Notation

We will use the following abbreviations: if $\mathbb{R}^n$ has coordinates x^i, we will use $\partial_i = \partial_{x^i} = \frac{\partial}{\partial x^i}$ to denote the differentiation with respect to x^i. In addition, we will use Einstein's summation notation, in which the sum symbol is not written if two indices appear, one as a lower index and one as an upper index, e.g. the directional derivative is

$$\mathrm{D}_A\varphi = \partial_i\varphi \cdot A^i.$$

In such a situation, the summation is implicit (and thus, it is necessary to state explicitly if the summation is not intended). We believe that this highly improves readability of formulas.

We will also use a notation $\varphi\colon V \dashrightarrow W$ that means that φ is defined on an open subset of V, denoted $\operatorname{dom}\varphi$.

1. Analysis in a vector space

A derivative of a mapping $\varphi\colon V \dashrightarrow W$ at a point x is the best linear approximation of φ at x and as such is a linear map $V \to W$ (more precisely, it is the linear part of the best affine approximation). We will use this geometric object in general considerations and, only in computations, we will use the matrix $(\partial_j\varphi^i)$ representing this linear map in coordinates on V and W.

Definition 1.1. Given a mapping $\varphi\colon V \dashrightarrow W$, its *derivative* (or differential) at x is a linear map $\mathrm{D}|_x\varphi\colon V \to W$ that satisfies $\varphi(x+v) = \varphi(x) + \mathrm{D}|_x\varphi(v) + o(v)$.

Remark. A function $h(v)$ lies in $o(v)$ if $\lim_{v\to 0}\frac{|h(v)|}{|v|} = 0$. Similarly, $h(v)$ lies in $O(v)$ if $\frac{|h(v)|}{|v|}$ is bounded in some neighbourhood of 0. Since any pair of norms is comparable, the notions of $O(v)$ and $o(v)$ do not depend on a particular choice of a norm. Applying to the $|-|_\infty$ norm on W, $h(v) \in o(v)$ if and only if each component $h^i(v) \in o(v)$.

Example 1.2. Functions of one variable, i.e. $V = \mathbb{R}$, $W = \mathbb{R}$. Then $\mathrm{D}|_t\varphi$ is a (1×1)-matrix with the sole entry $\varphi'(t)$, the usual derivative, i.e. $\mathrm{D}|_t\varphi = (\varphi'(t))$. We will usually identify these two objects.

Example 1.3. Functions of several variables, i.e. $V = \mathbb{R}^n$, $W = \mathbb{R}$. Then $\mathrm{D}|_x\varphi$ is a $(1 \times n)$-matrix, or a linear form. It associates to each vector the rate of growth of φ along this vector. The gradient of φ at x is then the direction of the largest growth (geometrically, it is perpendicular to the level set $\varphi^{-1}(\varphi(x))$ and of some particular magnitude) and as such depends on a scalar product: $\mathrm{D}|_x\varphi(v) = \langle \operatorname{grad}_x\varphi, v\rangle$. In orthonormal coordinates, it corresponds to transposing the linear form $\mathrm{D}|_x\varphi$. For these reasons, we will use mostly the more geometric $\mathrm{D}|_x\varphi$.

Example 1.4. Paths, i.e. $V = \mathbb{R}$, $W = \mathbb{R}^n$. Then $\mathrm{D}|_t\varphi$ is an $(n \times 1)$-matrix. There is a natural identification between matrices of this type and vectors (since $\operatorname{Hom}(\mathbb{R},\mathbb{R}^n) \cong \mathbb{R}^n$, given by evaluation at 1). Denoting $\varphi'(t) = \mathrm{D}|_t\varphi(1)$, we thus obtain $\mathrm{D}|_t\varphi(\tau) = \tau\cdot\varphi'(t)$. This is a generalization of the first example. The geometric meaning of $\varphi'(t)$ is that of a tangent vector to φ at time t. It is thus natural to think of it as a vector based at $\varphi(t)$ – we will continue to develop such formalism later and reserve the notation $\varphi'(t)$ for the based version.

Lemma 1.5 (chain rule)**.** *Given two mappings $\varphi\colon V \dashrightarrow W$, $\psi\colon W \dashrightarrow X$, we have*

$$\mathrm{D}|_x(\psi\circ\varphi) = \mathrm{D}|_{\varphi(x)}\psi \circ \mathrm{D}|_x\varphi.$$

It will be useful to consider the derivatives at different points and organize them into a mapping

$$\mathrm{D}\varphi\colon V \dashrightarrow \operatorname{Hom}(V,W), \quad x \mapsto \mathrm{D}|_x\varphi.$$

We also denote $\mathrm{D}|_x\varphi(v) = \mathrm{D}_v|_x\varphi = \mathrm{D}_v\varphi(x)$ (the directional derivative of φ along v at x) and, thus, obtain a mapping $\mathrm{D}_v\varphi\colon V \dashrightarrow W$. It is called the *directional derivative* of φ along v. Since the derivative of a linear map ψ is the very same map ψ, the chain rule gives:

Corollary 1.6. *For any linear map ψ, we have $\mathrm{D}_v(\psi\circ\varphi) = \psi\circ\mathrm{D}_v\varphi$, i.e. linear maps commute with directional derivatives.*

Lemma 1.7 (Leibniz rule)**.** *Given two mappings* $\varphi\colon V \dashrightarrow W$*,* $g\colon V \dashrightarrow X$*, we have*

$$\mathrