 Countable atlas

The definition of an atlas of a smooth manifold leaves no limitation on the size of the family of charts. Sometimes it is useful that the atlas is not too large. In this section we introduce an assumptions of this nature. In particular, we shall see that all manifolds in $\mathbb{R}^n$ satisfy the assumption.

Recall that a set A is said to be *countable* if it is finite or in one-to-one correspondence with $\mathbb{N}$.

Definition 2.9. An atlas of an abstract manifold M is said to be *countable* if the set of charts in the atlas is countable.

In a topological space X, a *base for the topology* is a collection of open sets V with the property that for every point $x \in X$ and every neighborhood U of x there exists a member V in the collection, such that $x \in V \subset U$. For example, in a metric space the collection of all open balls is a base for the topology.

Example 2.9 As just mentioned, the collection of all open balls $B(y, r)$ is a base for the topology of $\mathbb{R}^n$. In fact, the subcollection of all balls, for which the radius as well as the coordinates of the center are rational, is already a base for the topology. For if $x \in \mathbb{R}^n$ and a neighborhood U is given, there exists a rational number $r > 0$ such that $B(x, r) \subset U$. By density of the rationals, there exists a point $y \in B(x, r/2)$ with rational coordinates, and by the triangle inequality $x \in B(y, r/2) \subset B(x, r) \subset U$. We conclude that in $\mathbb{R}^n$ there is a *countable* base for the topology. The same is then true for any subset $X \subset \mathbb{R}^n$, since the collection of intersections with X of elements from a base for $\mathbb{R}^n$, is a base for the topology in X.

Lemma 2.9. *Let M be an abstract manifold. Then M has a countable atlas if and only if there exists a countable base for the topology.*

A topological space, for which there exists a countable base, is said to be *second countable.*

Proof. Assume that M has a countable atlas. For each chart $\sigma: U \to M$ in the atlas there is a countable base for the topology of U, according to the example above, and since σ is a homeomorphism it carries this to a countable base for $\sigma(U)$. The collection of all these sets for all the charts in the atlas, is then a countable base for the topology of M, since a countable union of countable sets is again countable.

Assume conversely that there is a countable base $(V_k)_{k\in I}$ for the topology. For each $k \in I$, we select a chart $\sigma: U \to M$ for which $V_k \subset \sigma(U)$, if such a chart exists. The collection of selected charts is clearly countable. It covers M, for if $x \in M$ is arbitrary, there exists a chart σ (not necessarily among the selected) around x, and there exists a member V_k in the base with $x \in V_k \subset \sigma(U)$. This member V_k is contained in a chart, hence also in a selected chart, and hence so is x. Hence the collection of selected charts is a countable atlas. □

Corollary 2.9. *Let $\mathcal{S}$ be a manifold in $\mathbb{R}^n$. There exists a countable atlas for $\mathcal{S}$.*

Proof. According to Example 2.9 there is a countable base for the topology. □

In Example 2.3.4 we introduced a 0-dimensional smooth structure on an arbitrary set X, with the discrete topology. Any basis for the topology must contain all singleton sets in X. Hence if the set X is not countable, there does not exist any countable atlas for this manifold.

2.10 Whitney's theorem

The following is a famous theorem, due to Whitney. The proof is too difficult to be given here. In Section 5.6 we shall prove a weaker version of the theorem.

Theorem 2.10. *Let M be an abstract smooth manifold of dimension m, and assume there exists a countable atlas for M. Then there exists a diffeomorphism of M onto a manifold in $\mathbb{R}^{2m}$.*

For example, the projective space $\mathbb{R}\mathrm{P}^2$ is diffeomorphic with a 2-dimensional manifold in $\mathbb{R}^4$. Notice that the assumption of a countable atlas cannot be avoided, due to to Corollary 2.9.

The theorem could give one the impression that if we limit our interest to manifolds with a countable atlas, then the notion of abstract manifolds is superfluous. This is not so, because in many circumstances it would be very inconvenient to be forced to perceive a particular smooth manifold as a subset of some high-dimensional $\mathbb{R}^k$. The abstract notion frees us from this limitation.

Chapter 3

The tangent space

In this chapter the tangent space for an abstract manifold is defined. Let us recall from Geometry 1 that for a parametrized surface $\sigma: U \to \mathbb{R}^3$ we defined the tangent space at point $x \in U$ as the linear subspace in $\mathbb{R}^3$ spanned by the derived vectors σ'_u and σ'_v at x. The generalization to $\mathbb{R}^n$ is straightforward, it is given below. In this chapter we shall generalize the concept to abstract manifolds. We shall do so in two steps. The first step is to consider manifolds in $\mathbb{R}^n$. Here we can still define the tangent space with reference to the ambient space $\mathbb{R}^n$. For the abstract manifolds, treated in the second step, we need to give an abstract definition.

3.1 The tangent space of a parametrized manifold

Let $\sigma: U \to \mathbb{R}^n$ be a parametrized manifold, where $U \subset \mathbb{R}^m$ is open.

Definition 3.1. The *tangent space* $T_{x_0}\sigma$ of σ at $x_0 \in U$ is the linear subspace of $\mathbb{R}^n$ spanned by the columns of the $n \times m$ matrix $D\sigma(x_0)$.

Assume that σ is regular at x_0. Then the the columns of $D\sigma(x_0)$ are linearly independent, and they form a basis for the space $T_{x_0}\sigma$. If $v \in \mathbb{R}^m$ then $D\sigma(x_0)v$ is the linear combination of these columns with the coordinates of v as coefficients, hence $v \mapsto D\sigma(x_0)v$ is a linear isomorphism of $\mathbb{R}^m$ onto the m-dimensional tangent space $T_{x_0}\sigma$.

Observe that the tangent space is a 'local' object, in the sense that if two parametrized manifolds $\sigma: U \to \mathbb{R}^n$ and $\sigma': U' \to \mathbb{R}^n$ are equal on some neighborhood of a point $x_0 \in U \cap U'$, then $T_{x_0}\sigma = T_{x_0}\sigma'$.

From Geometry 1 we recall the following result, which is easily generalized to the present setting.

Theorem 3.1. *The tangent space is invariant under reparametrization. In other words, if $\phi: W \to U$ is a diffeomorphism of open sets in $\mathbb{R}^m$ and $\tau = \sigma \circ \phi$, then $T_{y_0}\tau = T_{\phi(y_0)}\sigma$ for all $y_0 \in W$.*

Proof. The proof is essentially the same as in Geometry 1, Theorem 2.7. By the chain rule, we have the matrix identity

$$D\tau(y_0) = D\sigma(\phi(y_0))D\phi(y_0),$$

which implies that the columns of $D\tau(y_0)$ are linear combinations of the columns of $D\sigma(\phi(y_0))$, hence $T_{y_0}\tau \subset T_{\phi(y_0)}\sigma$. The opposite inclusion follows by the same argument with ϕ replaced by its inverse. $\square$

3.2 The tangent space of a manifold in $\mathbb{R}^n$

Let $\mathcal{S} \subset \mathbb{R}^n$ be a manifold (see Definition 1.6.1).

Definition 3.2. The *tangent space* of σ at a point $p \in \mathcal{S}$ is the m-dimensional linear space

$$T_p\mathcal{S} = T_{x_0}\sigma \subset \mathbb{R}^n, \tag{3.1}$$

where $\sigma: U \to \mathcal{S}$ is an arbitrary chart on $\mathcal{S}$ with $p = \sigma(x_0)$ for some $x_0 \in U$.

Notice that it follows from Theorem 3.1 that the tangent space $T_p\mathcal{S} = T_{x_0}\sigma$ is actually independent of the chart σ, since any other chart around p will be a reparametrization, according to Theorem 1.8 (more precisely, it is the restriction to the overlap of the two charts, which is a reparametrization). Notice also the change of notation, in that the footpoint p of $T_p\mathcal{S}$ belongs to $\mathcal{S} \subset \mathbb{R}^n$ in contrast to the previous footpoint $x_0 \in \mathbb{R}^m$ of $T_{x_0}\sigma$.

With the generalization to abstract manifolds in mind we shall give a different characterization of the tangent space $T_p\mathcal{S}$, which is also closely related to a property from Geometry 1 (Theorem 2.4).

We call a parametrized smooth curve $\gamma: I \to \mathbb{R}^n$ a *parametrized curve on* $\mathcal{S}$ if the image $\gamma(I)$ is contained in $\mathcal{S}$.

Theorem 3.2. *Let $p \in \mathcal{S}$. The tangent space $T_p\mathcal{S}$ is the set of all vectors in $\mathbb{R}^n$, which are tangent vectors $\gamma'(t_0)$ to a parametrized curve on $\mathcal{S}$, $\gamma: I \to \mathcal{S}$, with $\gamma(t_0) = p$ for some $t_0 \in I$.*

Proof. Let $\sigma: U \to \mathcal{S}$ be a chart with $p = \sigma(x_0)$ for some $x_0 \in U$. By definition $T_p\mathcal{S} = T_{x_0}\sigma$. As remarked below Definition 3.1, the map $v \mapsto D\sigma(x_0)v$ is a linear isomorphism of $\mathbb{R}^m$ onto $T_{x_0}\sigma$. For each $v \in \mathbb{R}^m$ we define a parametrized curve on $\mathcal{S}$ by

$$\gamma_v(t) = \sigma(x_0 + tv)$$

for t close to 0 (so that $x_0 + tv \in U$). It then follows from the chain rule that

$$\gamma_v{}'(0) = D\sigma(x_0)v,$$

so we conclude that the map

$$T^\sigma: \quad \mathbb{R}^m \to T_p\mathcal{S}, \quad v \mapsto \gamma_v{}'(0) \tag{3.2}$$

is a linear isomorphism. In particular, we see that every vector in $T_p\mathcal{S}$ is a tangent vector of a parametrized curve on $\mathcal{S}$.

Conversely, given a parametrized curve $\gamma: I \to \mathbb{R}^n$ on $\mathcal{S}$ with $\gamma(t_0) = p$, we define a parametrized curve μ in $\mathbb{R}^m$ by $\mu = \sigma^{-1} \circ \gamma$ for t in a neighborhood of t_0. This is the *coordinate expression* for γ with respect to σ. It follows

from Lemma 1.8 that μ is smooth, and since $\gamma = \sigma \circ \mu$, it follows from the chain rule that
$$\gamma'(t_0) = D\sigma(x_0)\mu'(t_0) \in T_{x_0}\sigma.$$
This proves the opposite inclusion. □

We have described the tangent space as a space of tangent vectors to curves on $\mathcal{S}$. However, it should be emphasized that the correspondence between curves and tangent vectors is not one-to-one. In general many different curves give rise to the same tangent vector in $\mathbb{R}^n$.

Example 3.2.1 Let $\Omega \subset \mathbb{R}^n$ be an open subset, then Ω is an n-dimensional manifold in $\mathbb{R}^n$, with the identity map as a chart (see Example 1.6.3). The tangent space is $\mathbb{R}^n$, and the map (3.2) is the identity map, since the derivative of $t \mapsto x_0 + tv$ at $t = 0$ is v.

Example 3.2.2 Suppose $\mathcal{S} \subset \mathbb{R}^n$ is the $n-k$-dimensional manifold given by an equation $f(p) = c$ as in Theorem 1.6, where in addition it is required that $Df(p)$ has rank k for all $p \in \mathcal{S}$. Then we claim that the tangent space $T_p\mathcal{S}$ is the null space for the matrix $Df(p)$, that is,
$$T_p\mathcal{S} = \{v \in \mathbb{R}^n \mid Df(p)v = 0\}.$$
Here we recall that by a fundamental theorem of linear algebra, the dimension of the null space is $n-k$ when the rank of $Df(p)$ is k.

The claim can be established as follows by means of Theorem 3.2. If v is a tangent vector, then $v = \gamma'(t_0)$ for some parametrized curve on $\mathcal{S}$ with $p = \gamma(t_0)$. Since γ maps into $\mathcal{S}$, the function $f \circ \gamma$ is constant (with the value c), hence $(f \circ \gamma)'(t_0) = 0$. It follows from the chain rule that $Df(p)v = 0$. Thus the tangent space is contained in the null space. Since both are linear spaces of dimension $n-k$, they must be equal.

Example 3.2.3 Let $\mathcal{C}$ denote the ∞-shaped set in Examples 1.2.2 and 1.3.5. The parametrized curve $\gamma(t) = (\cos t, \cos t \sin t)$ is a parametrized curve on $\mathcal{C}$, and it passes through $(0,0)$ for each $t = k\frac{\pi}{2}$ with k an odd integer. If $\mathcal{C}$ was a curve in $\mathbb{R}^2$, then $\gamma'(k\frac{\pi}{2})$ would belong to $T_{(0,0)}\mathcal{C}$, according to Theorem 3.2. But the vectors $\gamma'(\frac{\pi}{2}) = (-1,-1)$ and $\gamma'(3\frac{\pi}{2}) = (1,-1)$ are linearly independent, and hence $T_{(0,0)}\mathcal{C}$ cannot be a 1-dimensional linear space. Thus we reach a contradiction, and we conclude that $\mathcal{C}$ is *not* a curve (as mentioned in Example 1.3.5).

3.3 The abstract tangent space

Consider an abstract m-dimensional manifold M, and let $\sigma: U \to M$ be a chart. It does not make sense to repeat Definition 3.1 of the tangent space to σ, because the $n \times m$-matrix $D\sigma(x_0)$ is not defined (there is no n). Hence it does not make sense to repeat Definition 3.2 either. Instead, Theorem 3.2 will be our inspiration for the definition of the tangent space T_pM.

We define a *parametrized curve on M* to be a smooth map $\gamma: I \to M$, where $I \subset \mathbb{R}$ is open. By definition (see Definition 2.7.2), this means that $\sigma^{-1} \circ \gamma$ is smooth for all charts on M. As before, if σ is a chart on M, we call $\mu = \sigma^{-1} \circ \gamma$ the *coordinate expression* of γ with respect to σ. If $p \in M$ is a point, a parametrized curve on M *through* p, is a parametrized curve on M together with a point $t_0 \in I$ for which $p = \gamma(t_0)$.

Let $\gamma_1: I_1 \to M$ and $\gamma_2: I_2 \to M$ be parametrized curves on M with $p = \gamma_1(t_1) = \gamma_2(t_2)$, and let σ be a chart around p, say $p = \sigma(x)$. We say that γ_1 and γ_2 are *tangential* at p, if the coordinate expressions satisfy

$$(\sigma^{-1} \circ \gamma_1)'(t_1) = (\sigma^{-1} \circ \gamma_2)'(t_2). \tag{3.3}$$

Lemma 3.3.1. *Being tangential at p is an equivalence relation on curves through p. It is independent of the chosen chart σ.*

Proof. The first statement is easily seen. If $\tilde{\sigma}$ is another chart then the coordinate expressions are related by

$$\tilde{\sigma}^{-1} \circ \gamma_i = (\tilde{\sigma}^{-1} \circ \sigma) \circ (\sigma^{-1} \circ \gamma_i)$$

on the overlap. The chain rule implies

$$(\tilde{\sigma}^{-1} \circ \gamma_i)'(t_i) = D(\tilde{\sigma}^{-1} \circ \sigma)(x)\,(\sigma^{-1} \circ \gamma_i)'(t_i)$$

for each of the curves. Hence the relation (3.3) implies the similar relation with σ replaced by $\tilde{\sigma}$. $\square$

We write the equivalence relation as $\gamma_1 \sim_p \gamma_2$, and we interprete it as an abstract equality between the 'tangent vectors' of the two curves at p. In spite of the fact that we have not yet defined the tangent vector of a curve on M, we have thus accomplished to make sense out of the statement that that two curves have the same tangent vector. This is exactly what is needed in order to give an abstract version of Definition 3.2.

Definition 3.3. The *tangent space* T_pM is the set of $\sim_p$-classes of parametrized curves on M through p.

The following observation is of importance, because it shows that the abstract tangent space is a 'local' object, that is, it only depends on the structure of M in a vicinity of p. If $M' \subset M$ is an open subset, then M' is an abstract manifold in itself, according to Example 2.3.2. Obviously a parametrized curve on M' can also be regarded as a parametrized curve on M, and the notion of two curves being tangential at a point $p \in M'$ is independent of how we regard the curves. It follows that there is a natural inclusion of T_pM' in T_pM. Conversely, a parametrized curve $\gamma: I \to M$ through p is tangential at p to its own restriction $\gamma|_{I'}$ where $I' = \gamma^{-1}(M')$, and the latter is also a parametrized curve on M'. It follows that $T_pM' = T_pM$.

We need to convince ourselves that the abstractly defined tangent space agrees with the tangent space of Definition 3.2 in case $M = \mathcal{S} \subset \mathbb{R}^n$. This is the content of the following lemma.

Lemma 3.3.2. *Let $\mathcal{S}$ be a manifold in $\mathbb{R}^n$ and let γ_1, γ_2 be curves on $\mathcal{S}$ through p. Then $\gamma_1 \sim_p \gamma_2$ if and only if $\gamma_1'(t_1) = \gamma_2'(t_2)$.*

It follows that the map $\gamma \mapsto \gamma'(t_0)$ induces a bijective map from $\sim_p$-classes of parametrized curves onto the tangent space $T_p\mathcal{S}$ of Definition 3.2. Hence the abstract tangent space is in bijective correspondence with the old one.

Proof. Choose a chart σ around p, and let $\mu_i = \sigma^{-1} \circ \gamma_i$ be the coordinate expression for γ_i. Then by definition $\gamma_1 \sim_p \gamma_2$ means that $\mu_1'(t_1) = \mu_2'(t_2)$.

By applying the chain rule to the expressions $\gamma_i = \sigma \circ \mu_i$ and $\mu_i = \sigma^{-1} \circ \gamma_i$ we see that $\mu_1'(t_1) = \mu_2'(t_2)$ if and only if $\gamma_1'(t_1) = \gamma_2'(t_2)$. We have used Theorem 1.7 in order to be able to differentiate σ^{-1}. $\square$

3.4 The vector space structure

The definition we have given of T_pM has some disadvantages. In particular, it is not clear at all how to organize T_pM as a linear space. The elements of T_pM are classes of curves, and they are not easily seen as tangent *vectors* for any reasonable vector space structure. Thus it is in fact premature to denote T_pM a 'space'. We will remedy this by exhibiting an alternative description of T_pM, analogous to (3.2).

Theorem 3.4. *Let $\sigma\colon U \to M$ be a chart on M with $p = \sigma(x_0)$.*

(i) For each element $v \in \mathbb{R}^m$ let $\gamma_v(t) = \sigma(x_0 + tv)$ for t close to 0. The map

$$T^\sigma\colon v \mapsto [\gamma_v]_p$$

is a bijection of $\mathbb{R}^m$ onto T_pM. The inverse map is given by

$$[\gamma]_p \mapsto (\sigma^{-1} \circ \gamma)'(t_0),$$

for each curve γ on M with $\gamma(t_0) = p$.

(ii) There exists a unique structure on T_pM as a (necessarily m-dimensional) vector space such that the map T^σ is a linear isomorphism. This structure is independent of the chosen chart σ around p.

Proof. (i) Notice that $\sigma^{-1} \circ \gamma_v(t) = x_0 + tv$, and hence $(\sigma^{-1} \circ \gamma_v)'(0) = v$. It follows that if $\gamma_{v_1} \sim_p \gamma_{v_2}$ then $v_1 = v_2$. Hence T^σ is injective. To see that it is surjective, let an element $[\gamma]_p \in T_pM$ be given, with representative γ such that $\gamma(t_0) = p$. Let $\sigma^{-1} \circ \gamma$ be the coordinate expression of γ with respect to σ, and put $v = (\sigma^{-1} \circ \gamma)'(t_0)$. Then $\gamma_v \sim_p \gamma$ because $(\sigma^{-1} \circ \gamma_v)'(0) = (\sigma^{-1} \circ \gamma)'(t_0) = v$. Hence $T^\sigma(v) = [\gamma]_p$, and the map is surjective.

(ii) By means of the bijection $T^\sigma \colon v \mapsto [\gamma_v]_p$ we can equip T_pM as an m-dimensional vector space. We simply *impose* the structure on T_pM so that this map is an isomorphism, that is, we define addition and scalar multiplication by the requirement that $aT^\sigma(v) + bT^\sigma(w) = T^\sigma(av + bw)$ for all $v, w \in \mathbb{R}^m$ and $a, b \in \mathbb{R}$.

Only the independency of σ remains to be seen. Suppose that $\tilde{\sigma}$ is another chart on M around p. We will prove that

$$T^\sigma(v) = T^{\tilde{\sigma}}(D(\tilde{\sigma}^{-1} \circ \sigma)(x_0)v) \tag{3.4}$$

for all $v \in \mathbb{R}^m$. Let $\gamma(t) = \gamma_v(t) = \sigma(x_0 + tv)$. By the chain rule

$$(\tilde{\sigma}^{-1} \circ \gamma)'(0) = \frac{d}{dt}(\tilde{\sigma}^{-1} \circ \sigma)(x_0 + tv)|_{t=0} = D(\tilde{\sigma}^{-1} \circ \sigma)(x_0)v.$$

Applying the last statement of (i), for the chart $\tilde{\sigma}$, we conclude that

$$[\gamma]_p = T^{\tilde{\sigma}}((\tilde{\sigma}^{-1} \circ \gamma)'(0)) = T^{\tilde{\sigma}}(D(\tilde{\sigma}^{-1} \circ \sigma)(x_0)v),$$

from which (3.4) follows.

Since multiplication by $D(\tilde{\sigma}^{-1} \circ \sigma)(x_0)$ is a linear isomorphism of $\mathbb{R}^m$, it follows from (3.4) that $T^{\tilde{\sigma}}$ is a linear map if and only if T^σ is a linear map. Hence the two bijections induce the same linear structure on T_pM. $\square$

This theorem is helpful for our understanding of the tangent space. Based on it we can visualize the tangent space T_pM as a 'copy' of $\mathbb{R}^m$ which is attached to M with its base point at p.

We have introduced a linear structure on T_pM, but if M happens to be a manifold $\mathcal{S} \subset \mathbb{R}^n$ then $T_p\mathcal{S}$ already has such a structure, inherited from the ambient $\mathbb{R}^n$. In fact, it follows from the proof of Theorem 3.2, see (3.2), that the two structures agree in this case.

Example 3.4.1 Let V be an m-dimensional real vector space regarded as an abstract manifold as in Example 2.3.1. We will show that for each element $p \in V$ there exists a natural isomorphism $L \colon V \to T_pV$. For each vector $r \in V$ we define $L(r)$ to be the $\sim_p$-class of the line $p + tr$ with direction r, regarded as a curve through p. In order to prove that L is an isomorphism of vector spaces, we choose a basis $v_1, \dots, v_m$ for V, and let $\sigma \colon \mathbb{R}^m \to V$ denote the corresponding isomorphism $(x_1, \dots, x_m) \mapsto x_1v_1 + \cdots + x_mv_m$, which constitutes a chart on V (see Example 2.3.1). For $x \in \mathbb{R}^m$ and $r = \sigma(x)$ we have, with the notation of Theorem 3.4,

$$p + tr = p + t\sigma(x) = \sigma(x_0 + tx) = \gamma_x(t)$$

and hence $T^\sigma(x) = [\gamma_x]_p = L(r)$. Since T^σ is an isomorphism of vector spaces $\mathbb{R}^m \to T_pV$, and since $x \mapsto r = \sigma(x)$ is an isomorphism $\mathbb{R}^m \to V$, we conclude that $r \mapsto L(r)$ is an isomorphism $V \to T_pV$.

3.5 Directional derivatives

An important property of tangent vectors is that they can be brought to act on smooth functions as a kind of differentiation operators.

We recall the following concept from multivariable calculus. For a function $f\colon \Omega \to \mathbb{R}$ on an open set $\Omega \subset \mathbb{R}^n$, which is differentiable at a point $p \in \Omega$, we define the *directional derivative* in direction $X \in \mathbb{R}^n$ as the derivative at 0 of the function $t \mapsto f(p + tX)$, that is, we differentiate f along the line through p with direction X. We denote the directional derivative as follows:

$$\mathrm{D}_X(f) = \frac{d}{dt}\Big|_{t=0} f(p + tX). \tag{3.5}$$

The directional derivatives in the directions of the canonical basis vectors $e_1, \dots, e_n$ are the partial derivatives

$$\mathrm{D}_{e_i} = \frac{\partial f}{\partial x_i}(p)$$

and the general directional derivative is related to these partial derivatives through the formula

$$\mathrm{D}_X(f) = \sum_{i=1}^{n} a_i \frac{\partial f}{\partial x_i}(p), \tag{3.6}$$

when $X = (a_1, \dots, a_n)$. In particular, we see that $\mathrm{D}_X(f)$ depends linearly on X.

We shall now extend this concept to manifolds in $\mathbb{R}^n$. Let $\mathcal{S} \subset \mathbb{R}^n$ be a manifold and $p \in \mathcal{S}$ a point. Recall from Definition 2.6.1 that a function $f\colon \mathcal{S} \to \mathbb{R}$ is called smooth at p if it has a local smooth extension $F\colon W \to \mathbb{R}$ in a neighborhood $W \subset \mathbb{R}^n$ of p.

Lemma 3.5. *Let $f\colon \mathcal{S} \to \mathbb{R}$ be smooth at p, and let $F\colon W \to \mathbb{R}$ be a local smooth extension. Let $X \in T_p\mathcal{S}$ be a tangent vector, and let $\gamma\colon I \to \mathcal{S}$ be a parametrized curve on $\mathcal{S}$ with $\gamma(t_0) = p$ and $\gamma'(t_0) = X$. Then*

$$(f \circ \gamma)'(t_0) = \mathrm{D}_X F. \tag{3.7}$$

Proof. Since γ maps into $\mathcal{S}$ we have $f \circ \gamma = F \circ \gamma$ in a neighborhood of t_0, and it follows from the chain rule that $(F \circ \gamma)'(t_0) = DF(p)\gamma'(t_0)$, where $DF(p)$ is the $1 \times n$ Jacobian matrix of F at p. Since $\gamma'(t_0) = X = (X_1, \dots, X_n)$ we thus obtain

$$(f \circ \gamma)'(t_0) = \sum_{i=1}^{n} \frac{\partial F}{\partial x_i}(p)\, X_i, \tag{3.8}$$

and this is exactly $\mathrm{D}_X F$ according to (3.6) above. $\square$

Observe that the left side of (3.7) is independent of the choice of local smooth extension F of f, and the other side of is independent of the choice of curve γ with $\gamma'(t_0) = X$. We conclude that both sides only depend on f and X, so that the following definition is valid.

Definition 3.5. The *directional derivative* of f at p in direction $X \in T_p\mathcal{S}$ is

$$\mathrm{D}_X(f) = (f \circ \gamma)'(t_0),$$

where γ is chosen as above.

Notice that a tangent vector $X \in T_p\mathcal{S}$ is uniquely determined through its action D_X on smooth functions. For each $i = 1, \dots, n$ we can determine the i-th coordinate of X by letting it act on the function $f(x) = x_i$ (see Example 2.6.1), since it follows from (3.8) that $\mathrm{D}_X f = X_i$ for this function. For this reason it is quite common to identify X with the operator D_X and simply write Xf in place of $\mathrm{D}_X f$.

Example 3.5 Let $\mathcal{S}$ be the unit sphere in $\mathbb{R}^3$, let $p = (0,0,1) \in \mathcal{S}$ and let $X \in T_p\mathcal{S}$ denote the tangent vector $X = (1,3,0)$. Consider the function $f(x,y,z) = xz$ on $\mathcal{S}$. The directional derivative is

$$\mathrm{D}_X f = [\frac{\partial}{\partial x} + 3\frac{\partial}{\partial y}](xz)(p) = 1.$$

This follows from Lemma 3.5, since $F(x,y,z) = xz$, $(x,y,z) \in \mathbb{R}^3$, is a smooth extension of f.

3.6 Action on functions

We have seen in the preceding section that for a manifold in $\mathbb{R}^n$, we can bring its tangent vectors to act on smooth functions as a differentiation operator. The same can be done in the abstract case. Let M be an abstract manifold and let $p \in M$.

Lemma 3.6.1. *Let γ_1, γ_2 be parametrized curves on M with $p = \gamma_1(t_1) = \gamma_2(t_2)$. If $\gamma_1 \sim_p \gamma_2$ then $(f \circ \gamma_1)'(t_1) = (f \circ \gamma_2)'(t_2)$ for all $f \in C^\infty(M)$.*

Proof. Fix a chart $\sigma: U \to M$ around p, say with $p = \sigma(x_0)$. The composed map $f \circ \gamma_i$ can be written as follows in a neighborhood of t_i,

$$f \circ \gamma_i = (f \circ \sigma) \circ (\sigma^{-1} \circ \gamma_i).$$

Applying the chain rule we obtain

$$(f \circ \gamma_i)'(t_i) = D(f \circ \sigma)(x_0)(\sigma^{-1} \circ \gamma_i)'(t_i),$$

where $D(f \circ \sigma)(x_0)$ is the Jacobian $1 \times m$-matrix of $f \circ \sigma$ at x_0. The lemma follows immediately from this expression and (3.3). $\square$

It is a consequence of this lemma that the following definition is valid (it is independent of the choice of the representative γ).

Definition 3.6. Let $X \in T_pM$. The *directional derivative* in direction X of a function $f \in C^\infty(M)$ is

$$\mathrm{D}_X(f) = (f \circ \gamma)'(t_0),$$

where γ is any representative of X.

Notice that the action of D_X is *local*, in the sense that if two functions f_1, f_2 are equal in a neighborhood of p, then $\mathrm{D}_X(f_1) = \mathrm{D}_X(f_2)$. Notice also, that $\mathrm{D}_X f$ makes sense also for $f \in C^\infty(\Omega)$, when $p \in \Omega \subset M$ with Ω open (use the observation below Definition 3.3).

3.7 The differential of a smooth map

Let M, N be abstract manifolds of dimension m and n, respectively, and let $p \in M$. Let $f\colon M \to N$ be smooth. The *differential*, also called *tangent map*, of f at p will now be defined. We shall define it as a linear map from T_pM to $T_{f(p)}N$. It serves as a generalization of the Jacobian matrix $Df(p)$ of a map $f\colon \mathbb{R}^m \to \mathbb{R}^n$ (which, in turn, generalizes the ordinary derivative f' of $f\colon \mathbb{R} \to \mathbb{R}$).

If M and N are open subsets of $\mathbb{R}^m$ and $\mathbb{R}^n$, then by definition the differential df_p is the linear map $\mathbb{R}^m \to \mathbb{R}^n$ which has the Jacobian $Df(p)$ as its matrix with respect to the standard basis vectors. It satisfies

$$f(p+h) - f(p) = df_p(h) + o(h),$$

which means that

$$\frac{|f(p+h) - f(p) - df_p(h)|}{|h|} \to 0$$

as $h \to 0$. The interpretation is that the linear map df_p is an approximation to the increment of f at p.

Notice that if $m = 1$ and $M \subset \mathbb{R}$, then the linear map $df_p\colon \mathbb{R} \to \mathbb{R}^n$ is completely determined by its value $df_p(1) \in \mathbb{R}^n$. This value is exactly the derivative $f'(p) \in \mathbb{R}^n$, which constitutes the $1 \times n$-matrix $Df(p)$. This is for example the case when $f = \gamma$ is a curve in N.

The important *chain rule* is valid for the differential of a composed map $f \circ g$, where $g\colon L \to M$ and $f\colon M \to N$ are smooth maps between open sets $L \subset \mathbb{R}^l$, $M \subset \mathbb{R}^m$ and $N \subset \mathbb{R}^n$. It reads

$$d(f \circ g)_p = df_{g(p)} \circ dg_p,$$

and it is an immediate consequence of the corresponding chain rule for the Jacobians.

Assume that $\mathcal{S}$ and $\mathcal{T}$ are manifolds in $\mathbb{R}^k$ and $\mathbb{R}^l$, respectively, and let $f\colon \mathcal{S} \to \mathcal{T}$ be smooth. Recall (see Definition 2.6.1) that f has a local smooth extension $F\colon W \to \mathbb{R}^l$ in a neighborhood of each point $p \in \mathcal{S}$.

Definition 3.7.1. We define the *differential*

$$df_p\colon T_p\mathcal{S} \to T_{f(p)}\mathcal{T}$$

as the linear map, which is the restriction to the tangent space $T_p\mathcal{S} \subset \mathbb{R}^k$ of the differential $dF_p\colon \mathbb{R}^k \to \mathbb{R}^l$ of a local smooth extension F of f at p.

It follows from the chain rule, applied to $F \circ \gamma$, that

$$df_p(\gamma'(t_0)) = dF_p(d\gamma_{t_0}(1)) = d(F \circ \gamma)_{t_0}(1) = (f \circ \gamma)'(t_0) \tag{3.9}$$

for any smooth curve γ on $\mathcal{S}$ with $\gamma(t_0) = p$. The expression (3.9) shows that df_p is independent of the choice of the local smooth extension F, and it also implies that df_p maps into $T_{f(p)}N$.

Example 3.7.1 Let $f\colon \mathcal{S} \to \mathbb{R}$ be a smooth function on $\mathcal{S}$. The tangent space $T_y\mathbb{R}$ of $\mathbb{R}$ at any point $y \in \mathbb{R}$ is identified with $\mathbb{R}$ (see Example 3.2.1). The differential df_p is therefore a linear map $T_p\mathcal{S} \to \mathbb{R}$. It is determined from the expression (3.9).

Assume now that M and N are arbitrary abstract manifolds, and let again $f\colon M \to N$ be smooth. Recall that $[\gamma]_p$ denotes the $\sim_p$-class of γ. The following definition is inspired by (3.9).

Definition 3.7.2. The *differential* of f is the map $df_p\colon T_pM \to T_{f(p)}N$ defined by

$$df_p([\gamma]_p) = [f \circ \gamma]_{f(p)}$$

for all smooth curves γ on M through p.

Notice that $f \circ \gamma$ is a smooth curve on N, and $[f \circ \gamma]_{f(p)}$ is its equivalence class in $T_{f(p)}N$. In fact, in order for $df_p([\gamma]_p)$ to be well defined, we need to show that $[f \circ \gamma]_{f(p)}$ is independent of the choice of representative γ for the class $[\gamma]_p$. This is done in the lemma below. Notice also that if M and N are manifolds in $\mathbb{R}^k$ and $\mathbb{R}^l$, then it follows from (3.9) that the differentials defined in Definitions 3.7.1 and 3.7.2 agree with each other.

Lemma 3.7.1. *Let γ_1 and γ_2 be smooth curves on M with $\gamma_1(t_1) = \gamma_2(t_2) = p$, then*

$$\gamma_1 \sim_p \gamma_2 \quad \Rightarrow \quad f \circ \gamma_1 \sim_{f(p)} f \circ \gamma_2.$$

Proof. Let τ be a chart on N around $f(p)$, and let σ be a chart on M around p. Then for each curve on M through p,

$$\begin{aligned}(\tau^{-1} \circ f \circ \gamma)'(t_0) &= (\tau^{-1} \circ f \circ \sigma \circ \sigma^{-1} \circ \gamma)'(t_0) \\ &= D(\tau^{-1} \circ f \circ \sigma)(x_0)(\sigma^{-1} \circ \gamma)'(t_0)\end{aligned} \tag{3.10}$$

by the chain rule (where x_0 is determined by $\sigma(x_0) = p$). Applying this to γ_1 and γ_2, the implication of the lemma follows from the definition of the equivalence relations $\sim_p$ and $\sim_{f(p)}$. $\square$

The *chain rule* is an easy consequence of Definition 3.7.2. If $g: L \to M$ and $f: M \to N$ are smooth maps between abstract manifolds, then

$$d(f \circ g)_p = df_{g(p)} \circ dg_p. \tag{3.11}$$

Example 3.7.2 We have seen in Example 2.7.3 that a chart σ is a smooth map $U \to M$. We will determine its differential $d\sigma_{x_0}$, which is a smooth map from $\mathbb{R}^m$ (the tangent space of U at x_0) to T_pM. Let $v \in \mathbb{R}^m$. Then v is the tangent vector at 0 to the curve $x_0 + tv$ in U, and hence by definition $d\sigma_{x_0}(v)$ is the tangent vector at 0 to the curve $\sigma(x_0 + tv)$. This curve is exactly the curve γ_v, and hence $d\sigma_{x_0}(v) = [\gamma_v]_p$. We conclude that the differential $d\sigma_{x_0}$ is exactly the isomorphism $T^\sigma: v \mapsto \gamma_v'(0)$ of Theorem 3.4.

In particular, we conclude from the preceding example that the standard basis vectors (see Section 3.6) for T_pM with respect to σ are the vectors given by $d\sigma_{x_0}(e_i)$ for $i = 1, \dots, m$, where $p = \sigma(x_0)$.

Theorem 3.7. *The differential $df_p: T_pM \to T_{f(p)}N$ is a linear map.*

Proof. Let a chart σ on M with $p = \sigma(x_0)$ and a chart τ on N with $f(p) = \tau(y_0)$ be chosen. It follows from the chain rule (3.11) and Example 3.7.2 that the following diagram commutes

$$\begin{tikzcd} T_pM \arrow[r, "df_p"] & T_{f(p)}N \\ \mathbb{R}^m \arrow[u, "d\sigma_{x_0}"] \arrow[r, "d(\tau^{-1}\circ f\circ\sigma)_{x_0}"] & \mathbb{R}^n \arrow[u, "d\tau_{y_0}"'] \end{tikzcd}$$

where the vertical maps are the isomorphisms of Theorem 3.4 for M and N (see Example 3.7.2). Since the map in the bottom of this diagram is linear, we conclude that $df_p: T_pM \to T_{f(p)}N$ is a linear map. $\square$

The differential of the identity map $M \to M$ is the identity map $\mathrm{id}: T_pM \to T_pM$. Hence it follows from the chain rule that if $f: M \to N$ is a diffeomorphism, then

$$d(f^{-1})_{f(p)} \circ df_p = d(f^{-1} \circ f)_p = \mathrm{id}\,.$$

From this relation and the similar one with opposite order of f and f^{-1}, we see that when f is a diffeomorphism, then df_p is bijective with $(df_p)^{-1} = d(f^{-1})_{f(p)}$. We thus obtain that the differential of a diffeomorphism is a linear isomorphism between tangent spaces. In particular, two manifolds between which there exists a diffeomorphism, must have the same dimension.

Notice that the diagram above implies that the matrix for df_p, with respect to the standard bases for T_pM and $T_{f(p)}N$ for the given charts, is the Jacobian $D(\tau^{-1} \circ f \circ \sigma)(x_0)$.

Example 3.7.3 Let $f\colon V \to W$ be a linear map between finite dimensional vector spaces, as in Example 2.7.2. In Example 3.4.1 we have identified the tangent space T_pV of a vector space V with V itself. By applying this identification for both V and W, we shall see that *the differential at $p \in V$ of a linear map $f\colon V \to W$ is the map f itself.* An element $v \in V = T_pV$ is the tangent vector to the curve $t \mapsto p + tv$ at $t = 0$. This curve is mapped to $t \mapsto f(p+tv) = f(p) + tf(v)$ by f, and the tangent vector at $t = 0$ of the latter curve on W is exactly $f(v)$. Hence $df_p(v) = f(v)$ as claimed.

In Section 3.6 an action of tangent vectors by directional derivations was defined. We will now determine the differential in this picture.

Lemma 3.7.2. *Let $X \in T_pM$ and let $f\colon M \to N$ be smooth. Then*

$$\mathrm{D}_{df_p(X)}(\varphi) = \mathrm{D}_X(\varphi \circ f)$$

for all $\varphi \in C^\infty(N)$.

Proof. Let γ be a representative of X. Then $f \circ \gamma$ is a representative of $df_p(X)$ and hence by Definition 3.6

$$\mathrm{D}_{df_p(X)}(\varphi) = (\varphi \circ f \circ \gamma)'(t_0) = \mathrm{D}_X(\varphi \circ f). \quad \square$$

3.8 The standard basis

Let M be an abstract manifold. Given a chart σ with $p = \sigma(x_0)$, we obtain a basis for T_pM from Theorem 3.4, by taking the isomorphic image $T^\sigma(e_i)$ of each of the standard basis vectors $e_1, \dots, e_m$ for $\mathbb{R}^m$ (in that order), that is, the basis vectors will be the equivalence classes of the curves $t \mapsto \sigma(x_0 + te_i)$. This basis for T_pM is called the *standard basis* with respect to σ. With the notation from the preceding section, see Example 3.7.2, the standard basis vectors are

$$d\sigma_{x_0}(e_1), \dots, d\sigma_{x_0}(e_m).$$

These basis vectors, which depend on the chosen chart σ, can be seen as abstract analogs of the basis vectors

$$\sigma'_{x_1}, \sigma'_{x_2}$$

for the tangent space of a parametrized surface in $\mathbb{R}^3$ (the columns of the Jacobian $D\sigma(x_0)$).

The standard basis vectors determine the following operators $\mathrm{D}_1, \dots, \mathrm{D}_m$ of directional differentiation. Let $\mathrm{D}_i = \mathrm{D}_{X_i}$ where $X_i = d\sigma_{x_0}(e_i)$. Then by Definition 3.6

$$\mathrm{D}_i f = \frac{d}{dt}|_{t=0}\, f(\sigma(x_0 + te_i)) = \frac{\partial}{\partial x_i}(f \circ \sigma)(x_0). \tag{3.12}$$

Because of (3.12) it is customary to denote D_i by $\frac{\partial}{\partial x_i}$.

3.9 Orientation

Let V be a finite dimensional vector space. Two ordered bases $(v_1, \dots, v_n)$ and $(\tilde{v}_1, \dots, \tilde{v}_n)$ are said to be *equally oriented* if the transition matrix S, whose columns are the coordinates of the vectors $v_1, \dots, v_n$ with respect to the basis $(\tilde{v}_1, \dots, \tilde{v}_n)$, has positive determinant. Being equally oriented is an equivalence relation among bases, for which there are precisely two equivalence classes. The space V is said to be *oriented* if a specific class has been chosen, this class is then called the *orientation* of V, and its member bases are called *positive.* The Euclidean spaces $\mathbb{R}^n$ are usually oriented by the class containing the standard basis $(e_1, \dots, e_n)$. For the null space $V = \{0\}$ we introduce the convention that an orientation is a choice between the signs $+$ and $-$.

Example 3.9.1 For a two-dimensional subspace V of $\mathbb{R}^3$ it is customary to assign an orientation by choosing a normal vector N. The positive bases (v_1, v_2) for V are those for which (v_1, v_2, N) is a positive basis for $\mathbb{R}^3$ (in other words, it is a right-handed triple).

Let σ be a chart on an abstract manifold M, then the tangent space is equipped with the standard basis (see Section 3.6) with respect to σ. For each $p \in \sigma(U)$ we say that the orientation of T_pM, for which the standard basis is positive, is the orientation *induced* by σ.

Definition 3.9. An *orientation* of an abstract manifold M is an orientation of each tangent space T_pM, $p \in M$, such that there exists an atlas of M in which all charts induce the given orientation on each tangent space.

The manifold is called *orientable* if there exists an orientation. If an orientation has been chosen we say that M is an *oriented manifold* and we call a chart *positive* if it induces the proper orientation on each tangent space.

A diffeomorphism map $f\colon M \to N$ between oriented manifolds of equal dimension is said to be *orientation preserving* if for each $p \in M$, the differential df_p maps positive bases for T_pM to positive bases for $T_{f(p)}N$.

Example 3.9.2 Every manifold M, of which there exists an atlas with only one chart, is orientable. The orientation induced by that chart is of course an orientation of M.

Example 3.9.3 Suppose $\mathcal{S} \subset \mathbb{R}^n$ is the $n-k$-dimensional manifold given by an equation $f(p) = c$ as in Theorem 1.6. As seen in Example 3.2.2 the tangent space at $p \in \mathcal{S}$ is the null space

$$T_p\mathcal{S} = \{X \in \mathbb{R}^n \mid Df(p)X = 0\}.$$

for the $k \times n$ matrix $Df(p)$. Let $w_1, \dots, w_k \in \mathbb{R}^n$ denote the rows of $Df(p)$, in the same order as they appear in the matrix. We can then define an orientation of $T_p\mathcal{S}$ by declaring an ordered basis $(v_1, \dots, v_{n-k})$ positive if the combined basis $(v_1, \dots, v_{n-k}, w_1, \dots, w_k)$ for $\mathbb{R}^n$ is positive with respect to the standard order. It can be shown that this is an orientation of $\mathcal{S}$, which is thus orientable.

For a surface in $\mathbb{R}^3$ given by $f(x) = c$, where f is a scalar valued function, this means that we define the orientation by means of the normal given by the gradient vector $w = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z})$ of f.

Not all manifolds are orientable however, the most famous example being the two-dimensional *Möbius band*. A model for it in $\mathbb{R}^3$ can be made by gluing together the ends of a strip which has been given a half twist. It is impossible to make a consistent choice of orientations, because the band is ‘one-sided. Choosing an orientation in one point forces it by continuity to be given in neighboring points, etcetera, and eventually we are forced to the opposite choice in the initial point.

Chapter 4

Submanifolds

The notion of a submanifold of an abstract smooth manifold will now be defined. In fact, there exist two different notions of submanifolds, 'embedded submanifolds' and 'immersed submanifolds'. Here we shall focus on the stronger notion of embedded submanifolds, and for simplicity we will just use the term 'submanifold'.

4.1 Submanifolds in $\mathbb{R}^k$

Definition 4.1. Let $\mathcal{S}$ be a manifold in $\mathbb{R}^k$. A *submanifold* of $\mathcal{S}$ is a subset $\mathcal{T} \subset \mathcal{S}$, which is also a manifold in $\mathbb{R}^k$.

Lemma 4.1. *Let $\mathcal{T}$ be a submanifold of $\mathcal{S} \subset \mathbb{R}^k$. Then $\dim \mathcal{T} \leq \dim \mathcal{S}$ and $T_p\mathcal{T} \subset T_p\mathcal{S}$ for all $p \in \mathcal{T}$. Moreover, the inclusion map $i\colon \mathcal{T} \to \mathcal{S}$ is smooth, and its differential di_p at $p \in \mathcal{T}$ is the inclusion map $T_p\mathcal{T} \to T_p\mathcal{S}$.*

Proof. By Theorem 3.2, the tangent space $T_p\mathcal{T}$ is the space of all tangent vectors $\gamma'(t_0)$ to curves γ on $\mathcal{T}$ with $\gamma(t_0) = p$. Since $\mathcal{T} \subset \mathcal{S}$, a curve on $\mathcal{T}$ is also a curve on $\mathcal{S}$, and the inclusion $T_p\mathcal{T} \subset T_p\mathcal{S}$ follows. The inequality of dimensions is an immediate consequence.

The inclusion map $i\colon \mathcal{T} \to \mathcal{S}$ is the restriction of the identity map $\mathrm{id}\colon \mathbb{R}^k \to \mathbb{R}^k$, hence it is smooth, and by Definition 3.7.1 its differential at p is the restriction of the differential of id. The differential of the identity map is the identity map, so we conclude that di_p is the inclusion map $T_p\mathcal{T} \to T_p\mathcal{S}$. $\square$

Example 4.1 The circle $\{(x, y, 0) \mid x^2 + y^2 = 1\}$ is a one-dimensional manifold in $\mathbb{R}^3$ (for example by Theorem 1.6 with $f(x, y, z) = (x^2 + y^2, z)$ and $c = (1, 0)$). It is a submanifold (the equator) of the 2-sphere $\{(x, y, z) \mid x^2 + y^2 + z^2 = 1\}$.

4.2 Abstract submanifolds

Let M be an abstract manifold. A naive generalization of the definition above would be that a submanifold of M is subset of M, which is an abstract manifold of its own. This however, would not be a feasible definition, because it does not ensure any compatibility between the differential structures of M and the subset. In Definition 4.1 the compatibility, which is reflected in Lemma 4.1, results from the position of both manifolds inside an ambient space $\mathbb{R}^k$. For the general definition we use that lemma as inspiration.

Recall from Lemma 2.1.2 that every subset of a topological space is equipped as a topological space with the induced topology.

Definition 4.2. Let M be an abstract manifold. An *abstract submanifold* of M is a subset $N \subset M$ which is an abstract manifold on its own such that:
(i) the topology of N is induced from M,
(ii) the inclusion map $i\colon N \to M$ is smooth, and
(iii) the differential $di_p\colon T_pN \to T_pM$ is injective for each $p \in N$.

In this case, the manifold M is said to be *ambient* to N. In particular, since di_p is injective, the dimension of N must be smaller than or equal to that of M. By Definition 3.7.2, the differential di_p maps the equivalence class of a curve γ through p on N to the equivalence class of the same curve, now regarded as a curve on the ambient manifold (formally the new curve is $i \circ \gamma$). Based on the assumption in (iii) that this map is injective, we shall often regard the tangent space T_pN as a subspace of T_pM.

Example 4.2.1 Let M be an m-dimensional real vector space, regarded as an abstract manifold (see Example 2.3.1). Every linear subspace $N \subset M$ is then an abstract submanifold. The inclusion map $N \to M$ is linear, hence it is smooth and has an injective differential in each point $p \in N$ (see Examples 2.7.2 and 3.7.3).

Example 4.2.2 It follows directly from Lemma 4.1 that a submanifold $\mathcal{T}$ of a manifold $\mathcal{S}$ in $\mathbb{R}^k$ is also an abstract submanifold of $\mathcal{S}$, when $\mathcal{S}$ and $\mathcal{T}$ are regarded as abstract manifolds.

Example 4.2.3 A non-empty open subset of an m-dimensional abstract manifold M is an abstract submanifold. Indeed, as mentioned in Example 2.3.2 the subset is an abstract manifold of its own, also of dimension m. The conditions (i)-(iii) are easily verified in this case. Conversely, it follows by application of the inverse function theorem to the inclusion map i, that every m-dimensional abstract submanifold of M is an open subset of M.

Example 4.2.4 Let $M = \mathbb{R}^2$ with the standard differential structure, and let $N \subset M$ be the x-axis, equipped with standard topology together with the non-standard differential structure given by the chart $\tau(s) = (s^3, 0)$ (see Example 2.3.3). The inclusion map i is smooth, since $i \circ \tau : s \mapsto (s^3, 0)$ is smooth into $\mathbb{R}^2$. Its differential at $s = 0$ is 0, hence (iii) fails, and N (equipped with τ) is *not* a submanifold of M.

Notice that the property of being an abstract submanifold is transitive, that is, if N is a submanifold of M, and M is a submanifold of L, then N is a submanifold of L. This follows from application of the chain rule to the composed inclusion map.

The following lemma deals with the special case where the ambient manifold is $M = \mathbb{R}^k$.

Lemma 4.2. *Let $N \subset \mathbb{R}^k$ be a subset, equipped with the induced topology. Then N is an abstract submanifold of $\mathbb{R}^k$ if and only if it is a manifold in $\mathbb{R}^k$ (with the same smooth structure).*

Proof. The implication 'if' follows from Example 4.2.2 with $\mathcal{S} = \mathbb{R}^k$. For the converse we assume that N carries the structure of an abstract submanifold of $\mathbb{R}^k$, and we have to verify the properties in Definition 1.6.1, and that the smooth structure is the same.

Let $\sigma: U \to N$ be a chart (in the sense of an abstract manifold). We claim that σ is an embedded parametrized manifold with an image which is open in N. Since N is covered by such charts, all the desired properties follow from this claim.

It is part of the axioms for abstract manifolds that σ is a homeomorphism onto an open subset of N. Since it is required in (i) that N carries the induced topology, this means that $\sigma(U) = N \cap W$ as required in Definition 1.6.1.

We can regard σ as a map into $\mathbb{R}^k$, by composing it with the inclusion map $N \to \mathbb{R}^k$. The latter map is smooth by assumption (ii), hence $\sigma: U \to \mathbb{R}^k$ is smooth.

It remains to be seen that σ is regular, which means that the columns of the Jacobian matrix $D\sigma(p)$ are linearly independent. Since the differential $d\sigma_p: \mathbb{R}^n \to \mathbb{R}^k$ is the linear map represented by $D\sigma(p)$, an equivalent requirement is that this map is injective.

By the chain rule the differential of $\sigma: U \to \mathbb{R}^k$ is composed by the differential of $\sigma: U \to N$ and the differential of the inclusion map. The differential of $\sigma: U \to N$ is an isomorphism according to Example 3.7.2, and the differential di_p of $N \to \mathbb{R}^k$ is injective by assumption (iii). Hence the differential of $\sigma: U \to \mathbb{R}^k$ is injective. $\square$

The following corollary ensures that we do not have a conflict of definitions. Because of this we will often drop the word 'abstract' and just speak of submanifolds.

Corollary 4.2. *Let $\mathcal{S}$ be a smooth manifold in $\mathbb{R}^k$, and let $N \subset \mathcal{S}$ be a subset. Then N is an abstract submanifold according to Definition 4.2 if and only if it is a submanifold according to Definition 4.1.*

Proof. The implication 'if' was remarked in Example 4.2.2. To see the converse, assume that N is an abstract submanifold of $\mathcal{S}$, according to Definition 4.2. By the remark about transitivity before Lemma 4.2, N is an abstract submanifold of $\mathbb{R}^k$, so by Lemma 4.2 it is a manifold in $\mathbb{R}^k$. Since it is also contained in $\mathcal{S}$, it is a submanifold according to Definition 4.1. $\square$

4.3 The local structure of submanifolds

The following theorem, which shows that submanifolds are very well behaved, explains why they are said to be 'embedded'. It shows that locally all n-dimensional submanifolds of an m-dimensional smooth manifold look similar to the simplest example of an n-dimensional manifold in $\mathbb{R}^m$, the 'copy'

$$\{(x_1, \dots, x_n, 0, \dots, 0)\}$$

of $\mathbb{R}^n$ inside $\mathbb{R}^m$ (Example 1.6.2).

In the statement of the following theorem, we shall regard $\mathbb{R}^n$ as a subset of $\mathbb{R}^m$, given by the canonical inclusion

$$j(x_1, \dots, x_n) = (x_1, \dots, x_n, 0, \dots, 0) \in \mathbb{R}^m.$$

Theorem 4.3. *Let M be an m-dimensional abstract manifold, and $N \subset M$ a non-empty subset.*

If N is an n-dimensional submanifold, then for each $p \in N$ there exist a number $\epsilon > 0$ and a chart $\sigma\colon U =]-\epsilon, \epsilon[^m \to M$ with $\sigma(0) = p$, such that the following holds.

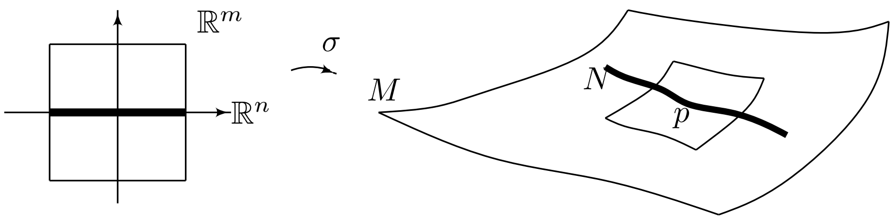
(1) $\sigma(U) \cap N = \sigma(U \cap \mathbb{R}^n)$

(2) *The restriction of σ to $U \cap \mathbb{R}^n$ is a chart on N.*

Conversely, if for each $p \in N$ there exist a chart $\sigma\colon U \to M$ with $\sigma(0) = p$ and

$$\sigma(U) \cap N = \sigma(U \cap \mathbb{R}^n),$$

then N is a submanifold of M, and the restricted maps $\sigma|_{U \cap \mathbb{R}^n}$ constitute the charts of an atlas for it.



Example 4.3 Let $M = \{(x, y, z) \mid x^2 + y^2 + z^2 = 1\}$ and let $N \subset M$ be the equator. We have seen in Example 4.1 that it is a submanifold. The sperical coordinates $\sigma(u, v) = (\cos u \cos v, \sin u \cos v, \sin v)$ have the properties required in the theorem for $p = (1, 0, 0)$ if we choose for example $\epsilon = \frac{\pi}{2}$.

Proof. The proof is based on a clever application of the inverse function theorem. Assume N is a submanifold, let $p \in N$, and choose arbitrary charts $\tilde{\sigma}\colon U \to M$ and $\tau\colon V \to N$ about p. We may assume that $0 \in U$ with $\tilde{\sigma}(0) = p$ and $0 \in V$ with $\tau(0) = p$.

In order to attain the properties (1) and (2), we would like to arrange that $V = U \cap \mathbb{R}^n$ and $\tau = \tilde{\sigma}|_V$. We will accomplish this by means of a modification of $\tilde{\sigma}$, and by shrinking the sets U and V. The modification will be of the form $\sigma = \tilde{\sigma} \circ \Phi$, with Φ a suitably defined diffeomorphism of a neighborhood of 0 in $\mathbb{R}^m$.

Let

$$\Psi = \tilde{\sigma}^{-1} \circ \tau \colon \mathbb{R}^n \to \mathbb{R}^m,$$

then Ψ is defined in a neighborhood of 0, and we have $\Psi(0) = 0$. Notice that we can view Ψ as the coordinate expression $\tilde{\sigma}^{-1} \circ i \circ \tau$ for the inclusion map in these coordinates around p.

The Jacobian matrix $D\Psi(0)$ of Ψ at 0 is an $m \times n$ matrix, and since di_p is injective this matrix has rank n. By a reordering of the coordinates in $\mathbb{R}^m$ we may assume that

$$D\Psi(0) = \begin{pmatrix} A \\ B \end{pmatrix}$$

where A is an $n \times n$ matrix with non-zero determinant, and B is an arbitrary $(m-n) \times n$ matrix.

Define $\Phi \colon \mathbb{R}^m \to \mathbb{R}^m$ on a neighborhood of 0 by $\Phi(x, y) = \Psi(x) + (0, y)$ for $x \in \mathbb{R}^n$, $y \in \mathbb{R}^{m-n}$. The Jacobian matrix of Φ at 0 has the form

$$D\Phi(0) = \begin{pmatrix} A & 0 \\ B & I \end{pmatrix}$$

where I denotes the $(m-n) \times (m-n)$ identity matrix. This Jacobian matrix has non-zero determinant and hence the inverse function theorem implies that Φ is a diffeomorphism in a neighborhood of 0. Notice that we have defined Φ so that $\Psi = \Phi \circ j$. Hence with coordinates $\sigma = \tilde{\sigma} \circ \Phi$ about p we obtain

$$\sigma \circ j = \tilde{\sigma} \circ \Phi \circ j = \tilde{\sigma} \circ \Psi = \tau,$$

in a neighborhood of 0. With our modified chart σ we have thus arranged that $\sigma \circ j = \tau$.

Let $\epsilon > 0$ be sufficiently small so that $]-\epsilon, \epsilon[^m$ and $]-\epsilon, \epsilon[^n$ are contained in all the mentioned neighborhoods of 0 in $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively. Let $U =]-\epsilon, \epsilon[^m$ and $V =]-\epsilon, \epsilon[^n$, and take the restriction of σ and τ to these sets. Then $j(V) = U \cap \mathbb{R}^n$, hence $\tau(V) = \sigma(j(V)) \subset \sigma(U) \cap N$, and (2) holds.

In order to arrange equality in (1) we shall replace ϵ by a smaller value. It is easily seen that such a change will not destroy the validity of the already established inclusion $\tau(V) \subset \sigma(U) \cap N$, nor does it destroy property (2).

Since $\tau(V)$ is open in N, which carries the induced topology, there exists an open set $W \subset M$ such that $\tau(V) = W \cap N$. Choose $\epsilon' > 0$ such that $\epsilon > \epsilon'$ and $\sigma(U') \subset W$, where $U' =]-\epsilon', \epsilon'[^m$. Let $V' =]-\epsilon, \epsilon[$. We claim that then $\tau(V') = \sigma(U') \cap N$. The inclusion $\subset$ has already been remarked, so let $u \in U'$ be an element with $\sigma(u) \in N$. Since $\tau(V) = W \cap N$ there exists $v \in V$ such that $\sigma(u) = \tau(v)$, and hence $\sigma(u) = \sigma(j(v))$. Since σ is injective it then follows that $u = j(v)$, from which we conclude that $v \in V'$. We have thus shown that $\sigma(U') \cap N \subset \tau(V')$. The equality in (1) has been established.

Conversely, assume there exists a chart σ of the mentioned type around each point $p \in N$. Then the maps $\tau = \sigma \circ j$ constitute an atlas on N. The condition of smooth transition maps follows from the same condition for M. Hence N can be given a smooth structure. The inclusion map, expressed in the local coordinates, is $\sigma^{-1} \circ i \circ \tau = j$. It follows that it is smooth and has an injective differential, so that N is a submanifold. $\square$

Corollary 4.3.1. *If $N \subset M$ is a submanifold of an abstract manifold, then there exists for each $p \in N$ an open neighborhood W in M and a smooth map $\pi\colon W \to N$ such that $\pi(x) = x$ for all $x \in W \cap N$.*

Proof. Choose a chart σ around p as in the theorem, and let $W = \sigma(U)$ and $\pi(\sigma(x_1, \dots, x_m)) = \tau(x_1, \dots, x_n)$ for $x \in U$. $\square$

Corollary 4.3.2. *Let L, M be abstract manifolds, let $N \subset M$ be a submanifold, and let $f\colon L \to N$ be a map. Then f is smooth $L \to M$ if and only if it is smooth $L \to N$.*

Proof. The statement 'if' follows from the fact that $f = i \circ f$, where $i\colon N \to M$ is the inclusion map, which is smooth. The statement 'only if' follows from the fact that, in the notation of the preceding corollary, $f = \pi \circ f$, where $\pi\colon M \to N$ is defined and smooth in a neighborhood of p. $\square$

A smooth structure on a subset $N \subset M$, which satisfies the properties (i)-(iii) in Definition 4.2, will be called a *submanifold structure*. It follows from Corollary 4.3.2 that the submanifold structure of a subset is unique, if it exists. For if there were two such structures, the identity map of N would be a diffeomorphism between them. Being a submanifold is thus really just a property of the subset $N \subset M$. The following corollary shows that this property is 'local'.

Corollary 4.3.3. *Let M be an abstract manifold and $N \subset M$ a subset. Then N is a submanifold if and only if for each $p \in N$ there exists an open neighborhood W of p in M such that $W \cap N$ is a submanifold of W.*

Proof. It follows directly from Definition 4.2 that if N is a submanifold of M, then $N \cap W$ is a submanifold of W for all open sets $W \subset M$. Conversely, assume that N satisfies the local condition. By applying the first part of Theorem 4.3 to the submanifold $N \cap W$ of W we obtain charts of the specified type within W, for each of the sets W. Since W is open in M, a chart on W is also a chart on M, and since all the sets W cover N, the second part of Theorem 4.3 implies that N is a submanifold. $\square$

Notice that it follows from Lemma 2.9 that if M has a countable atlas, then so has every submanifold N, since it inherits from M the property of having a countable base for the topology.

4.4 Level sets

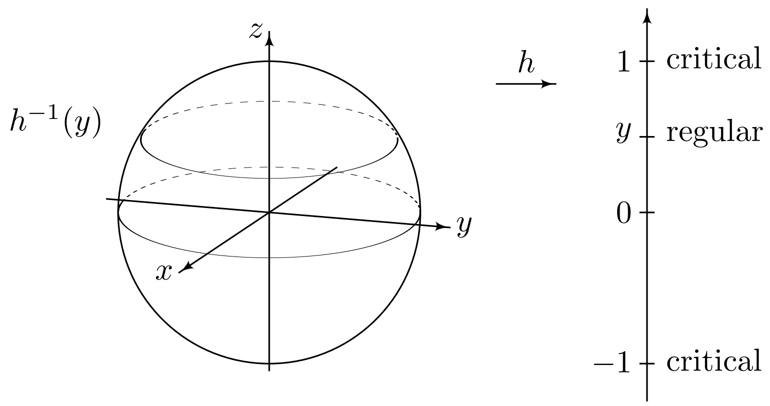
The following theorem is an efficient tool for the construction of manifolds. It represents the analog for abstract manifolds of Theorem 1.6. Let M, N be abstract smooth manifolds of dimension m and n, respectively, and let $f\colon M \to N$ be a smooth map.

Definition 4.4. An element $y \in N$ is called a *regular value* if for each $x \in f^{-1}(y) = \{x \in M \mid f(x) = y\}$, the differential $df_x\colon T_xM \to T_yN$ is surjective. A *critical value* is an element $y \in N$ which is not regular.

Notice that by this definition all elements $y \notin f(M)$ are regular values. Notice also that if there exists a regular value $y \in f(M)$, then necessarily $n \leq m$, since a linear map into a higher dimensional space cannot be surjective.

Example 4.4 Let $M \subset \mathbb{R}^3$ be the unit sphere, and let $h\colon M \to \mathbb{R}$ denote the 'height' function $h(x, y, z) = z$. The map $h\colon M \to \mathbb{R}$ is smooth, since it is the restriction of the map $H\colon \mathbb{R}^3 \to \mathbb{R}$, also given by $H(x, y, z) = z$. The differential of h at $p \in M$ is the restriction to T_pM of $dH_p\colon \mathbb{R}^3 \to \mathbb{R}$, hence $dh_p(v) = v \cdot (0, 0, 1)$. It follows that dh_p is surjective $T_pM \to \mathbb{R}$ if and only T_p is not perpendicular to the z-axis. This happens exactly when p is not one of the poles $\pm(0, 0, 1)$. We conclude that the critical values $y \in \mathbb{R}$ are $y = \pm 1$.

Theorem 4.4. *Let $y \in f(M)$ be a regular value for $f\colon M \to N$. Then $f^{-1}(y)$ is a submanifold of M of dimension $m-n$. Furthermore, the tangent space $T_p(f^{-1}(y))$ is the null space for the linear map $df_p\colon T_pM \to T_yN$ for each $p \in f^{-1}(y)$.*



The height function on the sphere, see Example 4.4

Proof. The proof, which resembles that of Theorem 4.3, is again based on a clever application of the inverse function theorem. Let $P\colon \mathbb{R}^m \to \mathbb{R}^n$ denote the projection

$$P\colon (x_1, \dots, x_m) \mapsto (x_1, \dots, x_n)$$

on the first n coordinates.

We first choose arbitrary coordinate systems $\sigma\colon U \to M$ and $\tau\colon V \to N$ about p and y such that $\sigma(0) = p$ and $\tau(0) = y$. Let

$$\Psi = \tau^{-1} \circ f \circ \sigma\colon \mathbb{R}^m \to \mathbb{R}^n,$$

then Ψ is defined in a neighborhood of 0, and we have $\Psi(0) = 0$. If we had $\Psi = P$, then we would have

$$f^{-1}(y) = \sigma(\Psi^{-1}(0)) = \sigma(\{x \in \mathbb{R}^m \mid x_1 = \dots = x_n = 0\})$$

in a neighborhood of p. This would imply that $f^{-1}(y)$ is an $m-n$-dimensional submanifold, according to the last statement in Theorem 4.3 (with n replaced by $m-n$). We will modify σ in order to obtain $\Psi = P$.

The modification will be of the form $\tilde{\sigma} = \sigma \circ \Phi^{-1}$, with Φ a suitably defined diffeomorphism of a neighborhood of 0 in $\mathbb{R}^m$. The Jacobian matrix $D\Psi(0)$ of Ψ at 0 is an $n \times m$ matrix, and since df_p is surjective this matrix has rank n. By a reordering of the coordinates in $\mathbb{R}^m$ we may assume

$$D\Psi(0) = \begin{pmatrix} A & B \end{pmatrix}$$

where A is an $n \times n$ matrix with non-zero determinant, and B is an arbitrary $n \times (m-n)$ matrix.

Define $\Phi\colon \mathbb{R}^m \to \mathbb{R}^m$ on a neighborhood of 0 by $\Phi(x, y) = (\Psi(x, y), y)$ for $x \in \mathbb{R}^n$, $y \in \mathbb{R}^{m-n}$. The Jacobian matrix of Φ at 0 has the form

$$\begin{pmatrix} A & B \\ 0 & I \end{pmatrix}$$

where I denotes the $(m-n) \times (m-n)$ identity matrix. This Jacobian matrix has non-zero determinant and hence the inverse function theorem implies that Φ is a diffeomorphism in a neighborhood of 0. Notice that we have defined Φ so that $\Psi = P \circ \Phi$. Hence with coordinates $\tilde{\sigma} = \sigma \circ \Phi^{-1}$ about p we obtain

$$\tau^{-1} \circ f \circ \tilde{\sigma} = \Psi \circ \Phi^{-1} = P,$$

in a neighborhood of 0, as desired.

Only the statement about the tangent space remains to be proved. Since f is a constant function on $f^{-1}(y)$, its differential is zero on the tangent space of $f^{-1}(y)$. It follows that $T_x(f^{-1}(y))$ is contained in the null space of df_x. That it is equal then follows by a comparison of dimensions. $\square$

4.5 The orthogonal group

As an example of the application of Theorem 4.4 we will show that the *orthogonal group* $\mathrm{O}(n)$ is a Lie group, and we will determine its dimension. Recall that by definition $\mathrm{O}(n)$ is the group of all $n \times n$ real matrices A, which are orthogonal, that is, which satisfy

$$AA^t = I.$$

Theorem 4.5. *$\mathrm{O}(n)$ is a Lie group of dimension $n(n-1)/2$.*

Proof. It follows from Theorem 2.8 that we only have to establish that $\mathrm{O}(n)$ is a submanifold of $\mathrm{GL}(n, \mathbb{R})$, of the mentioned dimension.

Let $\mathrm{Sym}(n)$ denote the set of all symmetric $n \times n$ real matrices, this is a vector space of dimension $n(n+1)/2$. Hence $\mathrm{Sym}(n)$ is a manifold of dimension $n(n+1)/2$.

Furthermore, let $f \colon \mathrm{GL}(n, \mathbb{R}) \to \mathrm{Sym}(n)$ be the map given by $f(A) = AA^t$, clearly this is a smooth map. Then $\mathrm{O}(n) = f^{-1}(I)$, where $I \in \mathrm{Sym}(n)$ is the identity matrix. It will be shown below that I is a regular value for f, hence it follows from Theorem 4.4 that $\mathrm{O}(n)$ is a submanifold of $\mathrm{GL}(n, \mathbb{R})$ of dimension $n^2 - n(n+1)/2 = n(n-1)/2$.

In order to show that I is a regular value, we must determine the differential df_A of f at each $A \in \mathrm{O}(n)$, and show that it is surjective from $T_A \,\mathrm{GL}(n, \mathbb{R})$ to $T_{f(A)} \,\mathrm{Sym}(n)$.

Recall that $\mathrm{GL}(n, \mathbb{R})$ is an open subset in the n^2-dimensional vector space $\mathrm{M}(n, \mathbb{R})$ of all real $n \times n$ matrices, hence its tangent space at A is identified with $\mathrm{M}(n, \mathbb{R})$. The tangent vector corresponding to a matrix $X \in \mathrm{M}(n, \mathbb{R})$ is the the derivative at $t = 0$ of the curve $\gamma(t) = A + tX$. By definition, the differential maps this tangent vector to the derivative at $t = 0$ of the curve $f \circ \gamma$. Since

$$f(A + tX) = (A + tX)(A + tX)^t = AA^t + t(AX^t + XA^t) + t^2 XX^t$$

we conclude that $df_A(X) = AX^t + XA^t \in \mathrm{Sym}(n)$.

The claim has thus been reduced to the claim that if A is orthogonal, then the linear map $X \mapsto AX^t + XA^t$ is surjective of $\mathrm{M}(n, \mathbb{R})$ onto $\mathrm{Sym}(n)$. This is easily seen, for if $B \in \mathrm{Sym}(n)$ is given, then $B = AX^t + XA^t$ where $X = \frac{1}{2}BA$. $\square$

Notice that it follows from Theorem 4.4 and the proof above that the tangent space of $\mathrm{O}(n)$ at A is

$$T_A \,\mathrm{O}(n) = \{X \in \mathrm{M}(n, \mathbb{R}) \mid AX^t + XA^t = 0\}.$$

In particular, the tangent space at the identity matrix I is the space

$$T_I \,\mathrm{O}(n) = \{X \in \mathrm{M}(n, \mathbb{R}) \mid X^t = -X\}$$

of antisymmetric matrices.

4.6 Domains with smooth boundary

As mentioned in Example 4.2.3, an open subset of an m-dimensional abstract manifold is an m-dimensional submanifold. In this section we will discuss some particularly well behaved open subsets.

Definition 4.6. Let M be an m-dimensional abstract manifold. A non-empty open subset $D \subset M$ is said to be a *domain with smooth boundary* if for each $p \in \partial D$ there exists an open neighborhood W of p in M, and a smooth function $f\colon W \to \mathbb{R}$ such that 0 is a regular value, $f(p) = 0$ and

$$W \cap D = \{x \in W \mid f(x) > 0\}.$$

In most examples D is given by a single function f as above, with $W = M$.

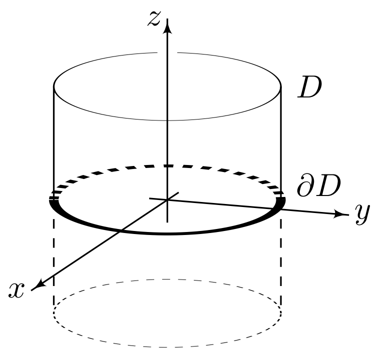
Example 4.6.1 Let M denote the cylinder

$$\{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 = 1\}$$

in $\mathbb{R}^3$. The half-cylinder

$$D = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 = 1,\ z > 0\}$$

is an open submanifold of M. The function $f\colon M \to \mathbb{R}$ defined by $f(x, y, z) = z$ has the required property, hence D is a domain with smooth boundary in M.



Example 4.6.2 A particularly simple case is that where $M = \mathbb{R}^m$. For example, the *half-space*

$$D = \mathbb{H}^m = \{x \in \mathbb{R}^m \mid x_m > 0\}$$

is a domain with smooth boundary. The boundary is $\partial D = \mathbb{R}^{m-1}$, standardly embedded as $\{x \mid x_m = 0\}$. The unit ball

$$D = \{x \in \mathbb{R}^m \mid \|x\|^2 < 1\}$$

is another example. The boundary is the unit sphere $\partial D = S^{m-1}$.

Lemma 4.6. *Let $D \subset M$ be a domain with smooth boundary, and let p, W and f be as in Definition 4.6. Then*

$$W \cap \partial D = \{x \in W \mid f(x) = 0\}.$$

Proof. Let $x \in W \cap \partial D$ be given. If $f(x) > 0$ then x belongs to D, which is open, and hence $x \notin \partial D$. If $f(x) < 0$ then it follows from the continuity of f that an open neighborhood of x is disjoint from D, and again we conclude that $x \notin \partial D$. Thus $f(x) = 0$.

For the opposite inclusion we use the condition of regularity in Definition 4.6. This condition implies that if $x \in W$ is a point with $f(x) = 0$ then $df_x(v) \neq 0$ for some tangent vector $v \in T_xW = T_xM$. We may assume $df_x(v) > 0$. If γ is a curve on M with $\gamma(0) = x$, representing v, then it follows that $(f \circ \gamma)'(0) > 0$. Hence $f(\gamma(t)) > 0$ for t positive and sufficiently close to 0, which shows that x is a limit of points in $W \cap D$. Hence $x \in \partial D$. ☐

Corollary 4.6. *Let $D \subset M$ be a domain with smooth boundary. If the boundary ∂D is not empty, it is an $m-1$-dimensional submanifold of M.*

Proof. We apply Corollary 4.3.3. Let $p \in \partial D$, and let W and f be as above. It follows from Lemma 4.6 and Theorem 4.4 that $\partial D \cap W$ is an $m-1$-dimensional submanifold of W. ☐

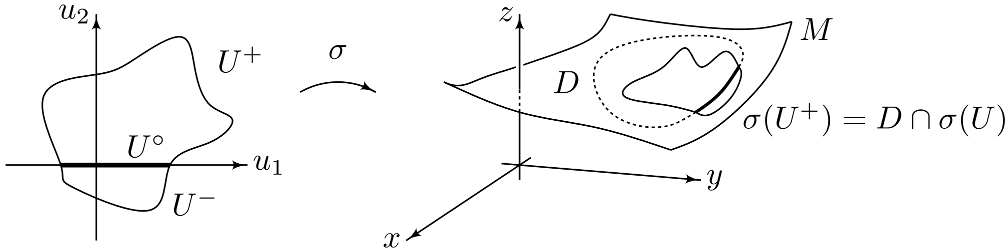
Theorem 4.6. *Let $D \subset M$ be open and non-empty. Then D is a domain with smooth boundary if and only if for each $p \in \partial D$ there exists a chart $\sigma: U \to M$ on M around $p = \sigma(0)$ such that*

$$D \cap \sigma(U) = \sigma(U^+) \quad \textit{where} \quad U^+ = \{x \in U \mid x_m > 0\}.$$

If this is the case and σ is such a chart, then in addition

$$\partial D \cap \sigma(U) = \sigma(U^\circ) \quad \textit{where} \quad U^\circ = \{x \in U \mid x_m = 0\},$$

and the map $(x_1, \dots, x_{m-1}) \mapsto \sigma(x_1, \dots, x_{m-1}, 0)$ is a chart on ∂D.



Proof. Assume the condition of existence of the charts σ. It is easily seen that 0 is a regular value for the smooth function f defined by $f(\sigma(x)) = x_m$

on the open set $W = \sigma(U)$. Hence D is a domain with smooth boundary, and Lemma 4.6 shows that $\partial D \cap \sigma(U) = \sigma(U^\circ)$. It follows from the last statement in Theorem 4.3 that $\sigma(x_1, \dots, x_{m-1}, 0)$ is a chart on ∂D.

Conversely, assume that D is a domain with smooth boundary, and let $p \in \partial D$. It follows from Corollary 4.6 and Theorem 4.3 that there exists a chart $\sigma\colon]-\epsilon, \epsilon[^m \to M$ with $\sigma(0) = p$, such that $\sigma(x) \in \partial D$ if and only if $x_m = 0$. Let $y, z \in U$ be two points, for which y_m and z_m are either both positive or both negative. Then $x_m \neq 0$ on the line segment from y to z, and hence the curve $\gamma(t) = \sigma(y + t(z - y))$, where $t \in [0, 1]$, does not meet ∂D. If one, but not both, of the points $\sigma(y)$ and $\sigma(z)$ belongs to D, we reach a contradiction, since in that case some point along γ would belong to ∂D (if $\sigma(y) \in D$ and $\sigma(z) \notin D$ then $\gamma(t) \in \partial D$ where $t_0 = \sup\{t \mid \gamma(t) \in D\}$). Thus the two points both belong to D or they both belong to the complement. It follows that $\sigma(U) \cap D$ equals either $\sigma(U^+)$ or $\sigma(U^-)$. In the latter case we change sign in the last coordinate of σ, and thus in any case we reach that $\sigma(U) \cap D = \sigma(U^+)$. $\square$

4.7 Orientation of the boundary

Let D be a domain with smooth boundary in M. If an orientation of M is given, we can establish an orientation of ∂D in the following manner. Let $p \in \partial D$, then we have to decide whether a given basis $v_1, \dots, v_{m-1}$ for $T_p\partial D$ is positive or negative. We choose a neighborhood W and a function $f \in C^\infty(W)$ as in Definition 4.6. Since 0 is a regular value, there exists a vector $v_0 \in T_pM$ with $df_p(v_0) < 0$. The inequality means that the tangent vector v_0 points *out* from D. We now say that $v_1, \dots, v_{m-1}$ is a positive basis for $T_p\partial D$ if and only if $v_0, v_1, \dots, v_{m-1}$ is a positive basis for T_pM.

We have to verify that what we have introduced is in fact an orientation of the tangent space $T_p\partial D$, that is, if $v_1', \dots, v_{m-1}'$ is another basis with the same property that $v_0, v_1', \dots, v_{m-1}'$ is positive, then the transition matrix from $v_1, \dots, v_{m-1}$ to $v_1', \dots, v_{m-1}'$ has positive determinant. At the same time we can verify that the definition is independent of the choices of W, f and v_0. The details are omitted, as they are completely straightforward.

After that we also have to verify that what we have introduced is in fact an orientation of ∂D, that is, there exists an atlas which induces the above on all tangent spaces. Let σ be a chart on M as in Theorem 4.6. We can arrange, by changing the sign on x_1 if necessary, that the basis $-e_m, e_1, \dots, e_{m-1}$ for $\mathbb{R}^m$ is mapped to a positive basis for T_pM by $d\sigma_p$, for each $p \in U$. It then follows from the definition above that the corresponding chart on ∂D, given by $(x_1, \dots, x_{m-1}) \mapsto \sigma(x_1, \dots, x_{m-1}, 0)$, induces a positive basis for $T_p\partial D$. These charts constitute an atlas on ∂D, as desired.

Theorem 4.7. *If M is orientable then so is ∂D, for any domain $D \subset M$ with smooth boundary.*

Proof. See the discussion above. □

Example 4.7.1. Assume that D is a domain with smooth boundary in $M = \mathbb{R}^2$. Then D is open in $\mathbb{R}^2$ and the boundary $\mathcal{C} = \partial D$ is a curve in $\mathbb{R}^2$. For example, D could be the ball $\{(x, y) \mid x^2 + y^2 < 1\}$ and $\mathcal{C}$ the circle $\{(x, y) \mid x^2+y^2 = 1\}$. When $\mathbb{R}^2$ is oriented in the standard fashion, a vector v in $T_p\mathcal{C}$ has positive direction if and only if the normal vector $\hat{v}$ points into D.

Example 4.7.2. Assume similarly that D is a domain with smooth boundary in $M = \mathbb{R}^3$ (for example, the unit ball). Then the boundary $\mathcal{S} = \partial D$ is a surface in $\mathbb{R}^3$ (in the example, the unit sphere). With the standard orientation of $\mathbb{R}^3$, a basis v_1, v_2 for $T_p\mathcal{S}$ is positive if and only if $v_1 \times v_2$ points outward of D.

Example 4.7.3. Assume that M is a curve in $\mathbb{R}^n$ given by a parametrization $\gamma\colon I \to \mathbb{R}^n$, where I is an open interval. The parametrization determines an orientation of M. Let $D = \gamma([a, b])$ where $[a, b] \subset I$, then D is a domain with smooth boundary. The boundary is $\partial D = \{\gamma(a)\} \cup \{\gamma(b)\}$. By definition, an orientation of a point set is a choice of a sign $\pm$ at each point. The definition above is interpreted so that the sign is $+$ in the case where v_0, the tangent vector of M that points out of D, is positively oriented in M. This means that the orientation of ∂D is $+$ at $\gamma(b)$ and $-$ at $\gamma(a)$.

Example 4.7.4. Let M be the cylinder $\{(x, y, z) \mid x^2 + y^2 = 1\}$ in $\mathbb{R}^3$, oriented with outward positive normal, and let D be a band around the waist, say where $0 < z < 1$. The boundary ∂D consists of two circles. The upper circle at $z = 1$ is oriented clockwise, and the lower circle at $z = 0$ is oriented oppositely.

Example 4.7.5. Let N be an arbitrary oriented manifold, and let M be the product manifold $M = N \times \mathbb{R}$. The subset $D = N \times [0, 1]$ is a domain with smooth boundary $\partial D = (N \times \{0\}) \cup (N \times \{1\})$. The two copies of N are oppositely oriented, as in the preceding example.

4.8 Immersed submanifolds

The other notion of submanifolds, which was mentioned first in the chapter, is briefly mentioned. For clarity, in this subsection we call a submanifold (according to Definition 4.2) an *embedded* submanifold.

Definition 4.8.1. Let M be an abstract smooth manifold. An *immersed submanifold* is a subset $N \subset M$ equipped as a topological space (but not necessarily with the topology induced from M) and a smooth structure such that the inclusion map $i\colon N \to M$ is smooth with a differential $di_p\colon T_pN \to T_pM$ which is injective for each $p \in N$.

This is a weaker notion, because an embedded submanifold always carries the induced topology, whereas here we are just requiring some topology. It is required that the inclusion map is smooth, in particular continuous. Hence every relatively open subset in N is open for this topology, but not vice-versa. For example, the figure ∞ in $\mathbb{R}^2$ can be obtained as an immersed submanifold (see Example 1.2.2), but it is not a submanifold of $\mathbb{R}^2$.

Definition 4.8.2. Let M and N be abstract manifolds, and let $f\colon N \to M$ be a map. Then f is called an *embedding/immersion* if its image $f(N)$ can be given the structure of an embedded/immersed submanifold of M onto which f is a diffeomorphism.

Chapter 5

Topological properties of manifolds

In this chapter we investigate the consequences for a manifold, when certain topological properties are assumed. Furthermore, we develop an important analytical tool, called partition of unity.

5.1 Compactness

Recall that in a metric space X, a subset K is said to be *compact*, if every sequence from K has a subsequence which converges to a point in K. Recall also that every compact set is closed and bounded, and that the converse statement is valid for $X = \mathbb{R}^n$ with the standard metric, that is, the compact subsets of $\mathbb{R}^n$ are precisely the closed and bounded ones.

The generalization of compactness to an arbitrary topological space X does not invoke sequences. It originates from another important property of compact sets in a metric space, called the *Heine-Borel property*, which concerns coverings of K.

Let X be a Hausdorff topological space, and let $K \subset X$.

Definition 5.1.1. A *covering* of K is a collection of sets $U_i \subset X$, where $i \in I$, whose union $\cup_i U_i$ contains K. A *subcovering* is a subcollection $(U_j)_{j \in J}$, where $J \subset I$, which is again a covering of K. An *open covering* is a covering by open sets $U_i \subset X$.

Definition 5.1.2. The set K is said to be *compact* if every open covering has a finite subcovering.

It is a theorem that for a metric space the property in Definition 5.1.2 is equivalent with the property that K is compact (according to the definition with sequences), hence there is no conflict of notations. The space X itself is called compact, if it satisfies the above definition with $K = X$.

The following properties of compact sets are well known and easy to prove. Let X and Y be Hausdorff topological spaces.

Lemma 5.1.1. *Let $f\colon X \to Y$ be a continuous map. If $K \subset X$ is compact, then so is the image $f(K) \subset Y$.*

Example 5.7.1 The canonical map $S^2 \to \mathbb{RP}^2$ is continuous. Hence $\mathbb{RP}^2$ is compact.

Lemma 5.1.2. *Let $K \subset X$ be compact. Then K is closed, and every closed subset of K is compact.*

Lemma 5.1.3. *Assume that X is compact, and let $f: X \to Y$ be a continuous bijection. Then f is a homeomorphism.*

Proof. We have to show that f^{-1} is continuous, or equivalently, that f carries open sets to open sets. By taking complements, we see that it is also equivalent to show that f carries closed sets to closed sets. This follows from Lemma 5.1.1, in view of Lemma 5.1.2. $\square$

5.2 Countable exhaustion by compact sets

The following property of a manifold with a countable atlas will be used in the next section.

Theorem 5.2. *Let M be an abstract manifold. The following conditions are equivalent.*

(i) *There exists a countable atlas for M.*

(ii) *There exists a sequence $K_1, K_2, \dots$ of compact sets with union M.*

(iii) *There exists a sequence $D_1, D_2, \dots$ of open sets with union M, such that each D_n has compact closure $\bar{D}_n$.*

(iv) *There exists a sequence $K_1, K_2, \dots$ of compact sets with union M such that*

$$K_1 \subset K_2^\circ \subset \cdots \subset K_n \subset K_{n+1}^\circ \subset \dots.$$

Example 5.2 In $\mathbb{R}^n$ the sequence of concentric closed balls $K_n = \bar{B}(0, n)$ of radius n has the property (iv).

Proof. The implications (iv)$\Rightarrow$(iii)$\Rightarrow$(ii) are easily seen. The implication (ii)$\Rightarrow$(i) follows from the fact, seen from Definition 5.1.2, that every compact set $K \subset M$ can be covered by finitely many charts.

We establish (i)$\Rightarrow$(iii). For each chart $\sigma: U \to M$ the collection of closed balls $\bar{B}(x, r)$ in U with rational center and rational radius is countable, and the corresponding open balls cover U (see Example 2.9). Since σ is a homeomorphism, the collection of all the images $\sigma(B(x, r))$ of these balls, for all charts in a countable atlas, is a countable collection of open sets with the desired property.

Finally we prove that (iii)$\Rightarrow$(iv). Let $D_1, D_2, \dots$ be as in (iii). Put $K_1 = \bar{D}_1$. By the compactness of K_1 we have

$$K_1 \subset D_1 \cup D_2 \cup \cdots \cup D_{i_1}$$

for some number i_1. Put

$$K_2 = \bar{D}_1 \cup \bar{D}_2 \cup \cdots \cup \bar{D}_{i_1},$$

then K_2 is compact and $K_1 \subset K_2^\circ$. Again by compactness we have

$$K_2 \subset D_1 \cup D_2 \cup \cdots \cup D_{i_2}$$

for some number $i_2 > i_1$. Put

$$K_3 = \bar{D}_1 \cup \bar{D}_2 \cup \cdots \cup \bar{D}_{i_2},$$

then K_3 is compact and $K_2 \subset K_3^\circ$. Proceeding inductively in this fashion we obtain the desired sequence. $\square$

5.3 Locally finite atlas

It follows from Definition 5.1.2 that a compact manifold has a finite atlas. We shall define a much more general property.

Definition 5.3. A collection of subsets of a topological space X is said to be *locally finite*, if for each element $x \in X$ there exists a neighborhood which intersects non-trivially with only finitely many of the subsets.

An atlas of an abstract manifold M is said to be locally finite if the collection of images $\sigma(U)$ is locally finite in M.

In the following lemma we give a useful criterion for the existence of a locally finite atlas.

Lemma 5.3.1. *Let M be an abstract manifold. There exists a locally finite atlas for M if and only if the following criterion holds.*

There exists a covering $M = \cup_{\alpha\in A} K_\alpha$ of M by compact sets $K_\alpha \subset M$, and a locally finite covering $M = \cup_{\alpha\in A} W_\alpha$ by open sets $W_\alpha \subset M$ (with the same set A of indices), such that $K_\alpha \subset W_\alpha$ for each $\alpha \in A$.

Before proving the lemma, we shall verify that the criterion holds in case of a manifold with a countable atlas.

Lemma 5.3.2. *Let M be an abstract manifold for which there exists a countable atlas. Then the criterion in Lemma 5.3.1 holds for M.*

Proof. Let $L_1, L_2, \dots$ be an increasing sequence of compact sets in M as in Theorem 5.2(iv). Put $K_1 = L_1$ and $K_n = L_n \setminus L_{n-1}^\circ$ for $n > 1$, and put $W_1 = L_2^\circ$, $W_2 = L_3^\circ$ and $W_n = L_{n+1}^\circ \setminus L_{n-2}$ for $n > 2$. It is easily seen that the criterion is satisfied by these collections of sets. $\square$

Proof of Lemma 5.3.1. Assume that M has a locally finite atlas. For each chart σ in this atlas, we can apply the preceding lemma to the manifold $\sigma(U)$ (which has an atlas of a single chart). It follows that there exist collections of sets as described, which cover $\sigma(U)$. The combined collection of these sets, over all charts in the atlas, satisfies the desired criterion for M.

Assume conversely that the criterion holds for M. For each $\alpha \in A$ we can cover the compact set K_α by finitely many charts σ, each of which maps into the open set W_α. The combined collection of all these charts for all α is an atlas because $\cup_\alpha K_\alpha = M$. Each $p \in M$ has a neighborhood which is disjoint from all but finitely many W_α, hence also from all charts except the finite collection of those which maps into these W_α's. Hence this atlas is locally finite. $\square$

Corollary 5.3. *Every abstract manifold with a countable atlas has a locally finite atlas.*

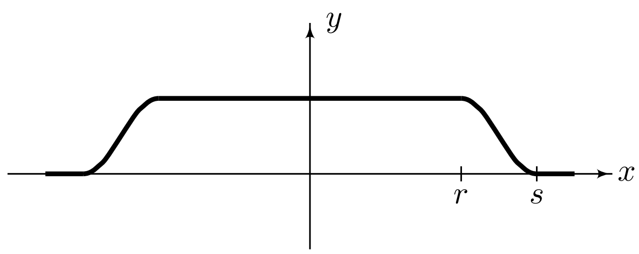
Proof. Follows immediately from the two lemmas (in fact, by going through the proofs above, one can verify that there exists an atlas which is both countable and locally finite). $\square$

5.4 Bump functions

The following lemma will be used several times in the future, when we need to pass between local and global objects on a manifold.

Lemma 5.4. *Let $\sigma\colon U \to M$ be a chart on an abstract manifold, and let $B(x_0, r) \subset \bar{B}(x_0, s) \subset U$ be concentric balls in U with $0 < r < s$. There exists a smooth function $g \in C^\infty(M)$, which takes values in $[0, 1]$, such that $g(q) = 1$ for $q \in \sigma(B(x_0, r))$ and $g(q) = 0$ for $q \notin \sigma(B(x_0, s))$.*

The function g is called a *bump* function around $p = \sigma(x_0)$, because of the resemblance with speed bumps used to reduce traffic. The following figure shows the graph of a bump function around $p = 0$ in $\mathbb{R}$.



Proof. It suffices to prove that there exists a function $\varphi \in C^\infty(\mathbb{R}^m)$ with values in $[0, 1]$ such that $\varphi(x) = 1$ for $|x| \le r$ and $\varphi(x) = 0$ for $|x| \ge s$, because then $g\colon M \to \mathbb{R}$ defined by

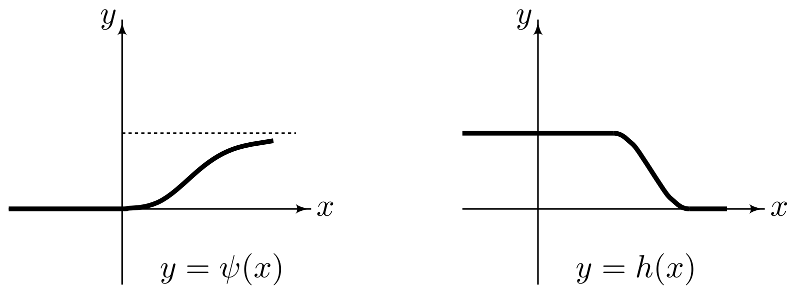
$$g(q) = \begin{cases} \varphi(y) & \text{if} \quad q = \sigma(y) \in \sigma(U) \\ g(q) = 0 & \text{otherwise,} \end{cases}$$

has the required properties. In particular g is smooth, because it is smooth on both $\sigma(B(x_0, s))$ (where it equals $\varphi \circ \sigma^{-1}$) and $M \setminus \sigma(\bar{B}(x_0, r))$ (where it is 0), hence on a neighborhood of every point in M.

In order to construct φ, we recall that the function ψ on $\mathbb{R}$, defined by $\psi(t) = e^{-1/t}$ for $t > 0$ and $\psi(t) = 0$ for $t \leq 0$, is smooth (see the graph below). The function

$$h(t) = \frac{\psi(s-t)}{\psi(s-t) + \psi(t-r)}$$

is smooth, and it takes the value 1 for $t \leq r$ and 0 for $t \geq s$. The function $\varphi(x) = h(|x|)$ has the required property. $\square$



5.5 Partition of unity

Recall that the *support*, $\operatorname{supp} f$, of a function $f: M \to \mathbb{R}$ is the closure of the set where f is non-zero.

Definition 5.5. Let M be an abstract manifold. A *partition of unity* for M is a collection $(f_\alpha)_{\alpha \in A}$ of functions $f_\alpha \in C^\infty(M)$ such that:
(1) $0 \leq f_\alpha \leq 1$,
(2) the collection of the supports $\operatorname{supp} f_\alpha$, $\alpha \in A$, is locally finite in M,
(3) $\sum_{\alpha \in A} f_\alpha(x) = 1$ for all $x \in M$.

Notice that because of condition (2) the (possibly infinite) sum in (3) has only finitely many non-zero terms for each x (but the non-zero terms are not necessarily the same for all x).

Theorem 5.5. *Let M be an abstract manifold for which there exists a locally finite atlas, and let $M = \cup_{\alpha \in A} \Omega_\alpha$ be an arbitrary open cover of M. Then there exists a partition of unity $(f_\alpha)_{\alpha \in A}$ for M (with the same set of indices), such that f_α has support inside Ω_α for each α.*

Proof. Let $M = \cup_{\beta \in B} K_\beta$ and $M = \cup_{\beta \in B} W_\beta$ be coverings as in Lemma 5.3.1, with $K_\beta \subset W_\beta$. Let $\beta \in B$ be arbitrary. For each $p \in K_\beta$ we choose $\alpha \in A$ such that $p \in \Omega_\alpha$, and we choose a chart $\sigma: U \to M$ (not necessarily from the given atlas) with $\sigma(x) = p$ and with image $\sigma(U) \subset \Omega_\alpha \cap W_\beta$. Furthermore, in U we choose a pair of concentric open balls $B(x,r) \subsetneq B(x,s)$ around x, such that the closure of the larger ball $B(x,s)$ is contained in U. Since K_β is compact a finite collection of the images of the smaller balls $\sigma(B(x,r))$ cover it. We choose such a finite collection of concentric balls for each β.

Since the sets K_β have union M, the combined collection of all the images of inner balls, over all β, covers M. Furthermore, for each $p \in M$ there exists a neighborhood which meets only finitely many of the images of the outer balls, since the W_β are locally finite. That is, the collection of the images of the outer balls is locally finite.

The rest of the proof is based on Lemma 5.4. For each of the above mentioned pairs of concentric balls we choose a smooth function $g \in C^\infty(M)$ with the properties mentioned in this lemma. Let $(g_i)_{i\in I}$ denote the collection of all these functions. By the remarks in the preceding paragraph, we see that for each $p \in M$ there exists some g_i with $g_i(p) = 1$, and only finitely many of the functions g_i are non-zero at p.

Hence the sum $g = \sum_i g_i$ has only finitely many terms in a neighborhood of each point p. It follows that the sum makes sense and defines a positive smooth function. Let $f_i = g_i/g$, then this is a partition of unity. Moreover, for each i there exists an $\alpha \in A$ such that g_i, hence also f_i, is supported inside Ω_α.

We need to fix the index set for the g_i so that it is A. For each i choose α_i such that f_i has support inside Ω_{α_i}. For each $\alpha \in A$ let f_α denote the sum of those f_i for which $\alpha_i = \alpha$, if there are any. Let $f_\alpha = 0$ otherwise. The result follows easily. $\square$

Corollary 5.5. *Let M be an abstract manifold with a locally finite atlas, and let C_0, C_1 be closed, disjoint subsets. There exists a smooth function $f \in C^\infty(M)$ with values in $[0,1]$, which is 1 on C_0 and 0 on C_1.*

Proof. Apply the theorem to the covering of M by the complements of C_1 and C_0. $\square$

5.6 Embedding in Euclidean space

As an illustration of the use of partitions of unity, we shall prove the following theorem, which is a weak version of Whitney's Theorem 2.10. Notice that the theorem applies to all compact manifolds. Recall that by definition an embedding is the same as a diffeomorphism onto a submanifold.

Theorem 5.6. *Let M be an abstract manifold, for which there exists a finite atlas. Then there exists a number $N \in \mathbb{N}$ and an embedding of M into $\mathbb{R}^N$.*

Proof. Let $\sigma_i\colon U_i \to M$, where $i = 1, \dots, n$ be a collection of charts on M, which comprise an atlas. We will prove the theorem with the value $N = n(m+1)$ where $m = \dim M$.

Let $f_1, \dots, f_n$ be a partition of unity for M such that $\operatorname{supp} f_i \subset \sigma_i(U_i)$ for each i. Its existence follows from Theorem 5.5. Then $f_i \in C^\infty(M)$ and $f_1 + \dots + f_n = 1$.

For each i, let $\varphi_i = \sigma_i^{-1}\colon \sigma_i(U_i) \to U_i \subset R^m$ denote the inverse of σ_i. It is a smooth map. We define a function $h^i\colon M \to \mathbb{R}^m$ by

$$h^i(p) = f_i(p)\varphi_i(p)$$

for $p \in \sigma_i(U_i)$ and by $h^i(p) = 0$ outside this set. Then h^i is smooth since the support of f_i is entirely within $\sigma_i(U_i)$, so that every point in M has a neighborhood either entirely inside $\sigma_i(U_i)$ or entirely inside the set where f_i, and hence also h^i, is 0.

We now define a smooth map $F\colon M \to \mathbb{R}^N = \mathbb{R}^m \times \cdots \times \mathbb{R}^m \times \mathbb{R}^n$ by

$$F(p) = (h^1(p), \dots, h^n(p), f_1(p), \dots, f_n(p)).$$

We will show that F is injective and a homeomorphism onto its image $F(M)$. Let $\Omega_i = \{p \in M \mid f_i(p) \neq 0\}$, for $i = 1, \dots, n$, then these sets are open and cover M. Likewise, let

$$W_i = \{(x^1, \dots, x^n, y_1, \dots, y_n) \in \mathbb{R}^N \mid y_i \neq 0, \frac{x_i}{y_i} \in U_i\},$$

where $x^i \in \mathbb{R}^m$ and $y_i \in \mathbb{R}$ for $i = 1, \dots, n$. Then W_i is open in $\mathbb{R}^N$, and $F(\Omega_i) \subset W_i$.

Define $g_i\colon W_i \to M$ by

$$g_i(x^1, \dots, x^n, y_1, \dots, y_n) = \sigma(x^i/y_i)$$

then g_i is clearly smooth, and we see that $g_i(F(p)) = p$ for $p \in \Omega_i$.

It follows now that F is injective, for if $F(p) = F(q)$, then since the values $f_1(p), \dots, f_n(p)$ are among the coordinates of $F(p)$, we conclude that $f_i(p) = f_i(q)$ for each i. In particular, $p \in \Omega_i$ if and only if $q \in \Omega_i$. We choose i such that $p, q \in \Omega_i$ and conclude $p = g_i(F(p)) = g_i(F(q)) = q$.

Furthermore, on the open set $F(M) \cap W_i$ in $F(M)$, the inverse of F is given by the restriction of g_i, which is continuous. Hence F is a homeomorphism onto its image.

For each $i = 1, \dots, n$ we now see that the restriction of $F \circ \sigma_i$ to $\sigma_i^{-1}(\Omega_i)$ is an embedded m-dimensional parametrized manifold in $\mathbb{R}^N$, whose image is $F(M) \cap W_i$. It follows that $F(M)$ is a manifold in $\mathbb{R}^N$.

The map F is smooth and bijective $M \to F(M)$, and we have seen above that the inverse has local smooth extensions g_i to each set W_i. Hence the inverse is smooth, and F is a diffeomorphism onto its image. $\square$

Example 5.6 In Section 2.4 we equipped the projective space $M = \mathbb{R}\mathrm{P}^2$ with an atlas consisting of $n = 3$ charts. Hence it follows from the proof just given that there exists a diffeomorphism of it into $\mathbb{R}^9$ (in fact, as mentioned below Theorem 2.10, there exists an embedding into $\mathbb{R}^4$).

5.7 Connectedness

In this section two different notions of connectedness for subsets of a topological space are introduced and discussed. Let X be a non-empty topological space.

Definition 5.7. (1) X is said to be *connected* if it cannot be separated in two disjoint non-empty open subsets, that is, if $X = A_1 \cup A_2$ with A_1, A_2 open and disjoint, then A_1 or A_2 is empty (and A_2 or A_1 equals X).

(2) X is called *pathwise connected* if for each pair of points $a, b \in S$ there exists real numbers $\alpha \leq \beta$ and a continuous map $\gamma\colon [\alpha, \beta] \to X$ such that $\gamma(\alpha) = a$ and $\gamma(\beta) = b$ (in which case we say that a and b can be joined by a continuous path in X).

(3) A non-empty subset $E \subset X$ is called connected or pathwise connected if it has this property as a topological space with the induced topology.

The above definition of "connected" is standard in the theory of topological spaces. However, the notion of "pathwise connected" is unfortunately sometimes also referred to as "connected". The precise relation between the two notions will be explained in this section and the following. The empty set was excluded in the definition, let us agree to call it both connected and pathwise connected.

Example 5.7.1 A singleton $E = \{x\} \subset X$ is clearly both connected and pathwise connected.

Example 5.7.2 A convex subset $E \subset \mathbb{R}^n$ is pathwise connected, since by definition any two points from E can be joined by a straight line, hence a continuous curve, inside E. It follows from Theorem 5.7.3 below that such a subset is also connected.

Example 5.7.3 It is a well-known fact, called the *intermediate value property*, that a continuous real function carries intervals to intervals. It follows from this fact that a subset $E \subset \mathbb{R}$ is pathwise connected if and only if it is an interval. We shall see below in Theorem 5.7.1 that likewise E is connected if and only if it is an interval. Thus for subsets of $\mathbb{R}$ the two definitions agree.

Lemma 5.7. *Let $(E_i)_{i \in I}$ be a collection of subsets of X, and let $E_0 \subset X$ be a subset with the property that $E_i \cap E_0 \neq \emptyset$ for all i.*

If both E_0 and all the sets E_i are connected, respectively pathwise connected, then so is their union $E = E_0 \cup (\cup_i E_i)$.

Proof. Assume that E_0 and the E_i are connected. and assume that E is separated in a disjoint union $E = A_1 \cup A_2$ where A_1, A_2 are relatively open in E. Then $A_1 = W_1 \cap E$ and $A_2 = W_2 \cap E$, where W_1, W_2 are open in X. Hence the intersections $A_1 \cap E_i = W_1 \cap E_i$ and $A_2 \cap E_i = W_2 \cap E_i$ are relatively open in E_i for all i (including $i = 0$).

It follows that $E_i = (A_1 \cap E_i) \cup (A_2 \cap E_i)$ is a disjoint separation in open sets, and since E_i is connected, one of the sets $A_1 \cap E_i$ and $A_2 \cap E_i$ must be empty, for each i (including $i = 0$). If for example $A_1 \cap E_0 = \emptyset$, then E_0 is contained in A_2, and since all the other E_i have nontrivial intersection with E_0, we conclude that then $E_i \cap A_2 \neq \emptyset$ for all i. Hence the intersection $E_i \cap A_1$ is empty for all i, and we conclude that A_1 is empty. We have shown that E is connected.

Assume next that E_0 and all the E_i are pathwise connected. Since all the E_i have non-trivial intersection with E_0, every point in E can be joined to a point in E_0 by a continuous path. Hence any two points of E can be joined by a continuous path, composed by the paths that join them to two points in E_0, and a path in E_0 that joins these two points. $\square$

Example 5.7.4 Let $H \subset \mathbb{R}^3$ be the surface

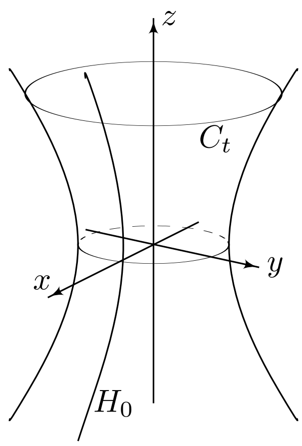
$$\{(x, y, z) \mid x^2 + y^2 - z^2 = 1\}$$

called the *one-sheeted hyperboloid*. We will prove by means of Lemma 5.7 that it is pathwise connected. Let

$$H_0 = \{(x, 0, z) \mid x^2 - z^2 = 1, x > 0\} \subset H$$

and for each $t \in \mathbb{R}$,

$$C_t = \{(x, y, t) \mid x^2 + y^2 = 1 + t^2\} \subset H.$$



Then H is the union of these sets. The set H_0 is pathwise connected because it is the image of the continuous curve $t \mapsto (\sqrt{1+t^2}, 0, t)$, and each set C_t is a circle, hence also pathwise connected. Finally, $H_0 \cap C_t$ is non-empty for all t, as it contains the point $(\sqrt{1+t^2}, 0, t)$.

Theorem 5.7.1. *Let $E \subset \mathbb{R}$ be non-empty. Then E is connected if and only if it is an interval.*

Proof. Assume that E is connected. Let $a = \inf E$ and $b = \sup E$. For each element $c \in \mathbb{R}$ with $a < c < b$ the sets $E \cap]-\infty, c[$ and $E \cap]c, \infty[$ are disjoint and open in E, and it follows from the definitions of a and b that they are both non-empty. Since E is connected, their union cannot be E, hence we conclude that $c \in E$. Hence E is an interval with endpoints a and b.

Assume conversely that E is an interval, and that $E = A \cup B$ where A and B are open, disjoint and non-empty. Since they have open complements, A and B are also closed in E. Choose $a \in A$ and $b \in B$, and assume for example that $a < b$. Then $[a, b] \subset E$, since E is an interval. The set $[a, b] \cap A$ is non-empty, since it contains a, let c be its supremum. Since A is closed it contains c, hence $c < b$ and $]c, b] \subset B$. Since B is closed it also contains c, contradicting that A and B are disjoint. □

One of the most fundamental property of connected sets is expressed in the following theorem, which generalizes the intermediate value property for real functions on $\mathbb{R}$ (see Example 5.7.3).

Theorem 5.7.2. *Let $f: X \to Y$ be a continuous map between topological spaces. If $E \subset X$ is connected, then so is the image $f(E) \subset Y$. Likewise, if E is pathwise connected then so is $f(E)$.*

Proof. We may assume $E = X$ (otherwise we replace X by E).

1) Assume $f(X) = B_1 \cup B_2$ with B_1, B_2 open and disjoint, and let $A_i = f^{-1}(B_i)$. Then A_1, A_2 are open, disjoint and with union X. Hence if X is connected then A_1 or A_2 is empty, and hence B_1 or B_2 is empty.

2) If $a, b \in X$ can be joined by a continuous path γ, then $f(a)$ and $f(b)$ are joined by the continuous path $f \circ \gamma$. □

Theorem 5.7.3. *A pathwise connected topological space is also connected.*

Proof. Suppose X were pathwise connected but not connected. Then $X = A \cup B$ with A, B open, disjoint and nonempty. Let $a \in A$, $b \in B$, then there exists a continuous path $\gamma: [\alpha, \beta] \to X$ joining a to b. The image $C = \gamma([\alpha, \beta])$ is the disjoint union of $C \cap A$ and $C \cap B$. These sets are open subsets of the topological space C, and they are nonempty since they contain a and b, respectively. Hence C is not connected. On the other hand, since the interval $[\alpha, \beta]$ is connected, it follows from Theorem 5.7.2 that $C = \gamma([\alpha, \beta])$ is connected, so that we have reached a contradiction. □

The converse statement is false. There exists subsets of, for example $\mathbb{R}^n$ ($n \geq 2$), which are connected but not pathwise connected (an example in $\mathbb{R}^2$ is given below). However, for open subsets of $\mathbb{R}^n$ the two notions of connectedness agree. This will be proved in the following section.

Example 5.7.5 The graph of the function

$$f(x) = \begin{cases} \sin(1/x) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

is connected but not pathwise connected.

5.8 Connected manifolds

In this section we explore the relation between the two types of connectedness, with the case in mind that the topological space X is an abstract manifold.

Definition 5.8. A topological space X is said to be *locally pathwise connected* if it has the following property. For each point $x \in X$ and each neighborhood V there exists an open pathwise connected set U such that $x \in U \subset V$.

Example 5.8 The space $X = \mathbb{R}^n$ is locally pathwise connected, since all open balls are pathwise connected. The same is valid for an abstract manifold, since each point has a neighborhood, which is the image by a chart of a ball in $\mathbb{R}^m$, hence pathwise connected.

Lemma 5.8. *In a locally pathwise connected topological space, all open connected sets E are pathwise connected.*

Proof. For $a, b \in E$ we write $a \sim b$ if a and b can be joined by a continuous path in E. It is easily seen that this is an equivalence relation. Since E is open there exists for each $a \in E$ an open neighborhood $V \subset E$, hence an open pathwise connected set U with $a \in U \subset E$. For all points x in U we thus have $a \sim x$. It follows that the equivalence classes for $\sim$ are open. Let A be an arbitrary of these equivalence classes, and let B denote the union of all other equivalence classes. Then A and B are open, disjoint and have union E. Since E is connected, A or B is empty. Since $a \in A$, we conclude that $B = \emptyset$ and $A = E$. Hence all points of E are equivalent with each other, which means that E is pathwise connected. $\square$

Theorem 5.8. *Let M be an abstract smooth manifold. Each open connected subset E of M is also pathwise connected.*

Proof. Follows immediately from Lemma 5.8, in view of Example 5.8. $\square$

In particular, every open connected subset of $\mathbb{R}^n$ is pathwise connected.

5.9 Components

Let X be topological space. We shall determine a decomposition of X as a disjoint union of connected subsets. For example, the set $\mathbb{R}^\times = \mathbb{R} \setminus \{0\}$ is the disjoint union of the connected subsets $]-\infty, 0[$ and $]0, \infty[$.

Definition 5.9. *A* component *(or* connected component*) of X is a subset $E \subset X$ which is maximal connected, that is, it is connected and not properly contained in any other connected subset.*

Components are always closed. This is a consequence of the following lemma (notice that in the example above, the components $]-\infty, 0[$ and $]0, \infty[$ are really closed in $\mathbb{R}^\times$).

Lemma 5.9. *The closure of a connected subset $E \subset X$ is connected.*

Proof. Assume that $\bar{E}$ is separated in a disjoint union $\bar{E} = A_1 \cup A_2$ where A_1, A_2 are relatively open in $\bar{E}$. Then $A_i = W_i \cap \bar{E}$ for $i = 1, 2$, where W_1 and W_2 are open in X. Hence each set $W_i \cap E = A_i \cap E$ is relatively open in E, and these two sets separate E in a disjoint union. Since E is connected, one of the two sets is empty. If for example $W_1 \cap E = \emptyset$, then E is contained in the complement of W_1, which is closed in X. Hence also $\bar{E}$ is contained in this complement, and we conclude that $A_1 = \bar{E} \cap W_1$ is empty. □

Theorem 5.9. *X is the disjoint union of its components. If X is locally pathwise connected, for example if it is an abstract manifold, then the components are open and pathwise connected.*

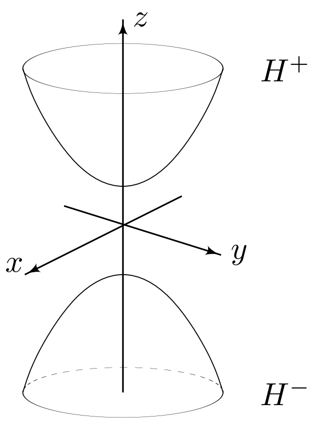
Proof. Let $x \in X$ be arbitrary. It follows from Lemma 5.7 with $E_0 = \{x\}$, that the union of all the connected sets in X that contain x, is connected. Clearly this union is maximal connected, hence a component. Hence X is the union of its components. The union is disjoint, because if two different components overlapped, their union would be connected, again by Lemma 5.7, and hence none of them would be maximal.

Assume that X is locally pathwise connected, and let $E \subset X$ be a component of X. For each $x \in E$ there exists a pathwise connected open neighborhood of x in X. This neighborhood must be contained in E (by maximality of E). Hence E is open. Now Theorem 5.8 implies that E is pathwise connected. □

Example 5.9 Let $H \subset \mathbb{R}^3$ be the surface $\{(x, y, z) \mid x^2 + y^2 - z^2 = -1\}$, called the *two-sheeted hyperboloid.* We claim it has the two components

$$H^+ = \{(x, y, z) \in H \mid z > 0\}, \quad H^- \{(x, y, z) \in H \mid z < 0\}.$$

It is easily seen that $|z| \geq 1$ for all $(x, y, z) \in H$, hence $H = H^+ \cup H^-$, a disjoint union. The verification that H^+ and H^- are pathwise connected is similar to that of Example 5.7.4.



Corollary 5.9. *Let M be an abstract manifold. The components of M are abstract manifolds (of the same dimension). If an atlas is given for each component, then the combined collection of the charts comprises an atlas for M.*

Proof. Follows from the fact that the components are open and cover M. □

In particular, since there is no overlap between the components, M is orientable if and only if each of its components is orientable.

5.10 The Jordan-Brouwer theorem

The proof of the following theorem is too difficult to be given here.

Theorem 5.10. *Let $M \subset \mathbb{R}^n$ be an $n-1$–dimensional compact connected manifold in $\mathbb{R}^n$. The complement of M in $\mathbb{R}^n$ consists of precisely two components, of which one, called the* outside *is unbounded, and the other, called the* inside *is bounded. Each of the two components is a domain with smooth boundary M.*

The Jordan curve theorem for smooth plane curves is obtained in the special case $n = 2$.

Example 5.10 Let $M = S^{n-1}$ denote the unit sphere in $\mathbb{R}^n$. Its inside is the open n-ball $\{x \mid \|x\| < 1\}$, and the outside is the set $\{x \mid \|x\| > 1\}$.

Corollary 5.10. *Let M be an $n-1$–dimensional compact manifold in $\mathbb{R}^n$. Then M is orientable.*

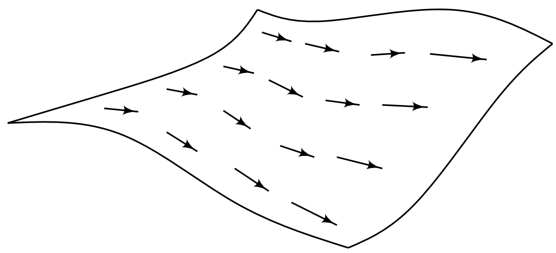
Proof. It suffices to prove that each component of M is orientable, so we may as well assume that M is connected. The inside of M is orientable, being an open subset of $\mathbb{R}^n$. Hence it follows from Theorem 4.7 that M is orientable. □

In particular, all compact surfaces in $\mathbb{R}^3$ are orientable.

Chapter 6

Vector fields and Lie algebras

By definition, a vector field on a manifold M assigns a tangent vector $Y(p) \in T_pM$ to each point $p \in M$. One can visualize the vector field as a collection of arrows, where $Y(p)$ is placed as an arrow tangent to the manifold and with its basis at p.



6.1 Smooth vector fields

We would like the vector field Y to vary smoothly with p. Since $Y: p \mapsto Y(p)$ maps into a space that varies with p, the precise formulation of smoothness requires some care. For vector fields on a manifold $\mathcal{S}$ in $\mathbb{R}^n$ the formulation is easy, since we can express the smoothness by means of the ambient space $\mathbb{R}^n$.

Definition 6.1.1. A *smooth vector field* on a manifold $\mathcal{S} \subset \mathbb{R}^n$ is a smooth map $Y: \mathcal{S} \to \mathbb{R}^n$ such that $Y(p) \in T_p\mathcal{S}$ for all $p \in \mathcal{S}$.

The following result gives a reformulation, which paves the way for the abstract case.

Lemma 6.1.1. *Let $\mathcal{S} \subset \mathbb{R}^n$ be a manifold and let $Y: \mathcal{S} \to \mathbb{R}^n$ be such that $Y(p) \in T_p\mathcal{S}$ for all $p \in \mathcal{S}$. The following conditions are equivalent*
(a) *Y is smooth.*
(b) *For each chart $\sigma: U \to \mathcal{S}$ we have*

$$Y(\sigma(u)) = \sum_{i=1}^{m} a_i(u)\sigma'_{u_i}(u) \tag{6.1}$$

for all $u \in U$, with coefficient functions $a_1, \dots, a_m$ in $C^\infty(U)$.

(c) *The same as* (b), *but only for σ in a given atlas.*

Proof. By Theorem 2.6 the first condition is equivalent to $Y \circ \sigma$ being smooth $U \to \mathbb{R}^n$ for all charts σ in a given atlas. It follows immediately from (6.1) that this is the case if the functions $a_1, \dots, a_m$ are smooth. Thus (c) implies (a).

The implication (b)$\Rightarrow$(c) is clear. Finally we will derive (b) from (a). Since the elements $\sigma'_{u_1}(u), \dots, \sigma'_{u_m}(u)$ form a basis for $T_{\sigma(u)}M$, there exist for each $u \in U$ some unique real numbers $a_1(u), \dots, a_m(u)$ such that (6.1) holds. The claim is that (a) implies these numbers depend smoothly on u.

Let $p = \sigma(u_0) \in \sigma(U)$ be given. It follows from Theorem 1.7 that σ^{-1} has a smooth extension φ to a neighborhood W of p in $\mathbb{R}^n$. Then $\varphi \circ \sigma$ is the identity map in a neighborhood of u_0. The entries in the Jacobian matrix $D\varphi(x)$ of φ depend smoothly on $x \in W$. According to the chain rule we have $D\varphi(\sigma(u))D\sigma(u) = I$ for all $u \in U$. Let $a(u) \in \mathbb{R}^m$ denote the column whose elements are the numbers $a_1(u), \dots, a_m(u)$. The vectors $\sigma'_{u_i}(u)$ in (6.1) are the columns of $D\sigma(u)$, and the expression on the right of (6.1) is $D\sigma(u)a(u)$. Hence it follows from (6.1) that

$$D\varphi(\sigma(u))Y(\sigma(u)) = D\varphi(\sigma(u))D\sigma(u)a(u) = a(u)$$

and hence $a(u)$ depends smoothly on u. $\square$

Let now M be an abstract smooth manifold. We will use a generalized version of the conditions (b)-(c) in our definition of a smooth vector field on M. Recall from Section 3.8, that we introduced the standard basis vectors for the tangent space T_pM at p of M,

$$d\sigma_u(e_1), \dots, d\sigma_u(e_m) \tag{6.2}$$

where $\sigma(u) = p$, and where $e_i \in \mathbb{R}^m$ are the canonical basis vectors. These are the analogs of the vectors $\sigma'_{u_i}(u)$ in (6.1).

Definition 6.1.2. A vector field on M is an assignment of a tangent vector $Y(p) \in T_pM$ to each $p \in M$. It is called a *smooth vector field* if the following condition holds for each chart $\sigma: U \to M$ in a given atlas of M. There exist $a_1, \dots, a_m$ in $C^\infty(U)$ such that

$$Y(\sigma(u)) = \sum_{i=1}^m a_i(u) d\sigma_u(e_i) \tag{6.3}$$

for all $u \in U$.

The space of smooth vector fields on M is denoted $\mathfrak{X}(M)$.

It is a consequence of Lemma 6.2.2, to be proved below, that the definition is unchanged if the atlas is replaced by a compatible one. The concept introduced thus only depends on the smooth structure of M.

Notice that each chart σ in the given atlas for M is a diffeomorphism $U \to \sigma(U)$, and the coefficients $a_1(u), \dots, a_n(u)$ in (6.3) are the components of the vector $d(\sigma^{-1})_p(Y(p)) \in \mathbb{R}^m$, where $p = \sigma(u)$ (this follows from the fact that $d(\sigma^{-1})_p = (d\sigma_u)^{-1}$). Hence the condition of smoothness can also be phrased as the condition that

$$u \mapsto d(\sigma^{-1})_{\sigma(u)}(Y(\sigma(u))) \tag{6.4}$$

is smooth.

Example 6.1 Let $M \subset \mathbb{R}^m$ be an open subset, regarded as a manifold with the chart of the identity map $\mathbb{R}^m \to \mathbb{R}^m$. We know that at each point $p \in M$ the tangent space T_pM can be identified with $\mathbb{R}^m$, and a smooth vector field on M is nothing but a smooth map $Y\colon M \to \mathbb{R}^m$.

If we regard tangent vectors to $\mathbb{R}^m$ as directional differentiation operators (see Section 3.6), the standard basis vectors are the partial derivative operators $\frac{\partial}{\partial x_i}$, for $i = 1, \dots, m$. The operator corresponding to Y is the partial differential operator

$$\sum_{i=1}^{m} a_i(x) \frac{\partial}{\partial x_i},$$

where the coefficients $a_1, \dots, a_m \in C^\infty(M)$ are the components of Y. In other words, we can think of a smooth vector field on $\mathbb{R}^m$ as a first order partial differential operator with smooth coefficients.

Definition 6.1.3. Let $f \in C^\infty(M)$ and $Y \in \mathfrak{X}(M)$. The *product* $fY \in \mathfrak{X}(M)$ is defined by $(fY)(p) = f(p)Y(p)$ for all $p \in M$.

The fact that Yf is smooth is easily seen from Definition 6.1.2. Smooth vector fields satisfy the following generalized Leibniz rule

Lemma 6.1.2. *Let* $f, g \in C^\infty(M)$ *and* $Y \in \mathfrak{X}(M)$. *Then*

$$Y(fg) = fY(g) + gY(f).$$

Proof. Let σ be a chart on M, and let (6.3) be the associated expression of Y. Recall

$$\mathrm{D}_{d\sigma_u(e_i)} f = \frac{\partial}{\partial u_i}(f \circ \sigma)(u) \tag{6.5}$$

(see (3.12)). Hence on the image $\sigma(U)$ we have

$$Y(fg) = \sum_i a_i \frac{\partial}{\partial u_i}[(fg) \circ \sigma].$$

The lemma now follows by applying the usual Leibniz rule to the differentiation of the product $[(fg) \circ \sigma] = (f \circ \sigma)(g \circ \sigma)$. $\square$

6.2 An equivalent formulation of smoothness

Let Y be a vector field on an abstract manifold M. The interpretation in Example 6.1 of a vector field as a first order differential operator can be rendered in this general case as well, by means of the definitions in Sections 3.6 and 3.8. Recall from Definition 3.6 that a member $X = [\gamma]_p$ of the tangent space T_pM can be brought to act on a smooth function f through operator D_X given by $\mathrm{D}_X(f) = (f \circ \gamma)'(t_0)$. For $f \in C^\infty(\Omega)$, where $\Omega \subset M$ is open, it thus makes sense to apply $Y(p)$ to f for each $p \in \Omega$. We denote by $Yf = Y(f)$ the resulting function $p \mapsto Y(p)f = \mathrm{D}_{Y(p)}f$ on Ω.

It is convenient to introduce the *coordinate functions* $\xi_1, \dots, \xi_m$ associated with a chart $\sigma\colon U \to M$. They map $\sigma(U) \to U$ and are defined by

$$\xi_i(\sigma(u)) = u_i, \quad u \in U.$$

Lemma 6.2.1. *The coordinate functions ξ_i belong to $C^\infty(\sigma(U))$. For each $X \in T_pM$ the coordinates of X with respect to the standard basis* (6.2) *are $\mathrm{D}_X(\xi_1), \dots, \mathrm{D}_X(\xi_m)$.*

Proof. The function ξ_i is equal to σ^{-1} composed with projection on the i-th coordinate of $\mathbb{R}^m$. Hence it is smooth (see Example 2.7.1). If $X = \sum_i a_i d\sigma_u(e_i) \in T_pM$ we derive from (6.5)

$$\mathrm{D}_X(\xi_j) = \sum_i a_i \frac{\partial}{\partial u_i}(\xi_j \circ \sigma)(u) = \sum_i a_i \frac{\partial}{\partial u_i}(u_j) = a_j. \quad \square$$

Notice that it follows from the preceding lemma, that a tangent vector $X \in T_pM$ is *uniquely* determined by its action D_X on functions. If we know $D_X f$ for all smooth functions f defined on arbitrary neighborhoods of p, then we can determine X by taking $f = \xi_1, \dots, f = \xi_m$ with respect to some chart.

Lemma 6.2.2. *Let Y be a vector field on an abstract manifold M. The following conditions are equivalent:*
(i) *Y is smooth,*
(ii) *$Yf \in C^\infty(M)$ for all $f \in C^\infty(M)$,*
(iii) *$Yf \in C^\infty(\Omega)$ for all open sets $\Omega \subset M$ and all $f \in C^\infty(\Omega)$.*

Proof. Let σ be a chart on M, and let (6.3) be the associated expression of Y. It follows from (6.5) that

$$Yf(\sigma(u)) = \sum_{i=1}^m a_i(u) \mathrm{D}_{d\sigma_u(e_i)} f = \sum_{i=1}^m a_i(u) \frac{\partial}{\partial u_i}(f \circ \sigma)(u). \tag{6.6}$$

The implication of (i)$\Rightarrow$(ii) will be proved from this equation.

Assume (i) and let $f \in C^\infty(M)$. Assume that σ belongs to the given atlas. Then the a_i are smooth functions of u, and it follows from (6.6) that $Yf \circ \sigma$ is smooth. Since σ was arbitrary within an atlas, Yf is smooth. This proves (ii).

We prove (ii)⇒(iii). Let $\Omega \subset M$ be open, and let $f \in C^\infty(\Omega)$. Let $p \in \Omega$ be given, and chose a chart σ on M with $p \in \sigma(U) \subset \Omega$. Let $g \in C^\infty(\sigma(U))$ be a smooth 'bump' function around p, as in Lemma 5.4 (more precisely, g is the bump function composed with σ^{-1}). Then, as g is zero outside a closed subset of $\sigma(U)$, the function h on M defined by the product $h = gf$ on $\sigma(U)$ and $h = 0$ otherwise, belongs to $C^\infty(M)$. Furthermore $h = f$ in a neighborhood of p, and hence $Yh = Yf$ in that neighborhood. Since Yh is smooth by assumption (ii), it follows that Yf is smooth in a neighborhood of p. Since p was arbitrary in Ω, (iii) follows,

For the last implication, (iii)⇒(i), we apply Lemma 6.2.1. Let σ be an arbitrary chart on M, and let ξ_i be a coordinate function. It follows from (iii) with $\Omega = \sigma(U)$ that $Y\xi_i$ is smooth on this set, and it follows from the last statement in Lemma 6.2.1 that (6.3) holds with $a_i = Y\xi_i$. Hence Y is smooth. □

Notice that in the preceding proof we determined the coefficient a_j of the vector field Y from the action of Y on the function ξ_j, in a neighborhood of the given point p. In particular, it follows that the vector field is uniquely determined by its action on functions. Because of this, it is quite customary to identify a smooth vector field Y on M with its action on smooth functions, and thus regard the operator $f \mapsto Yf$ as being the vector field itself.

6.3 The tangent bundle

In Definition 6.1.2 we defined a vector field on an abstract manifold M as an assignment of a tangent vector $Y(p)$ to each point $p \in M$. Thus Y is a map from M into the set of all tangent vectors at all points of M. It turns out that this set of tangent vectors can be given a differential structure of its own, by means of which the smoothness of Y can be elegantly expressed.

Definition 6.3. The *tangent bundle* of M is the union $TM = \cup_{p\in M} T_pM$ of all tangent vectors at all points.

For a given element $X \in TM$ the point $p \in M$ for which $X \in T_pM$ is called the *base point*. The map $\pi: TM \to M$ which assigns p to X, is called the *projection*.

Notice that the union we take is *disjoint*, that is, there is no overlap between $T_{p_1}M$ and $T_{p_2}M$ if $p_1 \neq p_2$. Formally, an element in TM is actually a *pair* (p, X) where $p \in M$ and $X \in T_pM$, and π maps (p, X) to the first member of the pair. Notationally it is too cumbersome to denote elements in this fashion, and hence the base point p is suppressed.

If $M = \mathcal{S}$, a manifold in $\mathbb{R}^n$, then we are accustomed to viewing tangent vectors as elements in the ambient space, and of course the same vector $v \in \mathbb{R}^n$ can easily be tangent vector to $\mathcal{S}$ at several points. For example, two antipodal points on a sphere have identical tangent spaces. However, in the tangent bundle TM it is important that we make distinction between copies of v which are attached to different base points (otherwise, for example, the projection would not be well-defined). Thus the tangent bundle of a manifold in $\mathbb{R}^n$ is *not* to be conceived as a subset of $\mathbb{R}^n$.

Example 6.3.1 Let $U \subset \mathbb{R}^m$ be an open set, which we view as an m-dimensional manifold. At each point $x \in U$ the tangent space T_xU is a copy of $\mathbb{R}^m$, hence we can identify the tangent bundle of U as $TU = U \times \mathbb{R}^m$. The projection π is given by $\pi(x, y) = x$.

Let $f\colon M \to N$ be a smooth map between manifolds, then its differential df_p maps T_pM into $T_{f(p)}N$ for each $p \in M$. The collection of these maps, for all $p \in M$, is a map from TM to TN, which we denote by df.

In the special case where $f\colon U \to \mathbb{R}^n$ maps an open set $U \subset \mathbb{R}^m$ smoothly into $\mathbb{R}^n$, the differential $df\colon U \times \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^n$ is the map given by $df(p, v) = (f(p), Df(p)v)$. Observe that this is a smooth map.

Theorem 6.3. *Let M be an abstract manifold. The collection consisting of the maps*

$$d\sigma\colon U \times \mathbb{R}^m \to TM,$$

for all charts σ in an atlas on M, is an atlas for a structure of an abstract manifold on the tangent bundle TM.

With this structure the projection $\pi\colon TM \to M$ is smooth, and a vector field Y on M is smooth if and only if it is smooth as a map from M to TM.

Proof. Before we can prove that the chosen collection is an atlas on TM, we need to give TM a Hausdorff topology. We declare a subset W of TM to be open if and only if, for each σ in the atlas of M, the preimage $d\sigma^{-1}(W)$ is open in $U \times \mathbb{R}^m$. The conditions for a topology are easily verified. Moreover, with this definition it is clear that each map $d\sigma$ is a homeomorphism onto its image, and it is easily seen that $\pi\colon TM \to M$ is continuous.

We need to verify the Hausdorff axiom. Let two distinct elements $X_1, X_2 \in TM$ be given, and let $p_1, p_2 \in M$ be their base points. If p_1 and p_2 are distinct, there exist disjoint open sets $V_1, V_2 \subset M$ around them (since M is Hausdorff), and then the sets $\pi^{-1}(V_1), \pi^{-1}(V_2)$ are open, disjoint neighborhoods of X_1, X_2. On the other hand, if $p_1 = p_2$ there exists a chart σ around this point, and there exist distinct elements $w_1, w_2 \in \mathbb{R}^m$ such that $X_1 = d\sigma_p(w_1)$ and $X_2 = d\sigma_p(w_2)$. Then (since $\mathbb{R}^m$ is Hausdorff) there exist disjoint open sets W_1, W_2 in $\mathbb{R}^m$ around w_1, w_2, and the sets $d\sigma(U \times W_1), d\sigma(U \times W_2)$ are open disjoint neighborhoods of X_1 and X_2.

We have to show that our charts on TM overlap smoothly, that is, we have to verify that if $\sigma_1\colon U_1 \to M$ and $\sigma_2\colon U_2 \to M$ are charts on M, then $d\sigma_1^{-1} \circ d\sigma_2$

is smooth. It follows from the chain rule that $d\sigma_1^{-1} \circ d\sigma_2 = d(\sigma_1^{-1} \circ \sigma_2)$, and thus the assertion follows from the smoothness of $\sigma_1^{-1} \circ \sigma_2$ (see the observation before the theorem). This completes the verification that TM is an abstract manifold. It is easily seen that π is smooth for this structure.

Let Y be a vector field on M, viewed as a map $M \to TM$. Let $p \in M$ be given, and let σ be a chart around it. Then $d\sigma$ is a chart on TM around $Y(p)$, and hence the map $Y: M \to TM$ is smooth at $p = \sigma(u)$ if and only if the coordinate expression $d\sigma^{-1} \circ Y \circ \sigma$ is smooth at u. As remarked earlier (see (6.4)), this is exactly the condition in Definition 6.1.2. $\square$

Let $f: M \to N$ be a smooth map between manifolds. It follows from the chain rule that if σ is a chart on M and τ a chart on N, then $d\tau^{-1} \circ df \circ d\sigma = d(\tau^{-1} \circ f \circ \sigma)$. Using the atlas on TM as above, and the corresponding one for TN, we now see that df is a smooth map between TM and TN.

6.4 The Lie bracket

Let $X, Y \in \mathfrak{X}(M)$ be smooth vector fields on M, and let $\Omega \subset M$ be open. For a given function $f \in C^\infty(\Omega)$ we obtain a new function $Yf \in C^\infty(\Omega)$ by application of Y (see Lemma 6.2.2). Applying X to this function we obtain a function $X(Yf) \in C^\infty(\Omega)$, which we denote XYf. The two vector fields thus give rise to a linear operator $XY: C^\infty(\Omega) \to C^\infty(\Omega)$.

For example, the mixed second derivative $\frac{\partial^2 f}{\partial x \partial y}$ of a function $f \in C^\infty(\mathbb{R}^2)$ is obtained in this fashion with $X = \frac{\partial}{\partial x}$ and $Y = \frac{\partial}{\partial y}$. In this example it is known that the two derivatives commute, that is, $XYf = YXf$ for all $f \in C^\infty(\mathbb{R}^2)$, and hence the operators XY and YX agree. The same commutation does *not* occur in general. The non-commutativity is expressed by the difference $XY - YX$ of the two operators, which turns out to be a fundamental object.

Definition 6.4. Let $X, Y \in \mathfrak{X}(M)$ be smooth vector fields on M. The *Lie bracket* $[X, Y]$ is the linear operator $C^\infty(\Omega) \to C^\infty(\Omega)$ defined by the equation

$$[X, Y]f = XYf - YXf$$

for $f \in C^\infty(\Omega)$.

Example 6.4 Let X be the vector field $\frac{d}{dx}$ on $\mathbb{R}$ and let Y be the vector field $x\frac{d}{dx}$ on $\mathbb{R}$. Then $XYf = \frac{d}{dx}(x\frac{d}{dx}f) = x\frac{d^2}{dx^2}f + \frac{d}{dx}f$ and $YXf = x\frac{d}{dx}(\frac{d}{dx}f) = x\frac{d^2}{dx^2}f$, and hence the Lie bracket is the first order operator given by $[X, Y]f = XYf - YXf = \frac{d}{dx}f$.

None of the operators XY and YX are vector fields – they are second order operators, whereas vector fields are first order operators. However, it turns out that their difference $[X, Y]$ is again a vector field, the essential reason being that the second order terms cancel with each other. This is established in the following theorem.

Theorem 6.4. *Let $X, Y \in \mathfrak{X}(M)$ be smooth vector fields on an abstract manifold. The Lie bracket $[X, Y]$ is again a smooth vector field on M, that is, there exist a unique element $Z \in \mathfrak{X}(M)$ such that $[X, Y]f = Zf$ for all $f \in C^\infty(\Omega)$ and all open sets $\Omega \subset M$.*

Proof. Let $p \in M$. We will show that there exists a tangent vector $Z(p) \in T_pM$ such that $Z(p)f = ([X, Y]f)(p)$ for all $f \in C^\infty(\Omega)$ where $p \in \Omega$ (as remarked below Lemma 6.2.1 such a tangent vector, if it exists, is unique).

Choose a chart $\sigma\colon U \to M$ around p, and let

$$X(\sigma(u)) = \sum_{i=1}^m a_i(u) d\sigma_u(e_i), \quad Y(\sigma(u)) = \sum_{j=1}^m b_j(u) d\sigma_u(e_j)$$

be the expressions for the smooth vector fields X and Y, as in Definition 6.1.2. Since $d\sigma_u(e_i)f = \frac{\partial}{\partial u_i}(f \circ \sigma)(u)$, we obtain

$$\begin{aligned}
([X, Y]f) \circ \sigma &= \sum_{i,j=1}^m a_i \frac{\partial}{\partial u_i}\big(b_j \frac{\partial}{\partial u_j}(f \circ \sigma)\big) - b_j \frac{\partial}{\partial u_j}\big(a_i \frac{\partial}{\partial u_i}(f \circ \sigma)\big) \\
&= \sum_{i,j=1}^m a_i \frac{\partial b_j}{\partial u_i} \frac{\partial}{\partial u_j}(f \circ \sigma) - b_j \frac{\partial a_i}{\partial u_j} \frac{\partial}{\partial u_i}(f \circ \sigma),
\end{aligned}$$

since the terms with both differentiations on f cancel. Let

$$\begin{aligned}
Z(p) &= \sum_{i,j=1}^m a_i \frac{\partial b_j}{\partial u_i} d\sigma_u(e_j) - b_j \frac{\partial a_i}{\partial u_j} d\sigma_u(e_i) \\
&= \sum_i \big(\sum_j a_j \frac{\partial b_i}{\partial u_j} - b_j \frac{\partial a_i}{\partial u_j}\big) d\sigma_u(e_i) \in T_pM,
\end{aligned}$$

where the coefficient functions are evaluated at $u \in U$ with $\sigma(u) = p$. This tangent vector $Z(p) \in T_pM$ has the desired property.

Since p was arbitrary, we have obtained a vector field Z on M, such that $Z(p)f = [X, Y]f(p)$ for all $f \in C^\infty(\Omega)$. It follows from Lemma 6.2.2 that this vector field is smooth, since the Lie bracket $[X, Y]$ has the properties (ii) and (iii) in that lemma. $\square$

6.5 Properties of the Lie bracket

Because of the result in Theorem 6.4, we regard the Lie bracket as a map $\mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$. It has the following elementary properties, which are easily verified.

Lemma 6.5.1. *The Lie bracket satisfies*

$$[aX + bY, Z] = a[X, Z] + b[Y, Z], \quad [X, aY + bZ] = a[X, Y] + b[X, Z],$$
$$[X, Y] = -[Y, X],$$
$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0,$$

for all $X, Y, Z \in \mathfrak{X}(M)$*, and* $a, b \in \mathbb{R}$*.*

The latter equation, which is a replacement for associativity, is known as the *Jacobi identity.* The following property of the Lie bracket is verified by means of Lemma 6.1.2.

Lemma 6.5.2. *Let* $X, Y \in \mathfrak{X}(M)$ *and* $f, g \in C^\infty(M)$*. Then*

$$[fX, gY] = fg[X, Y] + f(Xg)Y - g(Yf)X.$$

The following lemma shows that the Lie bracket of a pair of smooth vector fields is transformed in a natural way by a smooth map.

Lemma 6.5.3. *Let* $\phi\colon M \to N$ *be a smooth map between abstract manifolds, and let* $X, Y \in \mathfrak{X}(M)$ *and* $V, W \in \mathfrak{X}(N)$*. If*

$$d\phi \circ X = V \circ \phi \quad \textit{and} \quad d\phi \circ Y = W \circ \phi,$$

then

$$d\phi \circ [X, Y] = [V, W] \circ \phi,$$

where these are identities between maps $M \to TN$*.*

Proof. By definition $d\phi_p(X(p))(f) = X(f \circ \phi)(p)$ for $p \in M$ and $f \in C^\infty(N)$. Hence the assumption on X and V amounts to

$$X(f \circ \phi) = (Vf) \circ \phi$$

for all $f \in C^\infty(N)$. Likewise the assumption on Y and W amounts to

$$Y(f \circ \phi) = (Wf) \circ \phi,$$

and the desired conclusion for the Lie brackets amounts to

$$[X, Y](f \circ \phi) = ([V, W]f) \circ \phi,$$

for all $f \in C^\infty(N)$.

The proof is now a straightforward computation:

$$\begin{aligned} [X, Y](f \circ \phi) &= X(Y(f \circ \phi)) - Y(X(f \circ \phi)) \\ &= X((Wf) \circ \phi) - Y((Vf) \circ \phi) \\ &= (V(Wf)) \circ \phi - (W(Vf)) \circ \phi = ([V, W]f) \circ \phi. \quad \square \end{aligned}$$

Corollary 6.5. *Let $M \subset N$ be a submanifold, and let $V, W \in \mathfrak{X}(N)$. Assume that $V(p), W(p) \in T_pM$ for all $p \in M$. Then $V|_M, W|_N \in \mathfrak{X}(M)$ and*

$$[V|_M, W|_M] = [V, W]|_M.$$

Proof. Smoothness of the restrictions follows from Theorem 4.3. We can then apply Lemma 6.5.3 with ϕ equal to the inclusion map. $\square$

6.6 The Lie algebra of a Lie group

In the remainder of this chapter we describe an important application of Lie brackets, in the theory of Lie groups. The multiplication in a group can be quite difficult to handle algebraically, one reason being that it is not commutative in general. In a Lie group G we have the extra structure of a manifold, which allows the application of differential calculus. By means of this tool we shall develop an algebraic object related to G, called the Lie algebra of G. The Lie algebra is more easily handled, since it is linear in nature. In spite of the fact that it is a simpler object, the Lie algebra contains a wealth of information about the Lie group.

Let G be a Lie group and let $g \in G$. The map $x \mapsto \ell_g(X) = gx$ of G into itself is called *left translation* by g. Obviously, this is a smooth map.

Definition 6.6.1. A vector field X on G is said to be *left invariant* if

$$d(\ell_g)_x(X(x)) = X(gx)$$

for all elements g and x in G.

A shorter notation for the condition above is $d\ell_g \circ X = X \circ \ell_g$. It is easily seen that a linear combination of left invariant vector fields is left invariant.

Example 6.6.1 Let $G = \mathbb{R}^n$ with addition. The left translation by an element $y \in \mathbb{R}^n$ is the map $x \mapsto y + x$ of $\mathbb{R}^n$ to itself. The differential of this map is the identity map of $\mathbb{R}^n$ to itself, and hence the condition that a vector field $X\colon \mathbb{R}^n \to \mathbb{R}^n$ is left invariant is that $X(x) = X(y + x)$ for all $x \in \mathbb{R}^n$, or in other words, that X is a constant map. When X is viewed as a differentiation operator, the condition is that it should have *constant coefficients*, that is $Y = \sum_{i=1} a_i \frac{\partial}{\partial x_i}$ where $a_i \in \mathbb{R}$ is constant.

Lemma 6.6. *If X, Y are left invariant smooth vector fields on G, then their Lie bracket $[X, Y]$ is again left invariant.*

Proof. Lemma 6.5.3 is used with $\phi = \ell_g$ and $V = X$, $W = Y$. $\square$

Example 6.6.2 Again, let $G = \mathbb{R}^n$ with addition. The Lie bracket of two vector fields with constant coefficients,

$$[X, Y] = [\sum_i a_i \frac{\partial}{\partial x_i}, \sum_j b_j \frac{\partial}{\partial x_j}] = 0$$

is zero because the partial derivatives commute.

Definition 6.6.2. A *Lie algebra* is a vector space $\mathfrak{g}$ over $\mathbb{R}$, equipped with a map $[\cdot,\cdot]\colon \mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$ satisfying the properties

$$[aX+bY,Z]=a[X,Z]+b[Y,Z],\quad [X,aY+bZ]=a[X,Y]+b[X,Z],$$
$$[X,Y]=-[Y,X],$$
$$[[X,Y],Z]+[[Y,Z],X]+[[Z,X],Y]=0,$$

for all $X,Y,Z\in\mathfrak{g}$ and $a,b\in\mathbb{R}$.

Example 6.6.3 Let $\mathfrak{gl}(n,\mathbb{R})$ denote the vector space $\mathrm{M}(n,\mathbb{R})$ of all real $n\times n$ matrices, equipped with the *commutator bracket* $[A,B]=AB-BA$, where AB and BA are determined by ordinary matrix multiplication. Then $\mathfrak{gl}(n,\mathbb{R})$ is a Lie algebra (the verification is by straightforward computations).

Corollary 6.6. *The space $\mathfrak{g}$ of left invariant smooth vector fields on G, equipped with the Lie bracket, is a Lie algebra.*

The set $\mathfrak{g}$ defined in this corollary is called the *Lie algebra of G*. The choice, that the elements in the Lie algebra of G are *left* invariant vector fields, is not canonical. It can be shown that the use of right translation instead of left translation would lead to an isomorphic Lie algebra (but in general, it would be realized by a different set of vector fields).

Example 6.6.4 Let $G=\mathbb{R}^n$ with addition. It follows from Example 6.6.2 that $\mathfrak{g}=\mathbb{R}^n$, and that the Lie bracket is trivial $[X,Y]=0$ for all $X,Y\in\mathfrak{g}$.

In the preceding example the Lie group G is commutative, and all the Lie brackets $[X,Y]$ are trivial. This is not a coincidence, in fact it can be shown that a connected Lie group is commutative if and only if its Lie algebra has trivial brackets (we shall not prove this here). This is a simple instance of the fact mentioned above, that a lot of information about the Lie group can be retrieved from its Lie algebra.

6.7 The tangent space at the identity

Let G be a Lie group with Lie algebra $\mathfrak{g}$. We shall give a simple relation between the two objects. First, we need a lemma.

Lemma 6.7.1. *Every left invariant vector field is smooth.*

Proof. We apply Lemma 6.7.2 below. Assume X is a left invariant vector field on G, then according to Lemma 6.2.2 we must prove that Xf is smooth for all $f\in C^\infty(G)$. Since X is left invariant

$$Xf(g)=X(g)(f)=d\ell_g(X(e))(f)=X(e)(f\circ\ell_g).$$

The map $F(x,y)=f(xy)$ is smooth on $G\times G$, and the lemma shows that $g\mapsto X(e)(F(g,\cdot))$ is smooth on G. This is exactly the desired conclusion. $\square$

Lemma 6.7.2. *Let M, N be abstract manifolds, and let $F \in C^\infty(M \times N)$. For each $x \in M$ denote by $F(x, \cdot)$ the function $y \mapsto F(x, y)$ on N. Let $q \in N$ and $Y \in T_qN$ be fixed, then $x \mapsto Y(F(x, \cdot))$ is a smooth function on M.*

Proof. Let $\sigma: U \to M$ be a chart, and let $\gamma: I \to N$ be a parametrized curve through q, which represents Y. The map $(u, t) \mapsto F(\sigma(u), \gamma(t))$ belongs to $C^\infty(U \times I)$, hence so does its derivative with respect to t. As a function of u, this derivative is $Y(F(\sigma(u), \cdot))$. □

Theorem 6.7. *Let G be a Lie group. The evaluation map $X \mapsto X(e)$ is a linear isomorphism of $\mathfrak{g}$ of G onto the tangent space T_eG.*

Proof. The map is clearly linear, and it is injective, for if $X(e) = 0$ then $X(g) = d\ell_g(X(e)) = 0$ for all $g \in G$, hence $X = 0$. Finally, we show that it is surjective. Let $Y \in T_eG$ be given, and define $X(x) = d\ell_x(Y)$ for all $x \in G$. Then X is a left invariant vector field on G, since by the chain rule

$$d(\ell_g)_x(X(x)) = d(\ell_g)_x(d(\ell_x)_e(Y)) = d(\ell_g \circ \ell_x)_e(Y) = d(\ell_{gx})_e(Y) = X(gx),$$

where we used the multiplicative property $\ell_{gx} = \ell_g \circ \ell_x$ of left multiplication. It now follows from Lemma 6.7.1 that X is smooth. □

Corollary 6.7. *The Lie algebra $\mathfrak{g}$ is finite dimensional, with dimension equal to that of the manifold G.*

Notice that Theorem 6.7 does not reveal any information about the Lie bracket on $\mathfrak{g}$. It only describes $\mathfrak{g}$ and its structure as a vector space in terms of the manifold structure of G.

6.8 The Lie algebra of $\mathrm{GL}(n, \mathbb{R})$

In this section we determine the Lie algebra of the Lie group $G = \mathrm{GL}(n, \mathbb{R})$ (see Example 2.8.4). Recall that it is an n^2-dimensional Lie group with the manifold structure as an open subset of $\mathrm{M}(n, \mathbb{R}) \simeq \mathbb{R}^{n^2}$. The tangent space $T_g \mathrm{M}(n, \mathbb{R})$ of $\mathrm{M}(n, \mathbb{R})$ at a matrix x is conveniently also identified with $\mathrm{M}(n, \mathbb{R})$, such that the tangent vector in $T_x \mathrm{GL}(n, \mathbb{R})$ which corresponds to a given matrix $A \in \mathrm{M}(n, \mathbb{R})$ is represented by the curve $t \mapsto x + tA$ in $\mathrm{GL}(n, \mathbb{R}) \subset \mathrm{M}(n, \mathbb{R})$. Accordingly, the linear isomorphism $\mathfrak{g} \to T_eG$ in Theorem 6.7 is thus in this case a linear isomorphism $\mathfrak{g} \to \mathrm{M}(n, \mathbb{R})$. In order to describe the Lie algebra of $\mathrm{GL}(n, \mathbb{R})$, we need to determine the Lie bracket, that is, we need to describe the isomorphic image in $\mathrm{M}(n, \mathbb{R})$ of $[X, Y]$ in terms of the images of X and Y.

Recall from Example 6.6.3 that we equipped the vector space $\mathrm{M}(n, \mathbb{R})$, with the commutator bracket $AB - BA$. We claim that it is *this* structure on $\mathrm{M}(n, \mathbb{R})$, which represents the Lie bracket of $\mathrm{GL}(n, \mathbb{R})$, that is, if X and Y are left invariant vector fields on $\mathrm{GL}(n, \mathbb{R})$, then $[X, Y](e) = X(e)Y(e) - Y(e)X(e)$. In other words, we claim the following.

Theorem 6.8. *The Lie algebra of* $\mathrm{GL}(n,\mathbb{R})$ *is* $\mathfrak{gl}(n,\mathbb{R})$*.*

Proof. Let X, Y be left invariant vector fields with $X(e) = A$ and $Y(e) = B$, and let Z denote the left invariant vector field $[X, Y]$. The claim is that

$$Zf(e) = (AB - BA)f \tag{6.9}$$

for all $f \in C^\infty(\mathrm{GL}(n,\mathbb{R}))$. Let $\xi\colon \mathrm{M}(n,\mathbb{R}) \to \mathbb{R}$ be one of the n^2 maps, which takes a matrix to a given entry. These are the coordinate functions for our chart on $\mathrm{GL}(n,\mathbb{R})$. In Lemma 6.2.1 it was seen that a tangent vector is uniquely determined by its action on the coordinate functions, hence it follows that it suffices to prove (6.9) for $f = \xi$.

We shall determine the tangent vectors $X(g)$ and $Y(g)$ in terms of $X(e) = A$ and $Y(e) = B$. For a given $g \in \mathrm{GL}(n,\mathbb{R})$, the map $\ell_g\colon x \to gx$ of $\mathrm{GL}(n,\mathbb{R})$ to itself transforms the curve $t \mapsto x + tA$, representing A as a tangent vector, into the curve $t \mapsto gx + tgA$, of which the derivative at $t = 0$ is gA. Hence the differential of $d\ell_g$ at x is $d(\ell_g)_x(A) = gA$. The condition that a vector field X on $\mathrm{GL}(n,\mathbb{R})$ is left invariant, is thus that $X(gx) = gX(x)$ for all $g, x \in G$. It follows that $X(g) = gA$, and similarly $Y(g) = gB$.

It now follows that the action of the invariant vector field X on a function $f \in C^\infty(G)$ is given by

$$Xf(g) = \frac{d}{dt} f(g + tgA)|_{t=0}. \tag{6.10}$$

We insert $f = \xi$, and observe that ξ is a *linear* function of the matrices. Hence, for $X(g) = gA$,

$$X\xi(g) = \frac{d}{dt}\xi(g) + t\xi(gA)|_{t=0} = \xi(gA). \tag{6.11}$$

Replacing X by Y we thus have $Y\xi(g) = \xi(gB)$. We now insert $f = Y\xi$ and $g = e$ in (6.10), and we obtain

$$XY\xi(e) = \frac{d}{dt} Y\xi(e + tA)|_{t=0} = \frac{d}{dt}\xi((e + tA)B)|_{t=0} = \xi(AB).$$

Likewise $YX\xi(e) = \xi(BA)$, and thus $Z\xi(e) = \xi(AB - BA)$.

On the other hand, if we apply (6.11) with $g = e$ and with X replaced by the tangent vector $AB - BA$ at e, we obtain $(AB - BA)\xi = \xi(AB - BA)$. Now (6.9) follows. $\square$

6.9 Homomorphisms of Lie groups

Definition 6.9. Let G, H be Lie groups. A map $\phi: G \to H$ which is smooth and a homomorphism in the group theoretical sense, that is, which satisfies

$$\phi(xy) = \phi(x)\phi(y) \quad \forall x, y \in G,$$

is called a *homomorphism of Lie groups.*

For example, the determinant is a Lie homomorphism from $\mathrm{GL}(n, \mathbb{R})$ to $\mathbb{R}^\times$. The natural map from $\mathrm{SO}(2)$ to $\mathrm{GL}(2, \mathbb{R})$ is another example.

Let $\phi: G \to H$ be a homomorphism of Lie groups. The differential $d\phi_e$ at e maps the tangent space T_eG into T_eH. Hence it follows from Theorem 6.7 (applied both to G and to H), that for each element X in the Lie algebra $\mathfrak{g}$, there exists a unique element Y in the Lie algebra $\mathfrak{h}$, such that $Y(e) = d\phi_e(X(e))$. It follows from the homomorphism property of ϕ that then $Y(\phi(g)) = d\phi_g(X(g))$ for all $g \in G$. The map $X \mapsto Y$ from $\mathfrak{g}$ to $\mathfrak{h}$ is denoted by $d\phi$, and said to be *induced* by ϕ.

Theorem 6.9. *Let $\phi: G \to H$ be a homomorphism of Lie groups. The induced map $d\phi$ is a homomorphism of Lie algebras, that is, it is linear and satisfies*

$$d\phi([X_1, X_2]) = [d\phi(X_1), d\phi(X_2)]$$

for all $X_1, X_2 \in \mathfrak{g}$.

Proof. Follows from Lemma 6.5.3. □

Example 6.9 Let $G = \mathrm{O}(n)$. In Section 4.5 we determined the tangent space $T_I \, \mathrm{O}(n)$ at the identity as the set of $n \times n$ antisymmetric matrices. We denote this set by $\mathfrak{o}(n)$,

$$\mathfrak{o}(n) = \{A \in \mathrm{M}(n, \mathbb{R}) \mid A^t = -A\}.$$

The inclusion map $\mathrm{O}(n) \to \mathrm{GL}(n, \mathbb{R})$ is a homomorphism, hence its differential, which is the inclusion map of $\mathfrak{o}(n)$ in $\mathfrak{gl}(n, \mathbb{R})$, is a homomorphism of Lie algebras according to Theorem 6.9. It now follows from Theorem 6.8 that the Lie bracket of $\mathfrak{o}(n)$ again is the commutator bracket $[A, B] = AB - BA$.

Notice that $\mathrm{O}(n)$ is compact, since it is a closed and bounded subset of $\mathrm{M}(n, \mathbb{R}) \simeq \mathbb{R}^{n^2}$. Indeed, by definition all the columns of an orthogonal matrix have norm 1, hence all entries are bounded by 1, and the relation $A^tA = I$ is continuous in the entries of A, hence it defines closed set.

Chapter 7

Tensors

In this chapter some fundamental constructions for a real vector space V are introduced: The dual space V^*, the tensor spaces $T^k(V)$ and the alternating tensor spaces $A^k(V)$. The presentation is based purely on linear algebra, and it is independent of all the preceding chapters. The bridge to geometry will be built in the following chapter, where we shall apply the theory of the present chapter to the study of manifolds. The linear space V will then be the tangent space T_pM at a given point.

Let V be a vector space over $\mathbb{R}$. For our purposes only finite dimensional spaces are needed, so we shall assume $\dim V < \infty$ whenever it is convenient.

7.1 The dual space

We recall that the dual space of a vector space is defined as follows.

Definition 7.1. The *dual space* V^* is the space of linear maps $\xi\colon V \to \mathbb{R}$.

A linear map $\xi \in V^*$ is often called a *linear form*. Equipped with the natural algebraic operations of addition and scalar multiplication of maps, V^* becomes a vector space on its own.

The basic theorem about V^*, for V finite dimensional, is the following.

Theorem 7.1.1. *Assume* $\dim V = n \in \mathbb{N}$, *and let* $e_1, \dots, e_n$ *be a basis.*

(i) For each $i = 1, \dots, n$, *an element* $\xi_i \in V^*$ *is defined by*

$$\xi_i(a_1e_1 + \cdots + a_ne_n) = a_i, \qquad a_1, \dots, a_n \in \mathbb{R}.$$

(ii) The elements $\xi_1, \dots, \xi_n$ *form a basis for* V^* *(called the* dual basis*).*

Proof. (i) is easy. For (ii), notice first that two linear forms on a vector space are equal, if they agree on each element of a basis. Notice also that it follows from the definition of ξ_i that $\xi_i(e_j) = \delta_{ij}$. Let $\xi \in V^*$, then

$$\xi = \sum_{i=1}^{n} \xi(e_i)\xi_i, \tag{7.1}$$

because the two sides agree on each e_j. This shows that the vectors $\xi_1, \dots, \xi_n$ span V^*. They are also linearly independent, for if $\sum_i b_i\xi_i = 0$ then $b_j = \sum_i b_i\xi_i(e_j) = 0$ for all j. $\square$

Corollary 7.1. *If* $\dim V = n$ *then* $\dim V^* = n$.

If a linear form $\xi \in V^*$ satisfies that $\xi(v) = 0$ for all $v \in V$, then by definition $\xi = 0$. The similar result for elements $v \in V$ needs a proof:

Lemma 7.1.1. *Let* $v \in V$. *If* $v \neq 0$ *then* $\xi(v) \neq 0$ *for some* $\xi \in V^*$.

Proof. For simplicity we assume $\dim V < \infty$ (although the result is true in general). Assume $v \neq 0$. Then there exists a basis $e_1, \dots, e_n$ for V with $e_1 = v$. The element ξ_1 of the dual basis satisfies $\xi_1(v) = 1$. □

The dual space is useful for example in the study of subspaces of V. This can be seen from the following theorem, which shows that the elements of V^* can be used to detect whether a given vector belongs to a given subspace.

Lemma 7.1.2. *Let* $U \subset V$ *be a linear subspace, and let* $v \in V$. *If* $v \notin U$ *then* $\xi(v) \neq 0$ *for some* $\xi \in V^*$ *with* $\xi|_U = 0$.

Proof. The proof is similar to the previous one, which corresponds to the special case $U = \{0\}$. As before we assume $\dim V < \infty$. Let $e_1, \dots, e_m$ be a basis for U, let $e_{m+1} = v$, and extend to a basis $e_1, \dots, e_n$ for V. The element ξ_{m+1} of the dual basis satisfies $\xi_{m+1}|_U = 0$ and $\xi_{m+1}(v) = 1$. □

The term 'dual' suggests some kind of symmetry between V and V^*. The following theorem indicates such a symmetry for the finite dimensional case. We define $V^{**} = (V^*)^*$ as the dual of the dual space.

Theorem 7.1.2(duality theorem)**.** *The map* $\Phi : V \to V^{**}$ *defined by*

$$\Phi(v)(\xi) = \xi(v)$$

for $v \in V$, $\xi \in V^*$ *is linear and injective. If* $\dim V < \infty$ *it is an isomorphism.*

Notice that Φ is defined without reference to a particular basis for V.

Proof. It is easily seen that Φ maps into V^{**}, and that it is linear. If $\Phi(v) = 0$ then $\xi(v) = 0$ for all ξ, hence $v = 0$ by Lemma 7.1.1. Thus Φ is injective. If $\dim V < \infty$ then $\dim V^{**} = \dim V^* = \dim V$, and hence Φ is also surjective. □

7.2 The dual of a linear map

Let V, W be vector spaces over $\mathbb{R}$, and let $T : V \to W$ be linear. By definition, the *dual map* $T^* : W^* \to V^*$ takes a linear form $\eta \in W^*$ to its pull-back by T, that is, $T^*(\eta) = \eta \circ T$. It is easily seen that T^* is linear.

Lemma 7.2.1. *Assume V and W are finite dimensional, and let $T: V \to W$ be linear. If T is represented by a matrix (a_{ij}) with respect to some given bases, then T^* is represented by the transposed matrix (a_{ji}) with respect to the dual bases.*

Proof. Let $e_1, \dots, e_n$ and $f_1, \dots, f_m$ denote the given bases for V and W. The fact that T is represented by a_{ij} is expressed in the equality

$$Te_j = \sum_i a_{ij} f_i.$$

Let $\xi_1, \dots, \xi_n$ and $\eta_1, \dots, \eta_m$ denote the dual bases for V^* and W^*. Then

$$a_{ij} = \eta_i(Te_j).$$

We now obtain from (7.1) with $\xi = T^*\eta_k$ that

$$T^*\eta_k = \sum_j T^*\eta_k(e_j)\xi_j = \sum_j \eta_k(Te_j)\xi_j = \sum_j a_{kj}\xi_j. \quad \square$$

Lemma 7.2.2. *(i) T is surjective if and only if T^* is injective.*
(ii) T^* is surjective if and only if T is injective.

Proof. Assume for simplicity that V and W are finite dimensional. It is known from linear algebra that a map represented by a matrix is surjective if and only if the rank of the matrix equals the number of rows, and injective if and only if the rank equals the number of columns. The lemma follows from the preceding, since the rank of (a_{ij}) is equal to the rank of its transpose. $\square$

7.3 Tensors

We now proceed to define tensors. Let $k \in \mathbb{N}$. Given a collection of vector spaces $V_1, \dots, V_k$ one can define a vector space $V_1 \otimes \cdots \otimes V_k$, called their tensor product. The elements of this vector space are called tensors. However, we do not need to work in this generality, and we shall be content with the situation where the vector spaces $V_1, \dots, V_k$ are all equal to the same space. In fact, the tensor space T^kV we define below corresponds to $V^* \otimes \cdots \otimes V^*$ in the general notation.

Definition 7.3.1. Let $V^k = V \times \cdots \times V$ be the Cartesian product of k copies of V. A map φ from V^k to a vector space U is called *multilinear* if it is linear in each variable separately (i.e. with the other variables held fixed).

Definition 7.3.2. A (covariant) *k-tensor* on V is a multilinear map $T: V^k \to \mathbb{R}$. The set of k-tensors on V is denoted $T^k(V)$.

In particular, a 1-tensor is a linear form, $T^1(V) = V^*$. It is convenient to add the convention that $T^0(V) = \mathbb{R}$. The set $T^k(V)$ is called *tensor space*, it is a vector space because sums and scalar products of multilinear maps are again multilinear.

Lemma 7.3. *Let $\eta_1, \dots, \eta_k \in V^*$. The map*

$$(v_1, \dots, v_k) \mapsto \eta_1(v_1) \cdots \eta_k(v_k) \tag{7.2}$$

is a k-tensor.

Proof. This is clear. $\square$

The map (7.2) is denoted $\eta_1 \otimes \cdots \otimes \eta_k$. More generally, if $S \in T^k(V)$ and $T \in T^l(V)$ are tensors, we define the *tensor product* $S \otimes T \in T^{k+l}(V)$ by

$$S \otimes T(v_1, \dots, v_k, v_{k+1}, \dots, v_{k+l}) = S(v_1, \dots, v_k) T(v_{k+1}, \dots, v_{k+l}).$$

It is easily seen that $S \otimes T$ is a $k+l$-tensor, and that the algebraic operation $\otimes$ between tensors is associative and satisfies the proper distributive laws of a product. The associativity is expressed through

$$(R \otimes S) \otimes T = R \otimes (S \otimes T)$$

and the content of the distributive law is that $S \otimes T$ depends linearly on S and T. Because of the associativity no brackets are needed in the tensor product of three or more tensors. This explains the notation $\eta_1 \otimes \cdots \otimes \eta_k$.

Notice that for $k = 0$ the tensor $S \in T^k(V)$ is a number. In this case (and similarly when $l = 0$) the tensor product $S \otimes T$ is defined by multiplication of T with that number.

Theorem 7.3. *Assume $\dim V = n$ with $e_1, \dots, e_n$ a basis. Let $\xi_1, \dots, \xi_n \in V^*$ denote the dual basis. The elements $\xi_{i_1} \otimes \cdots \otimes \xi_{i_k}$, where $I = (i_1, \dots, i_k)$ is an arbitrary sequence of k numbers in $\{1, \dots, n\}$, form a basis for $T^k(V)$.*

Proof. Let $T_I = \xi_{i_1} \otimes \cdots \otimes \xi_{i_k}$. Notice that if $J = (j_1, \dots, j_k)$ is another sequence of k integers, and we denote by e_J the element $(e_{j_1}, \dots, e_{j_k}) \in V^k$, then

$$T_I(e_J) = \delta_{IJ},$$

that is, $T_I(e_J) = 1$ if $J = I$ and 0 otherwise. It follows that the T_I are linearly independent, for if a linear combination $T = \sum_I a_I T_I$ is zero, then $a_J = T(e_J) = 0$.

It follows from the multilinearity that a k-tensor is uniquely determined by its values on all elements in V^k of the form e_J. For any given k-tensor T we have that the k-tensor $\sum_I T(e_I) T_I$ agrees with T on all e_J, hence $T = \sum_I T(e_I) T_I$ and we conclude that the T_I span $T^k(V)$. $\square$

Corollary 7.3. *If* $\dim V = n$ *then* $\dim T^k(V) = n^k$.

It follows from Theorem 7.3 that every element of $T^k(V)$ is a sum of tensor products of the form $\eta_1 \otimes \cdots \otimes \eta_k$, where $\eta_1, \dots, \eta_k \in V^*$. For this reason it is customary to denote $T^k(V) = V^* \otimes \cdots \otimes V^*$.

Remark (especially for physicists). The coefficients of a vector $v \in V$ with respect to a basis $v_1, \dots, v_n$ are often numbered with superscripts instead of subscripts, so that $v = \sum a^i v_i$. Furthermore, the elements of the dual basis are numbered with superscripts, $\xi^1, \dots, \xi^n$, and the coefficients of a vector $\xi \in V^*$ with respect to this basis are then numbered with subscripts so that $\xi = \sum b_i \xi^i$. Then

$$\xi(v) = \sum a^i b_i.$$

These rules are adapted to the so-called "Einstein convention" which is to sum over an index which appears both above and below.

The coefficients of a covariant tensor are numbered with subscripts, so that the expression by means of the basis in Theorem 7.3 becomes

$$T = \sum a_{i_1,\dots,i_k} \xi^{i_1} \otimes \cdots \otimes \xi^{i_k}.$$

There is a similar construction of so-called contravariant tensors, which have superscript indices, and also mixed tensors which carry both types of indices, but we do not need them here. In the physics literature, a tensor T is often identified with the array of coefficients $a^{j_1,\dots,j_l}_{i_1,\dots,i_k}$ with respect to a particular basis. The array is then required to obey a transformation rule for the passage between different bases.

7.4 Alternating tensors

Let V be a real vector space. In the preceding section the tensor spaces $T^k(V)$ were defined, together with the tensor product

$$(S, T) \mapsto S \otimes T, \quad T^k(V) \times T^l(V) \to T^{k+l}(V).$$

There is an important construction of vector spaces which resemble tensor powers of V, but for which there is a more refined structure. These are the so-called exterior powers of V, which play an important role in differential geometry because the theory of differential forms is built on them. They are also of importance in algebraic topology and many other fields.

A multilinear map

$$\varphi: V^k = V \times \cdots \times V \to U$$

where $k > 1$, is said to be *alternating* if for all $v_1, \dots, v_k \in V$ the value $\varphi(v_1, \dots, v_k)$ changes sign whenever two of the vectors $v_1, \dots, v_k$ are interchanged, that is

$$\varphi(v_1, \dots, v_i, \dots, v_j, \dots, v_k) = -\varphi(v_1, \dots, v_j, \dots, v_i, \dots, v_k). \tag{7.3}$$

Since every permutation of the numbers $1,\dots,k$ can be decomposed into transpositions, it follows that

$$\varphi(v_{\sigma(1)},\dots,v_{\sigma(k)}) = \operatorname{sgn}\sigma\,\varphi(v_1,\dots,v_k) \tag{7.4}$$

for all permutations $\sigma \in S_k$ of the numbers $1,\dots,k$.

Examples 7.4.1. 1. Let $V = \mathbb{R}^3$. The vector product $(v_1, v_2) \mapsto v_1 \times v_2 \in V$ is alternating from $V \times V$ to V.

2. Let $V = \mathbb{R}^n$. The $n \times n$ determinant is multilinear and alternating in its columns, hence it can be viewed as an alternating map $(\mathbb{R}^n)^n \to \mathbb{R}$.

Lemma 7.4.1. *Let $\varphi\colon V^k \to U$ be multilinear. The following conditions are equivalent:*

(a) *φ is alternating,*

(b) *$\varphi(v_1,\dots,v_k) = 0$ whenever two of the vectors $v_1,\dots,v_k$ coincide,*

(c) *$\varphi(v_1,\dots,v_k) = 0$ whenever the vectors $v_1,\dots,v_k$ are linearly dependent.*

Proof. (a)$\Rightarrow$(b) If $v_i = v_j$ the interchange of v_i and v_j does not change the value of $\varphi(v_1,\dots,v_k)$, so (7.3) implies $\varphi(v_1,\dots,v_k) = 0$.

(b)$\Rightarrow$(a) Consider for example the interchange of v_1 and v_2. By linearity

$$\begin{aligned}
0 &= \varphi(v_1+v_2, v_1+v_2,\dots) \\
&= \varphi(v_1,v_1,\dots)+\varphi(v_1,v_2,\dots)+\varphi(v_2,v_1,\dots)+\varphi(v_2,v_2,\dots) \\
&= \varphi(v_1,v_2,\dots)+\varphi(v_2,v_1,\dots).
\end{aligned}$$

It follows that $\varphi(v_2,v_1,\dots) = -\varphi(v_1,v_2,\dots)$.

(b)$\Rightarrow$(c) If the vectors $v_1,\dots,v_k$ are linearly dependent then one of them can be written as a linear combination of the others. It then follows that $\varphi(v_1,\dots,v_k)$ is a linear combination of terms in each of which some v_i appears twice.

(c)$\Rightarrow$(b) Obvious. $\square$

In particular, if $k > \dim V$ then every set of k vectors is linearly dependent, and hence $\varphi = 0$ is the only alternating map $V^k \to U$.

Definition 7.4. An *alternating k-form* is an alternating k-tensor $V^k \to \mathbb{R}$. The space of these is denoted $A^k(V)$, it is a linear subspace of $T^k(V)$.

We define $A^1(V) = V^*$ and $A^0(V) = \mathbb{R}$.

Example 7.4.2. Let $\eta, \zeta \in V^*$. The 2-tensor $\eta\otimes\zeta - \zeta\otimes\eta$ is alternating.

The following lemma exhibits a standard procedure to construct alternating forms.

Lemma 7.4.2. *Let $k > 1$. For each $T \in T^k(V)$ the element* $\operatorname{Alt}(T) \in T^k(V)$ *defined by*

$$\operatorname{Alt}(T)(v_1, \dots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sgn} \sigma \, T(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \tag{7.5}$$

is alternating. Moreover, if T is already alternating, then $\operatorname{Alt}(T) = T$.

Proof. Let $\tau \in S_k$ be the transposition corresponding to an interchange of two vectors among $v_1, \dots, v_k$. We have

$$\operatorname{Alt}(T)(v_{\tau(1)}, \dots, v_{\tau(k)}) = \frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sgn} \sigma \, T(v_{\tau\circ\sigma(1)}, \dots, v_{\tau\circ\sigma(k)}).$$

Since $\sigma \mapsto \tau\circ\sigma$ is a bijection of S_k we can substitute σ for $\tau\circ\sigma$. Using $\operatorname{sgn}(\tau\circ\sigma) = -\operatorname{sgn}(\sigma)$, we obtain the desired equality with $-\operatorname{Alt}(T)(v_1, \dots, v_k)$.

If T is already alternating, all summands of (7.5) are equal to $T(v_1, \dots, v_k)$. Since $|S_k| = k!$ we conclude that $\operatorname{Alt}(T) = T$. ☐

The definition in Lemma 7.4.2 for $k > 1$ is supplemented with the following convention

$$\operatorname{Alt}(T) = T, \qquad T \in A^0(V) = \mathbb{R} \text{ or } T \in A^1(V) = V^*.$$

Notice that for $\eta_1, \dots, \eta_k \in V^*$ and $v_1, \dots, v_k \in V$ we obtain

$$\operatorname{Alt}(\eta_{i_1} \otimes \cdots \otimes \eta_{i_k})(v_1, \dots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sgn} \sigma \; \eta_{i_1}(v_{\sigma(1)}) \cdots \eta_{i_k}(v_{\sigma(k)})$$

and hence

$$\operatorname{Alt}(\eta_{i_1} \otimes \cdots \otimes \eta_{i_k})(v_1, \dots, v_k) = \frac{1}{k!} \det[(\eta_i(v_j))_{ij}]. \tag{7.6}$$

Let $e_1, \dots, e_n$ be a basis for V, and $\xi_1, \dots, \xi_n$ the dual basis for V^*. We saw in Theorem 7.3 that the elements $\xi_{i_1} \otimes \cdots \otimes \xi_{i_k}$ form a basis for $T^k(V)$. We will now exhibit a similar basis for $A^k(V)$. We have seen already that $A^k(V) = 0$ if $k > n$.

Theorem 7.4. *Assume $k \le n$. For each subset $I \subset \{1, \dots, n\}$ with k elements, let $1 \le i_1 < \cdots < i_k \le n$ be its elements, and let*

$$\xi_I = \operatorname{Alt}(\xi_{i_1} \otimes \cdots \otimes \xi_{i_k}) \in A^k(V). \tag{7.7}$$

These elements ξ_I form a basis for $A^k(V)$. In particular,

$$\dim A^k(V) = \frac{n!}{k!(n-k)!}.$$

Proof. It follows from the last statement in Lemma 7.4.2 that Alt: $T^k(V) \to A^k(V)$ is surjective. Applying Alt to all the basis elements $\xi_{i_1} \otimes \cdots \otimes \xi_{i_k}$ for $T^k(V)$, we therefore obtain a spanning set for $A^k(V)$. It follows from (7.6) that $\mathrm{Alt}(\xi_{i_1} \otimes \cdots \otimes \xi_{i_k}) = 0$ if there are repetitions among the $i_1, \dots, i_k$. Moreover, if we rearrange the order of the numbers $i_1, \dots, i_k$, the element $\mathrm{Alt}(\xi_{i_1} \otimes \cdots \otimes \xi_{i_k})$ is unchanged, apart from a possible change of the sign. Therefore $A^k(V)$ is spanned by the elements ξ_I in (7.7).

Consider an arbitrary linear combination $T = \sum_I a_I \xi_I$ with coefficients $a_I \in \mathbb{R}$. Let $J = (j_1, \dots, j_k)$ where $1 \leq j_1 < \cdots < j_k \leq n$, then it follows from (7.6) that

$$\xi_I(e_{j_1}, \dots, e_{j_k}) = \begin{cases} 1/k! & \text{if } I = J \\ 0 & \text{otherwise} \end{cases}$$

It follows that $T(e_{j_1}, \dots, e_{j_k}) = a_J/k!$ for $J = (j_1, \dots, j_k)$. Therefore, if $T = 0$ we conclude $a_J = 0$ for all the coefficients. Thus the elements ξ_I are independent. $\square$

Notice the special case $k = n$ in the preceding theorem. It follows that $A^n(V)$ is one-dimensional, and spanned by the map

$$(v_1, \dots, v_n) \mapsto \det[(\xi_i(v_j))_{ij}]$$

for an arbitrary (dual) basis for V^* (thus demonstrating the uniqueness of the determinant as an alternating n-multilinear map).

7.5 The wedge product

In analogy with the tensor product $(S, T) \mapsto S \otimes T$, from $T^k(V) \times T^l(V)$ to $T^{k+l}(V)$, there is a construction of a product $A^k(V) \times A^l(V) \to A^{k+l}$. Since tensor products of alternating tensors are not alternating, it does not suffice just to take $S \otimes T$.

Definition 7.5. Let $S \in A^k(V)$ and $T \in A^l(V)$. The *wedge product* $S \wedge T \in A^{k+l}(V)$ is defined by

$$S \wedge T = \mathrm{Alt}(S \otimes T).$$

Notice that in the case $k = 0$, where $A^k(V) = \mathbb{R}$, the wedge product is just scalar multiplication.

Example 7.5 Let $\eta_1, \eta_2 \in A^1(V) = V^*$. Then by definition

$$\eta_1 \wedge \eta_2 = \frac{1}{2}(\eta_1 \otimes \eta_2 - \eta_2 \otimes \eta_1).$$

Since the operator Alt is linear, the wedge product depends linearly on the factors S and T. It is more cumbersome to verify the associative rule for $\wedge$. In order to do this we need the following lemma.

Lemma 7.5.1. *Let $S \in T^k(V)$ and $T \in T^l(V)$. Then*

$$\mathrm{Alt}(\mathrm{Alt}(S) \otimes T) = \mathrm{Alt}(S \otimes \mathrm{Alt}(T)) = \mathrm{Alt}(S \otimes T).$$

Proof. We will only verify

$$\mathrm{Alt}(\mathrm{Alt}(S) \otimes T) = \mathrm{Alt}(S \otimes T).$$

The proof for the other expression is similar.

Let $G = S_{k+l}$ and let $H \subset G$ denote the subgroup of permutations leaving each of the last elements $k+1, \dots, k+l$ fixed. Then H is naturally isomorphic to S_k. Now

$$\begin{aligned}
&\mathrm{Alt}(\mathrm{Alt}(S) \otimes T)(v_1, \dots, v_{k+l}) \\
&= \tfrac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn} \sigma \, \Big(\mathrm{Alt}(S)(v_{\sigma(1)}, \dots, v_{\sigma(k)})\Big) T(v_{\sigma(k+1)}, \dots, v_{\sigma(k+l)}) \\
&= \tfrac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn} \sigma \, \Big(\tfrac{1}{k!} \sum_{\tau \in S_k} \operatorname{sgn} \tau \; S(v_{\sigma(\tau(1))}, \dots, v_{\sigma(\tau(k))})\Big) \\
&\qquad\qquad\qquad\qquad\qquad\qquad T(v_{\sigma(k+1)}, \dots, v_{\sigma(k+l)}) \\
&= \tfrac{1}{(k+l)!k!} \sum_{\sigma \in G} \sum_{\tau \in H} \operatorname{sgn}(\sigma \circ \tau) \; S(v_{\sigma(\tau(1))}, \dots, v_{\sigma(\tau(k))}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad T(v_{\sigma(\tau(k+1))}, \dots, v_{\sigma(\tau(k+l))}).
\end{aligned}$$

Since $\sigma \mapsto \sigma \circ \tau$ is a bijection of G we can substitute σ for $\sigma \circ \tau$, and we obtain the desired expression, since there are $k!$ elements in H. $\square$

Lemma 7.5.2. *Let $R \in A^k(V)$, $S \in A^l(V)$ and $T \in A^m(V)$. Then*

$$(R \wedge S) \wedge T = R \wedge (S \wedge T) = \mathrm{Alt}(R \otimes S \otimes T).$$

Proof. It follows from the preceding lemma that

$$(R \wedge S) \wedge T = \mathrm{Alt}(\mathrm{Alt}(R \otimes S) \otimes T) = \mathrm{Alt}(R \otimes S \otimes T)$$

and

$$R \wedge (S \wedge T) = \mathrm{Alt}(R \otimes \mathrm{Alt}(S \otimes T)) = \mathrm{Alt}(R \otimes S \otimes T). \quad \square$$

Since the wedge product is associative, we can write any product $T_1 \wedge \cdots \wedge T_r$ of tensors $T_i \in A^{k_i}(V)$ without specifying brackets. In fact, it follows by induction from Lemma 7.5.1 that

$$T_1 \wedge \cdots \wedge T_r = \mathrm{Alt}(T_1 \otimes \cdots \otimes T_r)$$

regardless of how brackets are inserted in the wedge product.

In particular, it follows from (7.6) that

$$\eta_1 \wedge \cdots \wedge \eta_k(v_1, \ldots, v_k) = \frac{1}{k!} \det[(\eta_i(v_j))_{ij}] \tag{7.8}$$

for all $v_1, \ldots, v_k \in V$ and $\eta_1, \ldots, \eta_k \in V^*$, where the vectors in V^* are viewed as 1-forms. The basis elements ξ_I in Theorem 7.4 are written in this fashion as

$$\xi_I = \xi_{i_1} \wedge \cdots \wedge \xi_{i_k}$$

where $I = (i_1, \ldots, i_k)$ is an increasing sequence from $1, \ldots, n$. This will be our notation for ξ_I from now on.

The wedge product is not commutative. Instead, it satisfies the following relation for the interchange of factors.

Lemma 7.5.3. *Let $\eta, \zeta \in V^*$, then*

$$\zeta \wedge \eta = -\eta \wedge \zeta. \tag{7.9}$$

More generally, if $S \in A^k(V)$ and $T \in A^l(V)$ then

$$T \wedge S = (-1)^{kl}\, S \wedge T \tag{7.10}$$

Proof. The identity (7.9) follows immediately from the fact that $\eta \wedge \zeta = \frac{1}{2}(\eta \otimes \zeta - \zeta \otimes \eta)$.

Since $A^k(V)$ is spanned by elements of the type $S = \eta_1 \wedge \cdots \wedge \eta_k$, and $A^l(V)$ by elements of the type $T = \zeta_1 \wedge \cdots \wedge \zeta_l$, where $\eta_i, \zeta_j \in V^*$, it suffices to prove (7.10) for these forms. In order to rewrite $T \wedge S$ as $S \wedge T$ we must let each of the k elements η_i pass the l elements ζ_j. The total number of sign changes is therefore kl. $\square$

7.6 The exterior algebra

Let

$$A(V) = A^0(V) \oplus A^1(V) \oplus \cdots \oplus A^n(V)$$

be the direct sum of all the spaces of alternating forms on V. Notice that the general elements in $A(V)$ are formal sums of k-forms with different values of k for the components, and it is only the individual components that are multilinear, alternating maps.

We extend the wedge product $\wedge$ to $A(V)$ by linearity, and in this fashion we obtain an algebra, called the *exterior algebra* of V. The element $1 \in A^0(V)$ is a unit for this algebra. The exterior algebra has a linear basis consisting of all the elements ξ_I, where I is an arbitrary subset of $\{1, \ldots, n\}$. Its dimension is

$$\dim A(V) = 2^n.$$

Chapter 8

Differential forms

It is customary to write the integral of a real function $f\colon \mathbb{R}^m \to \mathbb{R}$ in the form $\int f(x)\,dx_1 \dots dx_m$. The quantity $dx_1 \dots dx_n$ in the formula is regarded just as formal notation, which reminds us of the fact that in the definition of the Riemann integral (for functions of one variable), a limit is taken where the increments Δx tend to zero. Thus dx is regarded as an 'infinitesimal' version of Δx. In the theory of differential forms the infinitesimal quantity is replaced by an object $dx_1 \wedge \cdots \wedge dx_m$ which we shall see has a precise meaning as a k-form, acting on elements of the tangent space.

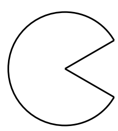
The main motivation is to develop theories of differentiation and integration which are valid for smooth manifolds without reference to any particular charts.

8.1 The cotangent space

Let M be an m-dimensional abstract manifold, and let $p \in M$.

Definition 8.1. The dual space $(T_pM)^*$ of T_pM is denoted T_p^*M and called the *cotangent space* of M at p. Its elements are called *cotangent vectors* or just *covectors.*

A tangent vector in T_pM is by definition an equivalence class $[\gamma]_p$ of smooth curves on M through p, say $p = \gamma(t_0)$. A cotangent vector associates a number in $\mathbb{R}$ with each such equivalence class. One can think of a covector as an "animal, which feeds on tangent vectors". By lack of better visualization, we draw it as follows:



a covector

Example 8.1 Consider a function $f \in C^\infty(M)$. Its differential df_p at $p \in M$ is a linear map from T_pM into the tangent space of $\mathbb{R}$, which we identify as $\mathbb{R}$. Hence

$$df_p \in T_p^*M.$$

We recall that $df_p(X) = \mathrm{D}_X f = (f \circ \gamma)'(t_0)$ for $X \in T_pM$, if γ is a parametrized curve through $\gamma(t_0) = p$, representing X.

Given a chart $\sigma: U \to M$ where $U \subset \mathbb{R}^m$ is open, we can now give dx_i a precise meaning as follows. Recall from Section 6.2 that we defined the *coordinate function* $x_i: \sigma(U) \to \mathbb{R}$ as the smooth map which takes $p = \sigma(u)$ to the i-th coordinate u_i of $u \in U$. Thus

$$\sigma^{-1}(p) = (x_1(p), \dots, x_m(p)) \in U.$$

See also Lemma 6.2.1, where the same function was denoted ξ_i. The change of notation is motivated by the desire to give dx_i the precise meaning, which it now obtains as the differential of x_i at p. In order to avoid the double subscript of $d(x_i)_p$, we denote this differential by $dx_i(p)$. According to Example 8.1 it is a covector,

$$dx_i(p) \in T_p^*M$$

for each $p \in \sigma(U)$.

Lemma 8.1. *Let $\sigma: U \to M$ be a chart on M, and let $p = \sigma(u) \in \sigma(U)$. The tangent vectors*

$$d\sigma_u(e_1), \dots, d\sigma_u(e_m) \tag{8.1}$$

form a basis for T_pM. The dual basis for T_p^*M is

$$dx_1(p), \dots, dx_m(p). \tag{8.2}$$

Thus, $dx_i(p)$ is the linear form on T_pM, which carries a vector to its i-th coordinate in the basis (8.1).

Proof. Recall from Section 3.8 that the standard basis for T_pM with respect to σ consists of the vectors (8.1), so the first statement is just a repetition of the fact that the standard basis is a basis.

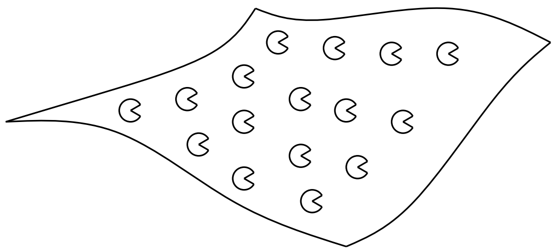
For the second statement, we notice that by the chain rule

$$dx_i(p)(d\sigma_u(e_j)) = d(x_i \circ \sigma)_u(e_j) = \frac{\partial u_i}{\partial u_j} = \delta_{ij}, \tag{8.3}$$

since $x_i \circ \sigma(u) = u_i$. $\square$

8.2 Covector fields

Recall that a vector field Y on M associates a tangent vector $Y(p) \in T_pM$ to each point $p \in M$. Similarly, a *covector field* associates a covector to each point. If we visualize a vector field as a collection of tangent arrows, one based in each point on M, then the corresponding picture of a covector field is that it is a collection of arrow-eating animals, one in each point of M.



a field of covectors

We will define what it means for a covector field to be smooth. The definition is analogous to Definition 6.1.2.

Definition 8.2. A *covector field* ξ on M is an assignment of a covector

$$\xi(p) \in T_p^*M$$

to each $p \in M$. It is called *smooth* if, for each chart $\sigma\colon U \to M$ in a given atlas of M, there exist smooth functions $a_1 \dots, a_m \in C^\infty(\sigma(U))$ such that

$$\xi(p) = \sum_i a_i(p) dx_i(p) \tag{8.4}$$

for all $p \in \sigma(U)$. The space of smooth covector fields is denoted $\mathfrak{X}^*(M)$.

It follows from the next lemma that the notion of smoothness is independent of the atlas.

Notice that a covector $\xi(p)$ can be applied to tangent vectors in T_pM. If $\Omega \subset M$ is open and Y a vector field on Ω, it thus makes sense to apply a covector field ξ to Y. The result is a function, $\xi(Y)$ on Ω, given by

$$\xi(Y)(p) = \xi(p)(Y(p)) \in \mathbb{R}.$$

Lemma 8.2.1. *Let ξ be a covector field on M. Then ξ is smooth if and only if for each open set $\Omega \subset M$ and each smooth vector field $Y \in \mathfrak{X}(\Omega)$, the function $\xi(Y)$ belongs to $C^\infty(\Omega)$.*

Proof. Assume first that ξ is smooth. Let Ω be open, and let σ be a chart (of the given atlas). For each $Y \in \mathfrak{X}(\Omega)$ we can write $Y(p) = \sum b_j(p) d\sigma_u(e_j)$ for $p \in \Omega \cap \sigma(U)$, with coefficients $b_j(p)$ that are smooth functions of p. If we write ξ according to (8.4), it follows from (8.3) that

$$\xi(Y)(p) = \sum_i a_i(p) dx_i(p) (\sum_j b_j(p) d\sigma_u(e_j)) = \sum_i a_i(p) b_i(p),$$

and hence $\xi(Y)$ is smooth.

Conversely, assume that $\xi(Y) \in C^\infty(\Omega)$ for all Ω and all $Y \in \mathfrak{X}(\Omega)$. In particular, we can take $\Omega = \sigma(U)$ and $Y = d\sigma(e_i)$ for a given chart, and conclude that the function a_i defined by $a_i(p) = \xi(p)(d\sigma_u(e_i))$ depends smoothly on p. By applying (7.1) to the covector $\xi(p)$ and the dual bases (8.1)-(8.2) in Lemma 8.1 we see that (8.4) holds. Hence ξ is smooth. □

Notice that if ξ is a smooth covector field on M, and $\varphi \in C^\infty(M)$ a smooth function, then the covector field $\varphi\xi$ defined by the pointwise multiplication $(\varphi\xi)(p) = \varphi(p)\xi(p)$ for each p, is again smooth. This follows from Definition 8.2, since the coefficients of $\varphi\xi$ in (8.4) are just those of ξ, multiplied by φ.

If $f \in C^\infty(M)$ then the differential df is a covector field, as it is explained in Example 8.1.

Lemma 8.2.2. *If $f \in C^\infty(M)$ then df is a smooth covector field on M. For each chart σ on M, the expression for df by means of the basis* (8.2) *is*

$$df = \sum_{i=1}^{m} \frac{\partial(f \circ \sigma)}{\partial u_i} \, dx_i. \tag{8.5}$$

The differentials satisfy the rule $d(fg) = g\,df + f\,dg$ for all $f, g \in C^\infty(M)$.

Proof. The smoothness of df follows from (8.5), because the coefficient functions are smooth. The rule for the differential of fg also follows from (8.5), by means of the Leibniz rule for the partial differentiation of a product. Thus we only need to prove this equality.

In order to do that we first note that by the chain rule,

$$df_p(d\sigma_u(e_j)) = d(f \circ \sigma)_u(e_j) = \frac{\partial(f \circ \sigma)}{\partial u_j}(u).$$

By applying (7.1) to the covector df_p and the dual bases (8.1)-(8.2) we obtain the following expression from which (8.5) then follows

$$df_p = \sum_i df_p(d\sigma_u(e_i))dx_i(p). \quad \square$$

Example 8.2 Let $M = \mathbb{R}^2$ with coordinates (x_1, x_2), and let $f \in C^\infty(\mathbb{R}^2)$. Then the differential of f is given by (8.5) as

$$df = \frac{\partial f}{\partial x_1} \, dx_1 + \frac{\partial f}{\partial x_2} \, dx_2,$$

and an arbitrary smooth covector field is given by

$$\xi = a_1 \, dx_1 + a_2 \, dx_2$$

with $a_1, a_2 \in C^\infty(\mathbb{R}^2)$. A covector field which has the form df for some function f is said to be *exact*. It is a well-known result that ξ is exact if and only if it is *closed*, that is, if and only if

$$\frac{\partial a_1}{\partial x_2} = \frac{\partial a_2}{\partial x_1}. \tag{8.6}$$

That this is a necessary condition follows immediately from the fact that

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1}. \tag{8.7}$$

We will sketch the proof that (8.6) is also sufficient. Let a closed form $\xi = a_1\, dx_1 + a_2\, dx_2$ on $\mathbb{R}^2$ be given. Choose a function $f_1 \in C^\infty(\mathbb{R}^2)$ with $\frac{\partial f_1}{\partial x_1} = a_1$ (for example, $f_1(x_1, x_2) = \int_0^{x_1} a_1(t, x_2)\, dt$). It follows from (8.6) and (8.7) that $\frac{\partial a_2}{\partial x_1} = \frac{\partial^2 f_1}{\partial x_1 \partial x_2}$. We deduce that $a_2 - \frac{\partial f_1}{\partial x_2}$ does not depend on x_1, and choose a one-variable function $f_2 \in C^\infty(\mathbb{R})$ with $f_2' = a_2 - \frac{\partial f_1}{\partial x_2}$, as a function of x_2. It is easily seen that $f(x_1, x_2) = f_1(x_1, x_2) + f_2(x_2)$ satisfies both $a_1 = \frac{\partial f}{\partial x_1}$ and $a_2 = \frac{\partial f}{\partial x_2}$.

The analogous result is not valid in an arbitrary 2-dimensional manifold (for example, it fails on $M = \mathbb{R}^2 \setminus \{(0,0)\}$).

8.3 Differential forms

The notion of a covector field can be generalized to a notion of fields of k-forms. Recall from Definition 7.4 that we denote by $A^k(V)$ the space of alternating k-tensors on the vector space V. Recall also that $A^0(V) = \mathbb{R}$ and $A^1(V) = V^*$, and if $\xi_1, \dots, \xi_m$ is a basis for V^*, then a basis for $A^k(V)$ is obtained from the set of elements

$$\xi_{i_1} \wedge \cdots \wedge \xi_{i_k}$$

where $1 \le i_1 < \cdots < i_k \le m$ (see Theorem 7.4).

Definition 8.3.1. A *k-form* ω on M is an assignment of an element

$$\omega(p) \in A^k(T_pM)$$

to each $p \in M$.

In particular, given a chart $\sigma\colon U \to M$, the elements $dx_{i_1} \wedge \cdots \wedge dx_{i_k}$, where $1 \le i_1 < \cdots < i_k \le m$, are k-forms on the open subset $\sigma(U)$ of M. For each $p \in \sigma(U)$, a basis for $A^k(T_pM)$ is obtained from these elements. Therefore, every k-form ω on M has a unique expression on $\sigma(U)$,

$$\omega = \sum_{I=\{i_1,\dots,i_k\}} a_I\ dx_{i_1} \wedge \cdots \wedge dx_{i_k}$$

where $a_I\colon \sigma(U) \to \mathbb{R}$.

Definition 8.3.2. We call ω *smooth* if all the functions a_I are smooth, for each chart σ in an atlas of M. A smooth k-form is also called a *differential k-form.* The space of differential k-forms on M is denoted $A^k(M)$.

In particular, since $A^0(T_pM) = \mathbb{R}$, we have $A^0(M) = C^\infty(M)$, that is, a differential 0-form on M is just a smooth function $\varphi \in C^\infty(M)$. Likewise, a differential 1-form is nothing but a smooth covector field.

It can be verified, analogously to Lemma 8.2.1, that a k-form ω is smooth if and only if $\omega(X_1, \dots, X_k) \in C^\infty(\Omega)$ for all open sets Ω and all $X_1, \dots, X_k \in \mathfrak{X}(M)$. In particular, it follows that the notion of smoothness of a k-form is independent of the chosen atlas for M.

Example 8.3 Let $\mathcal{S}$ be an oriented surface in $\mathbb{R}^3$. The *volume form* on $\mathcal{S}$ is the 2-form defined by $\omega_p(X_1, X_2) = (X_1 \times X_2) \cdot N$ for $X, Y \in T_p\mathcal{S}$, where $N \in \mathbb{R}^3$ is the positive unit normal. In the local coordinates of a chart, it is given by

$$\omega = (EG - F^2)^{1/2} dx_1 \wedge dx_2$$

where E, F and G are the coefficients of the first fundamental form. The volume form is smooth, since these coefficients are smooth functions.

Lemma 8.3. *Let ω be a k-form on M and φ a real function on M, and define the product $\varphi\omega$ pointwise by*

$$\varphi\omega(p) = \varphi(p)\omega(p).$$

Then $\varphi\omega$ is a smooth k-form if φ and ω are smooth.

Let θ be an l-form on M, and define the wedge product $\theta \wedge \omega$ pointwise by

$$(\theta \wedge \omega)(p) = \theta(p) \wedge \omega(p)$$

for each p. Then $\theta \wedge \omega$ is a smooth $k + l$-form if ω and θ are smooth.

Proof. Easy. $\square$

8.4 Pull back

An important property of differential forms is that they can be pulled back by a map from one manifold to another. The construction is analogous to that of the dual map in Section 7.2.

Definition 8.4. Let $f\colon M \to N$ be a smooth map of manifolds, and let $\omega \in A^k(N)$. We define $f^*\omega \in A^k(M)$, called the *pull back* of ω, by

$$f^*\omega(p)(v_1, \dots, v_k) = \omega(f(p))(df_p(v_1), \dots, df_p(v_k))$$

for all $v_1, \dots, v_k \in T_pM$ and $p \in M$.

It is easily seen that for each $p \in M$ the operator $f^*\omega(p)$ so defined is multilinear and alternating, so that it belongs to $A^k(T_pM)$. We have to verify that it depends smoothly on p. This will be done in the following lemma. It is clear that $f^*\omega$ depends linearly on ω, this will be used in the proof of the lemma.

Let

$$\omega = \sum_{I=\{i_1,\dots,i_k\}} a_I \; dy_{i_1} \wedge \dots \wedge dy_{i_k}$$

be the expression for ω by means of the coordinates of a chart $\tau: V \to N$ on N. Here y_i is the i-th coordinate of τ^{-1} for $i = 1, \dots, n$, and I is an increasing sequence of k numbers among $1, \dots, n$.

Lemma 8.4.1. *With notation from above,*

$$f^*\omega = \sum_I (a_I \circ f)\, d(y_{i_1} \circ f) \wedge \dots \wedge d(y_{i_k} \circ f) \tag{8.8}$$

on $f^{-1}(\tau(V))$. In particular, $f^*\omega$ is smooth.

Proof. In view of Lemma 8.3, the last statement follows from (8.8), since all the ingredients in this expression are smooth, and since M is covered by the open subsets $f^{-1}(\tau(V))$, for all τ in an atlas of N.

In order to verify (8.8), it suffices to prove the following rules for pull backs:

$$f^*(\varphi\omega) = (\varphi \circ f) f^*\omega, \tag{8.9}$$

$$f^*(\theta \wedge \omega) = f^*(\theta) \wedge f^*(\omega), \tag{8.10}$$

$$f^*(d\varphi) = d(\varphi \circ f). \tag{8.11}$$

These rules follow easily from the definitions. □

In particular, Definition 8.4 allows us to define the *restriction* of a k-form on a manifold N to a submanifold M. Let $i: M \to N$ be the inclusion map, then we define $\omega|_M = i^*\omega$. It is given at $p \in M$ simply by

$$\omega|_M(p)(v_1, \dots, v_k) = \omega(p)(v_1, \dots, v_k)$$

for $v_1, \dots, v_k \in T_pM \subset T_pN$.

The following result will be important later.

Lemma 8.4.2. *Let $f: M \to N$ be a smooth map between two manifolds of the same dimension k. Let $x_1, \dots, x_k$ denote the coordinate functions of a* *chart σ on M, and let $y_1, \dots, y_k$ denote the coordinate functions of a chart τ on N. Then*

$$f^*(dy_1 \wedge \cdots \wedge dy_k) = \det(df)\, dx_1 \wedge \cdots \wedge dx_k \tag{8.12}$$

where $\det(df)$ in p is the determinant of the matrix for df_p with respect to the standard bases for T_pM and $T_{f(p)}N$ given by the charts.

Proof. The element $f^*(dy_1 \wedge \cdots \wedge dy_k)$ is a k-form on M, hence a constant multiple of $dx_1 \wedge \cdots \wedge dx_k$ (recall that $\dim A^k(V) = 1$ when $\dim V = k$). We determine the constant by applying $f^*(dy_1 \wedge \cdots \wedge dy_k)$ to the vector $(d\sigma_p(e_1), \dots, d\sigma_p(e_k))$ in $(T_pM)^k$. We obtain, see (7.6),

$$\begin{aligned} &f^*(dy_1 \wedge \cdots \wedge dy_k)(d\sigma_p(e_1), \dots, d\sigma_p(e_k)) \\ &\quad = (dy_1 \wedge \cdots \wedge dy_k)(df_p(d\sigma_p(e_1)), \dots, df_p(d\sigma_p(e_k))) \\ &\quad = \frac{1}{k!} \det(dy_i(df_p(d\sigma_p(e_j)))). \end{aligned}$$

On the other hand, applying $dx_1 \wedge \cdots \wedge dx_k$ to the same vector in $(T_pM)^k$, we obtain, again by (7.6),

$$(dx_1 \wedge \cdots \wedge dx_k)(d\sigma_p(e_1), \dots, d\sigma_p(e_k)) = \frac{1}{k!} \det(\delta_{ij}) = \frac{1}{k!},$$

since $dx_1, \dots, dx_k$ is the dual basis. It follows that the constant we are seeking is the determinant $\det(dy_i(df_p(d\sigma_p(e_j))))$

Notice that $dy_i(df_p(d\sigma_p(e_j)))$ are exactly the matrix elements for df_p. The formula (8.12) now follows. $\square$

8.5 Exterior differentiation

In this section a differentiation operator d is defined for differential forms on a smooth manifold M. It is a generalization of the operator which maps a smooth function to its differential, in general it will map k-forms to $k+1$-forms. The operator simply consists of applying d to the coefficients a_I in the coordinate expression for the k-form.

Definition 8.5. Fix a chart $\sigma: U \to M$, and let $\omega \in A^k(M)$. On its image $\sigma(U) \subset M$ we can write

$$\omega = \sum_I a_I\, dx_{i_1} \wedge \cdots \wedge dx_{i_k} \tag{8.13}$$

where $I = (i_1, \dots, i_k)$ with $1 \leq i_1 < \cdots < i_k \leq m = \dim M$, and where $a_I \in C^\infty(\sigma(U))$ for each I (see below Definition 8.3.1). On the image $\sigma(U)$ we then define

$$d\omega = \sum_I da_I \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k}, \tag{8.14}$$

where da_I is the differential of a_I. Then $d\omega \in A^{k+1}(\sigma(U))$.

The operator d is called *exterior differentiation.* It is defined above by reference to a fixed chart σ, but we shall see in the following theorem that it can be extended to a globally defined operator, which is independent of any choice of chart or atlas. The theorem also mentions the most important properties of d.

Theorem 8.5. *There exists a unique map $d\colon A^k(M) \to A^{k+1}(M)$ such that (8.14) holds on $\sigma(U)$ for each chart σ in a given atlas for M. This map d is independent of the choice of the atlas. Furthermore, it has the following properties:*

(a) *if $k = 0$ it agrees with the differential d on functions,*

(b) *it is linear,*

(c) *$d(\varphi\omega) = d\varphi \wedge \omega + \varphi\, d\omega$ for $\varphi \in C^\infty(M), \omega \in A^k(M)$,*

(d) *$d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^k \omega_1 \wedge d\omega_2$ for $\omega_1 \in A^k(M), \omega_2 \in A^l(M)$,*

(e) *$d(df_1 \wedge \cdots \wedge df_k) = 0$ for all $f_1, \dots, f_k \in C^\infty(M)$,*

(f) *$d(d\omega) = 0$ for all $\omega \in A^k(M)$.*

Proof. Suppose an atlas is given. The map d is unique because (8.14) determines $d\omega$ uniquely on $\sigma(U)$, for each chart σ in the atlas. We have to verify the existence of d, the independency on the atlas, and the listed properties. However, before we commence on this we shall prove that the chart-dependent operator of Definition 8.5 satisfies (a)-(f).

(a) Let $k = 0$ and $\omega = f \in C^\infty(M)$. The only summand in the expression (8.13) corresponds to $I = \emptyset$, and the coefficient $a_\emptyset$ is equal to f itself. The expression (8.14) then reads $d\omega = df$ (on $\sigma(U)$).

(b) The coefficients a_I in (8.13) depend linearly on ω, and the expression (8.14) is linear with respect to these coefficients, hence $d\omega$ depends linearly on ω.

(c) Let $\varphi \in C^\infty(M)$, $\omega \in A^k(M)$, and write ω according to (8.13). In the corresponding expression for the product $\varphi\omega$ each coefficient a_I is then multiplied by φ, and hence definition (8.14) for this product reads

$$\begin{aligned}
d(\varphi\omega) &= \sum_I d(\varphi a_I) \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k} \\
&= \sum_I (a_I\, d\varphi + \varphi\, da_I) \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k} \\
&= d\varphi \wedge \sum_I a_I\, dx_{i_1} \wedge \cdots \wedge dx_{i_k} + \varphi \sum_I da_I \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k} \\
&= d\varphi \wedge \omega + \varphi\, d\omega
\end{aligned}$$

where the rule in Lemma 8.2.2 was applied in the second step.

(d) Notice first that if $I = (i_1, \dots, i_k)$ with $1 \le i_1 < \dots < i_k \le m$ then by definition

$$d(a\, dx_{i_1} \wedge \dots \wedge dx_{i_k}) = da \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k} \tag{8.15}$$

for $a \in C^\infty(M)$. The expression (8.15) holds also without the assumption that the indices $i_1, \dots, i_k$ are increasingly ordered, for the $dx_{i_1}, \dots, dx_{i_k}$ can be reordered at the cost of the same change of sign on both sides. The expression also holds if the indices are not distinct, because in that case both sides are 0.

Let $\omega_1 \in A^k(M)$, $\omega_2 \in A^l(M)$, and write both of these forms as follows, according to (8.13):

$$\omega_1 = \sum_I a_I\, dx_{i_1} \wedge \dots \wedge dx_{i_k}, \quad \omega_2 = \sum_J b_J\, dx_{j_1} \wedge \dots \wedge dx_{j_l},$$

where $I = (i_1, \dots, i_k)$ and $J = (j_1, \dots, j_l)$. Then

$$\omega_1 \wedge \omega_2 = \sum_{I,J} a_I b_J dx_{i_1} \wedge \dots \wedge dx_{i_k} \wedge dx_{j_1} \wedge \dots \wedge dx_{j_l}$$

and hence according to the remarks just made

$$\begin{aligned} d(\omega_1 \wedge \omega_2) &= \sum_{I,J} d(a_I b_J) \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k} \wedge dx_{j_1} \wedge \dots \wedge dx_{j_l} \\ &= \sum_{I,J} (b_J\, da_I + a_I\, db_J) \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k} \wedge dx_{j_1} \wedge \dots \wedge dx_{j_l} \\ &= \sum_{I,J} b_J\, da_I \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k} \wedge dx_{j_1} \wedge \dots \wedge dx_{j_l} \\ &\qquad + \sum_{I,J} a_I\, db_J \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k} \wedge dx_{j_1} \wedge \dots \wedge dx_{j_l} \\ &= d\omega_1 \wedge \omega_2 + (-1)^k \omega_1 \wedge\, d\omega_2 \end{aligned}$$

where the sign $(-1)^k$ in the last step comes from passing db_J to the right past the elements dx_{i_1} up to dx_{i_k}.

(e) is proved first for $k = 1$. In this case $f \in C^\infty(M)$ and

$$df = \sum_{i=1}^m \frac{\partial (f \circ \sigma)}{\partial x_i}\, dx_i.$$

Hence

$$d(df) = \sum_{i=1}^m d\Big[\frac{\partial (f \circ \sigma)}{\partial x_i}\Big] \wedge dx_i = \sum_{i=1}^m \sum_{j=1}^m \frac{\partial^2 (f \circ \sigma)}{\partial x_j \partial x_i} dx_j \wedge dx_i.$$

The latter expression vanishes because the coefficient $\frac{\partial^2(f\circ\sigma)}{\partial x_j \partial x_i} = \frac{\partial^2(f\circ\sigma)}{\partial x_i \partial x_j}$ occurs twice with opposite signs.

The general case is now obtained by induction on k. It follows from property (d) that $d(df_1 \wedge \cdots \wedge df_k)$ can be written as a sum of two terms each of which involve a factor where d is applied to a wedge product with fewer factors. Applying the induction hypothesis, we infer the statement.

(f) Let ω be given by (8.13). Then

$$d(d\omega) = d(\sum_I da_I \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k}) = 0$$

by property (e).

Having established these properties for the operator d on $\sigma(U)$, we proceed with the definition of the global operator d. Let $p \in M$ be given and choose a chart σ around p. We define $d\omega(p)$ for each $\omega \in A^k(M)$ by the expression in Definition 8.5. We need to prove that $d\omega(p)$ is independent of the choice of chart (this will also establish the independency of the atlas).

Suppose $d'\omega$ is defined on $\sigma'(U')$ for some other chart σ'. The claim is that then $d'\omega = d\omega$ on the overlap $\sigma'(U') \cap \sigma(U)$ of the two charts.

Assume first that $U' \subset U$ and $\sigma' = \sigma|_{U'}$. In the expression (8.13) for ω on $\sigma'(U')$ we then have the restriction of each a_I to this set, and since $d(a_I|_{\sigma'(U')}) = da_I|_{\sigma'(U')}$ it follows that the expression (8.14) for $d'\omega$ on $\sigma'(U')$ is just the restriction of the same expression for $d\omega$. Hence $d'\omega = d\omega$ on $\sigma'(U')$ as claimed.

For the general case the observation we just made implies that we can replace U by the open subset $\sigma^{-1}(\sigma'(U') \cap \sigma(U))$ and σ by its restriction to this subset, without affecting $d\omega$ on $\sigma'(U') \cap \sigma(U)$. Likewise, we can replace U' by $\sigma'^{-1}(\sigma'(U') \cap \sigma(U))$. The result is that we may assume $\sigma(U) = \sigma'(U')$. Moreover, we can replace M by the open subset $\sigma(U) = \sigma'(U')$, and ω by its restriction to this set, since this has no effect on the expressions (8.13) and (8.14). In conclusion, we may assume that $\sigma(U) = \sigma'(U') = M$.

We now apply the rules (a)-(f) for d to the expression

$$\omega = \sum_I a'_I \, dx'_{i_1} \wedge \cdots \wedge dx'_{i_k}$$

for ω with respect to the chart σ'. It follows from (b), (c) and (e) that

$$\begin{aligned} d\omega &= \sum_I d(a'_I \, dx'_{i_1} \wedge \cdots \wedge dx'_{i_k}) \\ &= \sum_I da'_I \wedge dx'_{i_1} \wedge \cdots \wedge dx'_{i_k}, \end{aligned}$$

and hence $d\omega = d'\omega$ as claimed.

It is now clear that $d\omega$ is a well defined smooth $k+1$-form. The properties (a)-(f) hold, since they hold for the restrictions to $\sigma(U)$, for any chart σ. $\square$

As in Section 7.6 we define

$$A(M) = A^0(M) \oplus A^1(M) \oplus \cdots \oplus A^m(M)$$

where $m = \dim M$. The elements of $A(M)$ are called differential forms. Thus a differential form is a map which associates to each point $p \in M$ a member of the exterior algebra $A(T_pM)$, in a smooth manner. The operator d maps the space of differential forms to itself.

8.6 Some examples

In the following we treat $M = \mathbb{R}^n$ for various values of n. In each case we denote by $\sigma\colon \mathbb{R}^n \to \mathbb{R}^n$ the standard chart on $\mathbb{R}^n$, the identity map.

Example 8.6.1 Let $M = \mathbb{R}$. Since $\dim M = 1$, we have $A(\mathbb{R}) = A^0(\mathbb{R}) \oplus A^1(\mathbb{R})$. Furthermore, $A^0(\mathbb{R}) = C^\infty(\mathbb{R})$ and $A^1(\mathbb{R}) = \{a\, dx \mid a \in C^\infty(\mathbb{R})\}$. The exterior differentiation is given by $df = f'\, dx$ for $f \in A^0(\mathbb{R})$ and $df = 0$ for $f \in A^1(\mathbb{R})$.

Example 8.6.2 Let $M = \mathbb{R}^2$. Since $\dim M = 2$, we have

$$A(\mathbb{R}^2) = A^0(\mathbb{R}^2) \oplus A^1(\mathbb{R}^2) \oplus A^2(\mathbb{R}^2).$$

Furthermore, $A^0(\mathbb{R}^2) = C^\infty(\mathbb{R}^2)$ and

$$\begin{aligned} A^1(\mathbb{R}^2) &= \{a_1\, dx_1 + a_2\, dx_2 \mid a_1, a_2 \in C^\infty(\mathbb{R}^2)\} \\ A^2(\mathbb{R}^2) &= \{a\, dx_1 \wedge dx_2 \mid a \in C^\infty(\mathbb{R}^2)\} \end{aligned}$$

The exterior differentiation is given by

$$df = \frac{\partial f}{\partial x_1}\, dx_1 + \frac{\partial f}{\partial x_2}\, dx_2$$

(see Example 8.2), and

$$\begin{aligned} d(a_1\, dx_1 + a_2\, dx_2) & \\ &= da_1 \wedge dx_1 + da_2 \wedge dx_2 \\ &= \left(\frac{\partial a_1}{\partial x_1}\, dx_1 + \frac{\partial a_1}{\partial x_2}\, dx_2\right) \wedge dx_1 + \left(\frac{\partial a_2}{\partial x_1}\, dx_1 + \frac{\partial a_2}{\partial x_2}\, dx_2\right) \wedge dx_2 \\ &= \left(-\frac{\partial a_1}{\partial x_2} + \frac{\partial a_2}{\partial x_1}\right) dx_1 \wedge dx_2 \end{aligned}$$

Example 8.6.3 Let $M = \mathbb{R}^3$. Then

$$A(\mathbb{R}^3) = A^0(\mathbb{R}^3) \oplus A^1(\mathbb{R}^3) \oplus A^2(\mathbb{R}^3) \oplus A^3(\mathbb{R}^3).$$

Furthermore, $A^0(\mathbb{R}^2) = C^\infty(\mathbb{R}^2)$ and

$$\begin{aligned}
A^1(\mathbb{R}^3) &= \{a_1\,dx_1 + a_2\,dx_2 + a_3\,dx_3 \mid a_1, a_2, a_3 \in C^\infty(\mathbb{R}^3)\}\\
A^2(\mathbb{R}^3) &= \{b_1 dx_2 \wedge dx_3 + b_2 dx_3 \wedge dx_1 + b_3 dx_1 \wedge dx_2 \mid b_1, b_2, b_3 \in C^\infty(\mathbb{R}^3)\}\\
A^3(\mathbb{R}^3) &= \{a\,dx_1 \wedge dx_2 \wedge dx_3 \mid a \in C^\infty(\mathbb{R}^3)\}.
\end{aligned}$$

The exterior differentiation is given by

$$df = \frac{\partial f}{\partial x_1}\,dx_1 + \frac{\partial f}{\partial x_2}\,dx_2 + \frac{\partial f}{\partial x_3}\,dx_3$$

and

$$\begin{aligned}
d(a_1\,dx_1 &+ a_2\,dx_2 + a_3\,dx_3)\\
&= da_1 \wedge dx_1 + da_2 \wedge dx_2 + da_3 \wedge dx_3\\
&= \left(\frac{\partial a_1}{\partial x_1}\,dx_1 + \frac{\partial a_1}{\partial x_2}\,dx_2 + \frac{\partial a_1}{\partial x_3}\,dx_3\right) \wedge dx_1\\
&\qquad + \left(\frac{\partial a_2}{\partial x_1}\,dx_1 + \frac{\partial a_2}{\partial x_2}\,dx_2 + \frac{\partial a_2}{\partial x_3}\,dx_3\right) \wedge dx_2\\
&\qquad\qquad + \left(\frac{\partial a_3}{\partial x_1}\,dx_1 + \frac{\partial a_3}{\partial x_2}\,dx_2 + \frac{\partial a_3}{\partial x_3}\,dx_3\right) \wedge dx_3\\
&= \left(\frac{\partial a_3}{\partial x_2} - \frac{\partial a_2}{\partial x_3}\right) dx_2 \wedge dx_3\\
&\qquad + \left(\frac{\partial a_1}{\partial x_3} - \frac{\partial a_3}{\partial x_1}\right) dx_3 \wedge dx_1\\
&\qquad\qquad + \left(\frac{\partial a_2}{\partial x_1} - \frac{\partial a_1}{\partial x_2}\right) dx_1 \wedge dx_2
\end{aligned}$$

Furthermore,

$$\begin{aligned}
d\,(b_1 dx_2 &\wedge dx_3 + b_2 dx_3 \wedge dx_1 + b_3 dx_1 \wedge dx_2)\\
&= db_1 \wedge dx_2 \wedge dx_3 + db_2 \wedge dx_3 \wedge dx_1 + db_3 \wedge dx_1 \wedge dx_2\\
&= \left(\frac{\partial b_1}{\partial x_1} + \frac{\partial b_2}{\partial x_2} + \frac{\partial b_3}{\partial x_3}\right) dx_1 \wedge dx_2 \wedge dx_3
\end{aligned}$$

and

$$d(a\,dx_1 \wedge dx_2 \wedge dx_3) = 0.$$

The expressions that occur as coefficients of $d\omega$ in the three degrees are familiar from vector calculus. The vector field

$$\operatorname{grad} f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}\right)$$

is the *gradient* of f, and the vector field

$$\operatorname{curl} a = \left(\frac{\partial a_3}{\partial x_2} - \frac{\partial a_2}{\partial x_3}, \frac{\partial a_1}{\partial x_3} - \frac{\partial a_3}{\partial x_1}, \frac{\partial a_2}{\partial x_1} - \frac{\partial a_1}{\partial x_2}\right)$$

is the *curl* (or rotation) of a. Finally, the function

$$\operatorname{div} b = \frac{\partial b_1}{\partial x_1} + \frac{\partial b_2}{\partial x_2} + \frac{\partial b_3}{\partial x_3}$$

is the *divergence* of b.

8.7 Exterior differentiation and pull back

The following property of exterior differentiation plays an important role in the next chapter.

Theorem 8.7. *Let $g\colon M \to N$ be a smooth map of manifolds, and let*

$$g^*\colon A^k(N) \to A^k(M)$$

denote the pull-back, for each k (see Definition 8.4). Then

$$d_M \circ g^* = g^* \circ d_N,$$

where d_M and d_N denotes the exterior differentiation for the two manifolds.

Proof. By Lemma 8.4.1, and properties (b), (c) and (e) above,

$$\begin{aligned} d(g^*\omega) &= d(\sum_I (a_I \circ g)\, d(y_{i_1} \circ g) \wedge \cdots \wedge d(y_{i_k} \circ g)) \\ &= \sum_I d(a_I \circ g) \wedge d(y_{i_1} \circ g) \wedge \cdots \wedge d(y_{i_k} \circ g). \end{aligned}$$

On the other hand, by the properties of g^* listed in the proof of Lemma 8.4.1,

$$\begin{aligned} g^*(d\omega) &= g^*(\sum_I da_I \wedge dy_{i_1} \wedge \cdots \wedge dy_{i_k}) \\ &= \sum_I d(a_I \circ g) \wedge d(y_{i_1} \circ g) \wedge \cdots \wedge d(y_{i_k} \circ g). \quad \square \end{aligned}$$

Corollary 8.7. *Let M be a submanifold of the abstract manifold N, and let ω be a differential form on N. Then the restriction of ω to M is a differential form on M, and it satisfies $d(\omega|_M) = d\omega|_M$, that is, exterior differentiation and restriction commute.*

Proof. The restriction $\omega|_M$ is exactly the pull back $i^*\omega$ by the inclusion map $i\colon M \to N$. The result is an immediate consequence of Theorem 8.7. $\square$

Chapter 9

Integration

The purpose of this chapter is to define integration on smooth manifolds, and establish its relation with the differentiation operator d of the previous chapter.

9.1 Null sets

Recall from Geometry 1 that a compact set $D \subset \mathbb{R}^2$ is said to be a null set, if for each $\epsilon > 0$ it can be covered by a finite union of rectangles with total area $< \epsilon$. We will generalize this concept to subsets D of an abstract manifold M.

We first treat the case of sets in $\mathbb{R}^n$. The notion of a rectangle is generalized as a *box*, by which we mean a product set $D = I_1 \times \cdots \times I_n \subset \mathbb{R}^n$, where each I_i is a closed interval $[a_i, b_i]$ with $-\infty < a_i \le b_i < \infty$. The *volume* of D is the product of the lengths of the intervals. The box is called a *cube* if the intervals have equal length.

Definition 9.1.1. A set $A \subset \mathbb{R}^n$ is said to be a *null set* if for each $\epsilon > 0$ it can be covered by a countable collection of boxes, of which the sum of volumes is $< \epsilon$.

It is clear that any subset of a null set is again a null set.

Example 9.1.1 Clearly a singleton $A = \{x\}$ is a null set. More generally every countable subset of $\mathbb{R}^n$ is a null set, since we can cover it by countably many boxes, each of volume 0.

Example 9.1.2 Let $m < n$ and consider a set in $\mathbb{R}^n$ of the form

$$A = \{(x_1, \dots, x_m, 0, \dots, 0) \mid x \in \mathbb{R}^m\}.$$

By writing $\mathbb{R}^m$ as a countable union of boxes B_i we can arrange that A is covered by the boxes $B_i \times \{0\}$, all of volume 0.

In fact for our present purpose we need only consider compact sets A. In this case, we can replace the word *countable* by *finite* in the above definition. Indeed, suppose a compact set A is covered by countably many boxes with volume sum $< \epsilon$. By increasing each box slightly, say the volume of the i-th box is increased at most by $2^{-i}\epsilon$, we can arrange that A is contained in a countable union of *open* boxes with volume sum $< 2\epsilon$. By Definition 5.1.2, A is then covered by a finite subcollection, which of course also has volume sum $< 2\epsilon$.

Lemma 9.1.2. *Let $A \subset \mathbb{R}^n$ be a null set, and let $f: A \to \mathbb{R}^n$ be a smooth map. Then $f(A)$ is a null set.*

Proof. We only give the proof for the case that A is compact. By Definition 2.6.1, f has an extension F to a smooth map defined on an open set W containing A. Let $D \subset W$ be compact, then there exists an upper bound C on the absolute values of the first order derivatives of F over D. Using the mean value theorem, one then obtains an estimate $|F(x) - F(y)| \leq C|x - y|$ for all $x, y \in D$. It follows from this estimate that the image by F of a ball inside D of radius r is contained in a ball of radius Cr.

Let $\epsilon > 0$ be given and choose a finite collection of boxes, which cover A and whose volumes add to $< \epsilon$. By possibly increasing some of the boxes slightly, we can arrange that all the boxes have positive and rational side lengths, still with volumes adding to $< \epsilon$. We can then write each box as a union of cubes of any size which divides the lengths of all sides of the box, and which only intersect along their boundary. We thus obtain a covering of A by finitely many cubes, whose volumes add to $< \epsilon$. The circumscribed ball of a cube has a volume which is proportional to that of the cube, with a factor c that depends only on n. It follows that we can cover A by finitely many arbitrarily small closed balls, whose volumes add to $< c\epsilon$. In particular, we may assume that each of these balls is contained in W. The discussion in the first part of the proof (with D equal to the union of the balls) then implies that $f(A)$ is covered by finitely many balls of total volume $< C^n c\epsilon$. By circumscribing these balls with cubes, we obtain a covering of $F(A)$ by boxes as desired. $\square$

Definition 9.1.2. Let M be an abstract manifold of dimension m. A subset $A \subset M$ is called a *null set* in M, if there exists a countable collection of charts $\sigma_i: U_i \to M$ and null sets $A_i \subset U_i$ such that $A = \cup_i \sigma_i(A_i)$.

In order for this to be a reasonable definition, we should prove that a diffeomorphism of open sets in $\mathbb{R}^m$ carries a null set to a null set, so that the condition on A_i is unchanged if σ_i is reparametrized. This follows from Lemma 9.1.2.

Example 9.1.3 Let N be a submanifold of M, with $\dim N < \dim M$, and assume that N has a countable atlas. Then N is a null set in M. It follows from Theorem 4.3 that for each $p \in N$ there exists a chart σ on M around p, such that $N \cap \sigma(U)$ is the image of a null set in U (see Example 9.1.2). A countable collection of these charts will already cover N, since N is second countable (see Lemma 2.9).

9.2 Integration on $\mathbb{R}^n$

Recall from Geometry 1, Chapter 3, that a non-empty subset $D \subset \mathbb{R}^2$ is called an *elementary domain* if it is compact and if its boundary is a finite union of smooth curves $[a, b] \to \mathbb{R}^2$. If D is an elementary domain, $U \subset \mathbb{R}^2$ an open set and $f\colon U \to \mathbb{R}$ a continuous function, the *plane integral* $\int_D f\,dA$ was given a sense in Geometry 1. The crucial property of D in this definition is that D is compact and that its boundary is a null set. In particular, if $D = [a, b] \times [c, d]$ is a rectangle,

$$\int_D f\,dA = \int_a^b \int_c^d f(u, v)\,dv\,du = \int_c^d \int_a^b f(u, v)\,du\,dv.$$

These definitions can be generalized to $\mathbb{R}^n$ in analogous fashion. A non-empty compact subset $D \subset \mathbb{R}^n$ is called a *domain of integration* if its boundary is a null set in $\mathbb{R}^n$. Assume that D is a domain of integration and that $D \subset U$, where U is open. In this case, if $f\colon U \to \mathbb{R}$ is continuous, the *integral* $\int_D f\,dV$ can be given a sense. The details are omitted since they are analogous to those in Geometry 1.

In particular, if $D = [a_1, b_1] \times \cdots \times [a_n, b_n]$ is a generalized rectangle, then

$$\int_D f\,dV = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, \dots, x_n)\,dx_n \dots dx_1,$$

and the order of the consecutive integrals over $x_1, \dots, x_n$ can be interchanged freely. Of course, for $n = 1$ the concept coincides with the ordinary integral of a continuous function if D is an interval. In the trivial case $n = 0$, where $\mathbb{R}^n = \{0\}$, we interprete the definition so that the integral over $D = \{0\}$ of f is the evaluation of f in 0.

Example 9.2.1 If $\Omega \subset \mathbb{R}^n$ is bounded and is a domain with smooth boundary, then the closure $D = \bar{\Omega}$ is a domain of integration. The boundary $\partial\Omega$ is an $n-1$-dimensional manifold, hence a null set according to Example 9.1.3.

The integral $\int_D f\,dV$ has the following properties, also in analogy with the two-dimensional case:

$$\begin{gathered}
\int_D f + g\,dA = \int_D f\,dA + \int_D g\,dA \\
\int_D cf\,dA = c\int_D f\,dA \\
|\int_D f\,dA\,| \le \int_D |f|\,dA \\
\int_{D_1 \cup D_2} f\,dA = \int_{D_1} f\,dA + \int_{D_2} f\,dA,
\end{gathered}$$

where in the last line D_1 and D_2 are domains of integration with disjoint interiors.

There is an important rule for change of variables, which reads as follows. It will not be proved here.

Theorem 9.2. *Let $\phi: U_1 \to U_2$ be a diffeomorphism of open sets, and let $D \subset U_1$ be a domain of integration. Then $\phi(D) \subset U_2$ is a domain of integration and*

$$\int_{\phi(D)} f\, dV = \int_D (f \circ \phi)\, |\det(D\phi)|\, dV$$

for all continuous $f: U_2 \to \mathbb{R}$.

9.3 Integration on a chart

Let $\mathcal{S} \subset \mathbb{R}^3$ be a surface, and let $\sigma: U \to M$ be a chart on it. If $D \subset U$ is a domain of integration, the *area* of the set $R = \sigma(D) \subset M$ is defined as the integral $\int_D (EG - F^2)^{1/2}\, dA$, where E, F, G are the coefficients of the first fundamental form. It was shown in Geometry 1 that the area is independent of the chosen parametrization σ, thanks to the presence of the function $(EG - F^2)^{1/2}$. The Jacobian determinant that arises in a change of coordinates, is exactly compensated by the change of this square root when E, F and G are replaced by their counterparts for the new coordinates.

Similarly, if $f: \mathcal{S} \to \mathbb{R}$ is a continuous function, its integral over R is defined to be

$$\int_R f\, dA = \int_D (f \circ \sigma)\, (EG - F^2)^{1/2}\, dA.$$

This integral is independent of the chart, for the same reason as before.

These considerations can be generalized to an m-dimensional smooth surface in $\mathbb{R}^k$. The factor $(EG - F^2)^{1/2}$ in the integral over D is generalized as the square root of the determinant $\det(\sigma'_{u_i} \cdot \sigma'_{u_j})$, $i, j = 1, \ldots, m$, for a given chart σ. The dot is the scalar product from $\mathbb{R}^n$.

When trying to generalize further to an abstract m-dimensional manifold M we encounter the problem that in general there is no dot product on T_pM, hence there is no way to generalize E, F and G. As a consequence, it is not possible to define the integral of a function on M in a way that is invariant under changes of of charts. One way out is to introduce the presence of an inner product on tangent spaces as an extra axiom in the definition of the concept of a manifold - this leads to so-called *Riemannian geometry*, which is the proper abstract framework for the theory of the first fundamental form.

Another solution is to replace the integrand including $(EG - F^2)^{1/2}$ by a differential form of the highest degree. This approach, which we shall follow here, leads to a theory of integration for m-forms, but not for functions. In particular, it does not lead to a definition of area, since that would be obtained from the integration of the constant 1, which is a function, not a form. However, we can still view the theory as a generalization to the theory for surfaces, by defining the integral of a function f over a surface as the integral of f times the volume form (see Example 8.3).

For some applications it is important to be able to integrate forms which are continuous, but not smooth. The definition of continuity of forms is analogous to Definition 8.3.2: A *continuous k-form* on an abstract manifold M is a k-form ω for which all the functions a_I are continuous, for each chart in an atlas of M. In particular, an m-form on an m-dimensional manifold thus has the form

$$\omega = \phi\, dx_1 \wedge \cdots \wedge dx_m \tag{9.1}$$

on each chart σ, with ϕ a continuous function on $\sigma(U)$.

Definition 9.3. Let M be an m-dimensional oriented abstract manifold, and let $\sigma\colon U \to M$ be a positive chart on M. Let $\omega \in A^m(M)$ be a continuous m-form on M. Let $D \subset U$ be a domain of integration, and let $R = \sigma(D)$. Then we define

$$\int_R \omega = \int_D \phi \circ \sigma\, dV \tag{9.2}$$

where ϕ is the function determined in (9.1).

We shall now see that the definition is independent of the choice of chart. The point is that when coordinates are changed, the function ϕ that represents ω changes by a factor that exactly compensates for the Jacobian determinant resulting from the change. The Jacobian determinant has simply been built into the object that we integrate. This is the central idea behind the theory.

Theorem 9.3. *The integral $\int_R \omega$ is independent of the choice of positive chart σ. More precisely, if R is contained in the image of some other chart, then the integral defined in* (9.2) *has the same value for the two charts.*

Proof. The statement is the special case of the identity mapping $M \to M$ in the following more general result. □

Lemma 9.3. *Let M, $\tilde{M}$ be oriented abstract manifolds, and let $g\colon \tilde{M} \to M$ be an orientation preserving diffeomorphism. Let $\sigma\colon U \to M$ and $\tilde{\sigma}\colon \tilde{U} \to \tilde{M}$ be charts, and let $D \subset U$ and $\tilde{D} \subset \tilde{U}$ be domains of integration. Assume that $g(\tilde{R}) = R$, where $R = \sigma(D)$ and $\tilde{R} = \tilde{\sigma}(\tilde{D})$.*

Furthermore, let ω be a continuous m-form on M and let $\tilde{\omega} = g^*\omega$ be its pull back by g (defined as in Definition 8.5). Then

$$\int_{\tilde{R}} \tilde{\omega} = \int_R \omega.$$

Proof. The open subset $\tilde{\sigma}(\tilde{U}) \cap g^{-1}(\sigma(U))$ of $\tilde{\sigma}(\tilde{U})$ contains $\tilde{R}$. Hence we can replace $\tilde{\sigma}$ by its restriction to the preimage in $\tilde{U}$ of this set with no effect on the integral over $\tilde{R}$. That is, we may assume that $\tilde{\sigma}(\tilde{U}) \subset g^{-1}(\sigma(U))$, or equivalently, $g(\tilde{\sigma}(\tilde{U})) \subset \sigma(U)$.

Similarly, since $g(\tilde{\sigma}(\tilde{U}))$ contains R, we can replace σ by its restriction to the preimage in U of this set with no effect on the integral over R. In other words, we may in fact assume that $g(\tilde{\sigma}(\tilde{U})) = \sigma(U)$.

Let Φ denote the coordinate expression for g,

$$\Phi = \sigma^{-1} \circ g \circ \tilde{\sigma} \colon \tilde{U} \to U.$$

It is a diffeomorphism of $\tilde{U}$ onto U. The claim is that

$$\int_{\tilde{D}} \tilde{\phi} \circ \tilde{\sigma}\, dV = \int_D \phi \circ \sigma\, dV$$

where $\omega = \phi\, dx_1 \wedge \cdots \wedge x_m$ and $\tilde{\omega} = \tilde{\phi}\, d\tilde{x}_1 \wedge \cdots \wedge \tilde{x}_m$ are the expressions for ω and $\tilde{\omega}$.

It follows from Lemma 8.4.2 that

$$g^*(dx_1 \wedge \cdots \wedge dx_m) = \det(dg)\, d\tilde{x}_1 \wedge \cdots \wedge d\tilde{x}_m, \tag{9.3}$$

where $\det(dg_p)$ is the determinant of the matrix for dg_p with respect to the standard bases for the tangent spaces. In other words, $\det(dg_p)$ is the determinant of the Jacobian matrix $D\Phi$ of the coordinate expression Φ. Applying (8.9) and (9.3) we conclude that

$$\tilde{\omega} = g^*\omega = g^*(\phi\, dx_1 \wedge \cdots \wedge dx_m) = (\phi \circ g) \det(D\Phi)\, d\tilde{x}_1 \wedge \cdots \wedge d\tilde{x}_m.$$

It follows that

$$\tilde{\phi} = \det(D\Phi)\, (\phi \circ g). \tag{9.4}$$

Since $D = \Phi(\tilde{D})$ we obtain from Theorem 9.2

$$\int_D \phi \circ \sigma\, dV = \int_{\tilde{D}} \phi \circ \sigma \circ \Phi\, |\det(D\Phi)|\, dV. \tag{9.5}$$

Observe that $\sigma \circ \Phi = g \circ \tilde{\sigma}$, and that $\det(D\Phi)$ is positive since g preserves orientation. It now follows from (9.4) that the expression (9.5) equals

$$\int_D \tilde{\phi} \circ \tilde{\sigma}\, dV,$$

and the lemma is proved. $\square$

Example 9.3 Let $\gamma\colon I \to \mathbb{R}^2$ be an embedded parametrized curve, and let ω be a continuous 1-form on $\mathbb{R}^2$, say $\omega = f_1(x_1, x_2)\, dx_1 + f_2(x_1, x_2)\, dx_2$. Let $[a, b] \subset I$ and $R = \gamma([a, b])$. Then

$$\int_R \omega = \int_a^b F(\gamma(u)) \cdot \gamma'(u)\, du \tag{9.6}$$

where $F(x_1, x_2) = (f_1(x_1, x_2), f_2(x_1, x_2))$. This is seen as follows. The basis for the 1-dimensional tangent space of γ at $p = \gamma(u)$ is $\gamma'(u)$, and the dual basis vector $du(p)$ is defined by $du(p)(\gamma'(u)) = 1$. On the other hand, $\omega_p(v) = F(p) \cdot v$ for $v \in \mathbb{R}^2$, and in particular $\omega_p(\gamma'(u)) = F(p) \cdot \gamma'(u)$. Hence $\omega|_{\gamma(I)} = \phi\, du$ with $\phi(p) = F(p) \cdot \gamma'(u)$. Now (9.6) follows from Definition 9.3. The expression (9.6) is called the *line integral* of ω along R.

9.4 Integration on a manifold

We shall generalize the definition of $\int_R \omega$ to the case where R is not necessarily contained inside a single chart.

Definition 9.4.1. A compact subset R of an abstract manifold M is said to be a *domain of integration*, if the boundary ∂R is a null set.

For example, if $\Omega \subset M$ is a domain with smooth boundary, such that the closure $R = \bar{\Omega}$ is compact, then this closure is a domain of integration, by Corollary 4.6 and Example 9.1.3.

It is easily seen that the union $R_1 \cup R_2$, the intersection $R_1 \cap R_2$ and the set difference $R_1 \setminus R_2^\circ$, of two domains of integration $R_1, R_2 \subset M$ is again a domain of integration (in all three cases, the boundary is contained in the union of the boundaries of R_1 and R_2).

In particular, if $R \subset M$ is a domain of integration, and if $\sigma: U \to M$ is a chart with another domain of integration $D \subset U$, then $\sigma(D) \cap R$ is a domain of integration inside $\sigma(U)$, to which we can apply the preceding definition of the integral. We shall define the integral over R as a sum of such integrals, for a collection of charts that cover R. In order to combine these integrals defined by different charts, we employ a partition of unity, see Section 5.4. We therefore assume that M has a locally finite atlas.

We choose around each point $p \in M$ a chart $\sigma: U \to M$ and a domain of integration $D \subset U$ such that p is contained in the image $\sigma(D^\circ)$ of the interior. The collection of all these charts is denoted $\sigma_\alpha: U_\alpha \to M$, where A is the index set. The open sets $\sigma_\alpha(D_\alpha^\circ)$ for all $\alpha \in A$ cover M. We choose a partition of unity $(\varphi_\alpha)_{\alpha \in A}$ such that each function φ_α has support inside the corresponding set $\sigma(D_\alpha^\circ)$ (see Theorem 5.5). Notice that the collection of the supports of the functions φ_α is locally finite, hence only finitely many of these functions are non-zero on the compact set R.

Definition 9.4.2. Let ω and be a continuous m-form on M, and let $R \subset M$ be a domain of integration. We define

$$\int_R \omega = \sum_{\alpha \in A} \int_{R \cap \sigma_\alpha(D_\alpha)} \varphi_\alpha \omega,$$

where the integrals on the right are defined with respect to the charts σ_α, as in Definition 9.3. The sum is finite, since only finitely many φ_α are non-zero on R.

Theorem 9.4. *The definition given above is independent of the choice of the charts σ_α and the subsequent choice of a partition of unity.*

Proof. Indeed, if $\tilde{\varphi}_{\tilde{\alpha}}$, where $\tilde{\alpha} \in \tilde{A}$, is a different partition of unity relative to a different atlas, then

$$\sum_{\alpha\in A}\int_{R\cap\sigma_\alpha(D_\alpha)}\varphi_\alpha\,\omega=\sum_{\alpha\in A}\int_{R\cap\sigma_\alpha(D_\alpha)}(\sum_{\tilde\alpha\in\tilde A}\tilde\varphi_{\tilde\alpha})\varphi_\alpha\,\omega$$
$$=\sum_{\tilde\alpha\in\tilde A}\sum_{\alpha\in A}\int_{R\cap\sigma_\alpha(D_\alpha)}\varphi_\alpha\tilde\varphi_{\tilde\alpha}\,\omega$$

and since $\varphi_\alpha\tilde\varphi_{\tilde\alpha}\,\omega$ is supported inside the intersection $R\cap\sigma_\alpha(D_\alpha)\cap\sigma_{\tilde\alpha}(D_{\tilde\alpha})$, the latter expression equals

$$\sum_{\tilde\alpha\in\tilde A}\sum_{\alpha\in A}\int_{R\cap\sigma_\alpha(D_\alpha)\cap\tilde\sigma_{\tilde\alpha}(D_{\tilde\alpha})}\varphi_\alpha\tilde\varphi_{\tilde\alpha}\,\omega$$

The integral over $R\cap\sigma_\alpha(D_\alpha)\cap\tilde\sigma_{\tilde\alpha}(D_{\tilde\alpha})$ has the same value for the two charts by Theorem 9.3. Hence the last expression above is symmetric with respect to the partitions indexed by A and $\tilde A$, and hence the original sum has the same value if the partition is replaced by the other one. □

It is easily seen that $\int_M\omega$ depends linearly on ω. Moreover, in analogy with Lemma 9.3:

Lemma 9.4. *Let $g\colon\tilde M\to M$ be an orientation preserving diffeomorphism, and $\tilde R\subset\tilde M$ a domain of integration. Then $R=g(\tilde R)\subset M$ is a domain of integration. Furthermore, let ω be a continuous m-form with pull back $\tilde\omega=g^*\omega$. Then*

$$\int_{\tilde R}\tilde\omega=\int_R\omega.$$

Proof. Let $(\varphi_\alpha)_{\alpha\in A}$ be a partition of unity on M, as in Definition 9.4.2, and put $\psi_\alpha=\varphi_\alpha\circ g\in C^\infty(\tilde M)$. Since g is a diffeomorphism, this is a partition of unity on $\tilde M$. It follows from (8.9), that $g^*(\varphi_\alpha\omega)=\psi_\alpha\tilde\omega$. The proof is now straightforward from Definition 9.4.2 and Lemma 9.3. □

The case of a 0-dimensional manifold needs special interpretation. Recall that a 0-dimensional abstract manifold is a discrete set M. By definition, an orientation of M assigns $+$ or $-$ to each point in M, it can thus be viewed as a function $o\colon M\to\{+,-\}$. A subset R is compact if and only if it is finite. A 0-form on M is a function $f\colon M\to\mathbb{R}$, and the interpretation of its integral over a finite set D is that it is the signed sum

$$\sum_{p\in D}o(p)f(p).$$

This interpretation may seem peculiar, but it is extremely convenient (for example, in Example 9.6.1 below).

9.5 A useful formula

The definition of $\int_R \omega$ by means of a partition of unity is almost impossible to apply, when it comes to explicit computations, because it is very difficult to write down explicit partitions of unity. The following theorem, of which the proof is omitted, can be used instead. The idea is to chop up R in finitely many pieces and treat each of them separately. The pieces are chosen such that they are convenient for computations.

Theorem 9.5. *Let M be an abstract oriented manifold, let ω be a continuous m-form on M, and let $R \subset M$ be a domain of integration.*

a) *Assume that $R = R_1 \cup \dots \cup R_n$ where each set $R_i \subset M$ is a domain of integration, such that all the interiors R_i° are disjoint from each other. Then*

$$\int_R \omega = \sum_{i=1}^{n} \int_{R_i} \omega.$$

b) *Let $i = 1, \dots, n$ and assume that there exist an open set $U_i \subset \mathbb{R}^m$, a smooth map $g_i\colon U_i \to M$ and a domain of integration $D_i \subset U_i$ such that $R_i = g_i(D_i)$, and such that the restriction $g_i|_{D_i^\circ}$ is a positive chart. Then*

$$\int_{R_i} \omega = \int_{D_i} g_i^* \omega.$$

Let $\phi_i\colon U_i \to \mathbb{R}$ be the function determined by $g_i^*\omega = \phi_i \, du_1 \wedge \dots \wedge du_m$, with $u_1, \dots, u_m$ the coordinates of $\mathbb{R}^m$. Then the last formula above reads

$$\int_{R_i} \omega = \int_{D_i} \phi_i \, dV.$$

Notice that the requirement on the map g_i is just that its restriction to the interior of D_i is a chart. If g_i itself is a chart, $\sigma = g_i$, then the present function ϕ_i is identical with the function $\phi \circ \sigma$ in (9.2), and the formula for $\int_{R_i} \omega$ by means of $g_i^*\omega$ is identical with the one in Definition 9.3.

For example, the unit sphere S^2 is covered in this fashion by a single map $g\colon D \to S^2$ of spherical coordinates

$$g(u, v) = (\cos u \cos v, \cos u \sin v, \sin u)$$

with $D = [-\pi/2, \pi/2] \times [-\pi, \pi]$, and thus we can compute the integral of a 2-form over S^2 by means of its pull-back by spherical coordinates, in spite of the fact that this is only a chart on a part of the sphere (the point being, of course, that the remaining part is a null set).

9.6 Stokes' theorem

Stokes' theorem is the central result in the theory of integration on manifolds. It gives the relation between exterior differentiation and integration, and it generalizes the fundamental theorem of calculus.

Let M be an m-dimensional oriented abstract manifold, and let $\Omega \subset M$ be a domain with smooth boundary. The boundary $\partial\Omega$ (which may be empty) is equipped with the induced orientation, see Section 4.8. For example, Ω could be the inside of a compact connected manifold in $\mathbb{R}^n$, as in the Jordan-Brouwer theorem, Theorem 5.10.

As mentioned below Definition 9.4.1, if $\bar{\Omega}$ is compact, then $\bar{\Omega}$ is a domain of integration, and hence the integral $\int_{\bar{\Omega}} \omega$ is defined for all continuous m-forms on M. The requirement that $\bar{\Omega}$ is compact can be relaxed as follows. Observe that if R is a domain of integration, then so is the intersection $\bar{\Omega} \cap R$, since the boundary of each set is a null set. Let now ω be a continuous m-form. The *support* $\operatorname{supp}\omega$ of ω is the closure of the set $\{p \in M \mid \omega(p) \neq 0\}$. Assume that $\bar{\Omega} \cap \operatorname{supp}\omega$ is compact, and choose a domain of integration $R \subset M$ which contains this set. We can then define the integral of ω over Ω (or $\bar{\Omega}$) by

$$\int_\Omega \omega = \int_{\bar{\Omega}\cap R} \omega,$$

the point being that the right hand side is independent of the choice of R.

Theorem 9.6. Stokes' theorem. *Let ω be a differential $m-1$-form on M, and assume that $\bar{\Omega} \cap \operatorname{supp}\omega$ is compact. Then*

$$\int_\Omega d\omega = \int_{\partial\Omega} \omega. \tag{9.7}$$

The right-hand side is interpreted as 0 if $\partial\Omega$ is empty.

The result is remarkable, because it shows that by knowing ω only on the boundary of Ω, we can predict a property of it on the interior, namely the total integral of its derivative.

Proof. Let $R \subset M$ be a domain of integration containing $\bar{\Omega} \cap \operatorname{supp}\omega$. Then this set also contains $\bar{\Omega} \cap \operatorname{supp} d\omega$. Hence by definition

$$\int_\Omega d\omega = \int_{\bar{\Omega}\cap R} d\omega \quad \text{and} \quad \int_{\partial\Omega} \omega = \int_{\partial\Omega\cap R} \omega.$$

Choose around each point $p \in M$ a chart $\sigma\colon U \to M$ and an m-dimensional rectangle $D \subset U$ such that $p \in \sigma(D^\circ)$, where D° is the interior of D. In addition we assume that these charts are chosen so that if $p \in \partial\Omega$ then $\sigma(U) \cap \Omega = \sigma(U^+)$ where $U^+ = \{x \in U \mid u_m > 0\}$ (see Theorem 4.6), if $p \in \Omega$ then $\sigma(U) \subset \Omega$ and if $p \in M \setminus \bar{\Omega}$ then $\sigma(U) \subset M \setminus \bar{\Omega}$.

We may now assume that ω is supported inside $\sigma(D^\circ)$ for one of these charts. For if the result has been established in this generality, we can apply it to $\varphi_\alpha\omega$ for each element φ_α in a partition of unity as in Definition 9.4.2. Since the set R is compact, only finitely many φ_α's are non-zero on it, and as both sides of (9.7) depend linearly on ω, the general result then follows.

The equation (9.7) clearly holds if $\sigma(U) \subset M \setminus \bar{\Omega}$, since then both sides are 0. This leaves the other two cases, $\sigma(U) \cap \Omega = \sigma(U^+)$ and $\sigma(U) \subset \Omega$, to be checked.

We may also assume that

$$\omega = f\, dx_1 \wedge \cdots \wedge dx_{j-1} \wedge dx_{j+1} \wedge \cdots \wedge dx_m$$

for some j between 1 and m, because in general ω will be a sum of such forms.

Since $df = \sum_{i=1}^m \frac{\partial f}{\partial x_i}\, dx_i$ (see Lemma 8.2.2), it then follows that

$$d\omega = (-1)^{j-1} \frac{\partial f}{\partial x_j}\, dx_1 \wedge \cdots \wedge dx_m,$$

where the sign appears because dx_j has been moved past the 1-forms from dx_1 up to dx_{j-1}.

Assume first that $j < m$. We will prove that then

$$\int_\Omega d\omega = \int_{\partial\Omega} \omega = 0.$$

By definition

$$\int_\Omega d\omega = (-1)^{j-1} \int_\Omega \frac{\partial f}{\partial x_j}\, dx_1 \wedge \cdots \wedge dx_m = (-1)^{j-1} \int \frac{\partial (f \circ \sigma)}{\partial u_j}\, dV,$$

where the last integration takes place over the set $\{x \in D \mid \sigma(x) \in \bar{\Omega}\}$, that is, over $D \cap \overline{U^+}$ or D, in the two cases.

In the integral over the rectangle $D \cap U^+$ or D, we can freely interchange the order of integrations over $u_1, \dots, u_m$. Let us take the integral over u_j first (innermost), and let us denote its limits by a and b. Notice that since $j < m$ these are the limits for the u_j variable both in $D \cap U^+$ and D. Now by the fundamental theorem of calculus

$$\begin{aligned}\int_a^b \frac{\partial (f \circ \sigma)}{\partial u_j}\, du_j = f(\sigma(&u_1, \dots, u_{j-1}, b, u_{j+1}, \dots, u_m)) \\ &- f(\sigma(u_1, \dots, u_{j-1}, a, u_{j+1}, \dots, u_m)).\end{aligned}$$

However, since ω is supported in $\sigma(D^\circ)$, it follows that these values are zero for all $u_1, \dots u_{j-1}, u_{j+1}, \dots u_m$, and hence $\int_\Omega d\omega = 0$ as claimed.

On the other hand, if $\sigma(U)\cap\Omega = \sigma(U^+)$, then $x_m = 0$ along the boundary, and it follows immediately that dx_m, and hence also ω, restricts to zero on $\partial\Omega$. Therefore also $\int_{\partial\Omega}\omega = 0$. The same conclusion holds trivially in the other case, $\sigma(U) \subset \Omega$.

Next we assume $j = m$. Again we obtain 0 on both sides of (9.7) in the case $\sigma(U) \subset \Omega$, and we therefore assume the other case. We will prove that then

$$\int_\Omega d\omega = \int_{\partial\Omega} \omega = (-1)^m \int f(\sigma(u_1, \dots, u_{m-1}, 0))\, du_1 \dots du_{m-1},$$

where the integral runs over the set of $(u_1, \dots, u_{m-1}) \in \mathbb{R}^{m-1}$ for which $\sigma(u_1, \dots, u_{m-1}, 0) \in \partial\Omega$

Following the preceding computation of $\int_\Omega d\omega$, we take the integral over $u_j = u_m$ first. This time, however, the lower limit a is replaced by the value 0 of x_m on the boundary, and the previous conclusion fails, that f vanishes here. Instead we obtain the value $f(\sigma(u_1, \dots, u_{m-1}, 0))$, with a minus in front because it is the lower limit in the integral. Recall that there was a factor $(-1)^{j-1} = (-1)^{m-1}$ in front. Performing the integral over the other variables as well, we thus obtain the desired integral expression

$$(-1)^m \int f(\sigma(u_1, \dots, u_{m-1}, 0))\, du_1 \dots du_{m-1}$$

for $\int_\Omega d\omega$.

On the other hand, the integral $\int_{\partial\Omega}\omega$ can be computed by means of the restricted chart $\sigma|_{U\cap\mathbb{R}^{m-1}}$ on $\partial\Omega$. However, we have to keep track of the orientation of this chart. By definition (see Section 4.8), the orientation of the basis

$$d\sigma(e_1), \dots, d\sigma(e_{m-1})$$

for $T_p\partial M$ is the same as the orientation of the basis

$$-d\sigma(e_m), d\sigma(e_1), \dots, d\sigma(e_{m-1})$$

for T_pM. The orientation of the latter is $(-1)^m$, because σ is a positive chart. It follows that

$$\int_{\partial\Omega} \omega = (-1)^m \int f(\sigma(u_1, \dots, u_{m-1}, 0))\, du_1 \dots du_{m-1},$$

as claimed. $\square$

9.7 Examples from vector calculus

In the following examples we always use the standard coordinates on the Euclidean space $\mathbb{R}^n$.

Example 9.6.1. Let $M = \mathbb{R}$, then $m = 1$ and ω is a 0-form, that is, a function $f \in C^\infty(\mathbb{R})$. The differential of f is the 1-form $df = f'(x)dx$ (see Example 8.6.1). Let $\Omega =]a, b[$, then the boundary $\partial\Omega$ consists of the two points a and b, oriented by -1 and $+1$, respectively (see Example 4.7.3). Hence

$$\int_\Omega d\omega = \int_a^b df = \int_a^b f'(x)\,dx$$

and

$$\int_{\partial\Omega} \omega = f(b) - f(a)$$

Thus we see that in this case Stokes' theorem reduces to the fundamental theorem of calculus.

Example 9.6.2. Let $M = \mathbb{R}^2$, then $m = 2$ and ω is a 1-form. The boundary of the open set Ω is a union of smooth curves, we assume for simplicity it is a single simple closed curve. Write $\omega = f(x, y)\,dx + g(x, y)\,dy$, then (see Example 8.6.2)

$$d\omega = (-\frac{\partial f}{\partial y} + \frac{\partial g}{\partial x})\,dx \wedge dy$$

and hence

$$\int_\Omega d\omega = \int_\Omega -\frac{\partial f}{\partial y} + \frac{\partial g}{\partial x}\,dA.$$

On the other hand, the integral $\int_{\partial\Omega} \omega$ over the boundary can be computed as follows. Assume $\gamma\colon [0, T] \to \partial\Omega$ is the boundary curve, with end points $\gamma(0) = \gamma(T)$ (and no other self intersections). Theorem 9.5 can be applied with $D_1 = [0, T]$ and $g_1 = \gamma$. Write $\gamma(t) = (x(t), y(t))$, then by definition

$$g_1^*\omega = (f \circ \gamma)\,d(x \circ \gamma) + (g \circ \gamma)\,d(y \circ \gamma),$$

and $\int_{\partial\Omega} \omega$ is the line integral

$$\int_{\partial\Omega} \omega = \int_\gamma f(x, y)dx + g(x, y)dy.$$

We thus see that in this case Stokes' theorem reduces to the classical *Green's theorem* (which is equivalent with the divergence theorem for the plane):

$$\int_\Omega -\frac{\partial f}{\partial y} + \frac{\partial g}{\partial x}\,dA = \int_\gamma f(x, y)dx + g(x, y)dy.$$

The orientation of γ is determined as follows. If $\mathbf{n}$ and $\mathbf{t}$ are the outward normal vector and the positive unit tangent vector, respectively, in a given point of γ, then $(\mathbf{n}, \mathbf{t})$ should be positively oriented, that is, $\mathbf{t} = \hat{\mathbf{n}}$, which is exactly the standard counter clockwise orientation of a closed curve in $\mathbb{R}^2$.

Example 9.6.3. Let $M = \mathbb{R}^3$, and let Ω be a domain with smooth boundary. The boundary $\partial\Omega$ is then a surface $\mathcal{S}$ in $\mathbb{R}^3$. Let ω be a 2-form on M:

$$\omega = f(x, y, z)\, dy \wedge dz + g(x, y, z)\, dz \wedge dx + h(x, y, z)\, dx \wedge dy.$$

Then (see Example 8.6.3)

$$d\omega = (\frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} + \frac{\partial h}{\partial z})\, dx \wedge dy \wedge dz$$

and hence

$$\int_\Omega d\omega = \int_\Omega \operatorname{div}(f, g, h)\, dV,$$

where $\operatorname{div}(f, g, h) = \frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} + \frac{\partial h}{\partial z}$ is the divergence of the vector field (f, g, h).

On the other hand, the integral $\int_{\partial\Omega} \omega$ over the boundary $\partial\Omega = \mathcal{S}$ can be computed as follows. Suppose $D_1, \dots, D_n$ and $g_i\colon D_i \to S$ are as in Proposition 9.5 for the manifold $\mathcal{S}$. Then $\int_{\mathcal{S}} \omega = \sum_{i=1}^n \int_{D_i} g_i^* \omega$. Let $\sigma(u, v) = g_i(u, v)$ be one of the functions g_i. Then

$$\sigma^*\omega = (f\circ\sigma)\, d(y\circ\sigma)\wedge d(z\circ\sigma) + (g\circ\sigma)\, d(z\circ\sigma)\wedge d(x\circ\sigma) + (f\circ\sigma)\, d(x\circ\sigma)\wedge d(y\circ\sigma).$$

Furthermore

$$\begin{aligned} d(y \circ \sigma) \wedge d(z \circ \sigma) &= (\frac{\partial \sigma_2}{\partial u}\, du + \frac{\partial \sigma_2}{\partial v}\, dv) \wedge (\frac{\partial \sigma_3}{\partial u}\, du + \frac{\partial \sigma_3}{\partial v}\, dv) \\ &= (\frac{\partial \sigma_2}{\partial u}\frac{\partial \sigma_3}{\partial v} - \frac{\partial \sigma_2}{\partial v}\frac{\partial \sigma_3}{\partial u})\, du \wedge dv \end{aligned}$$

and similarly

$$\begin{aligned} d(z \circ \sigma) \wedge d(x \circ \sigma) &= (\frac{\partial \sigma_3}{\partial u}\frac{\partial \sigma_1}{\partial v} - \frac{\partial \sigma_1}{\partial v}\frac{\partial \sigma_3}{\partial u})\, du \wedge dv \\ d(x \circ \sigma) \wedge d(y \circ \sigma) &= (\frac{\partial \sigma_1}{\partial u}\frac{\partial \sigma_2}{\partial v} - \frac{\partial \sigma_2}{\partial v}\frac{\partial \sigma_1}{\partial u})\, du \wedge dv. \end{aligned}$$

Notice that the three expressions in front of $du\wedge dv$ are exactly the coordinates of the normal vector $\sigma'_u \times \sigma'_v$. Thus we see that

$$\sigma^*\omega = (f \circ \sigma, g \circ \sigma, h \circ \sigma) \cdot (\sigma'_u \times \sigma'_v)\, du \wedge dv.$$

Let $\mathbf{N}(u, v)$ denote the outward unit normal vector in $\sigma(u, v)$, then $\sigma'_u \times \sigma'_v = \|\sigma'_u \times \sigma'_v\|\, \mathbf{N}$ and we see that $\int_D \sigma^*\omega$ is the surface integral of the function

$(f, g, h) \cdot \mathbf{N}$ over the image of $\sigma = g_i$ (recall that the surface integral of a function includes a factor $\|\sigma'_u \times \sigma'_v\| = (EG - F^2)^{1/2}$). Summing over $i = 1, \dots, n$ we finally obtain the surface integral

$$\int_{\mathcal{S}} \omega = \int_{\mathcal{S}} (f, g, h) \cdot \mathbf{N} \, dA.$$

We conclude that in this case Stokes' theorem reduces to the *divergence theorem* (sometimes also called the Gauss theorem):

$$\int_{\Omega} \operatorname{div}(f, g, h) \, dV = \int_{\partial\Omega} (f, g, h) \cdot \mathbf{N} \, dA.$$

Example 9.6.4. Let $M = \mathcal{S}$ be an oriented smooth surface in $\mathbb{R}^3$, and let $\Omega \subset M$ be a domain with smooth boundary. Let ω be a 1-form on M, and assume ω is the restriction of a smooth 1-form

$$f(x, y, z) \, dx + g(x, y, z) \, dy + h(x, y, z) \, dz$$

defined on an open neighborhood of M in $\mathbb{R}^3$. A similar analysis as in the previous example shows that in this case Stokes' theorem reduces to the original theorem of Stokes, which is

$$\int_{\Omega} (\operatorname{curl} F \cdot \mathbf{N}) \, dA = \int_{\partial\Omega} f \, dx + g \, dy + h \, dz.$$

where

$$\operatorname{curl} F = (\frac{\partial h}{\partial y} - \frac{\partial g}{\partial z}, \frac{\partial f}{\partial z} - \frac{\partial h}{\partial x}, \frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}).$$

The preceding examples show that in addition to the fundamental theorem of calculus, Stokes' theorem also generalizes the three main theorems of classical vector calculus.

Index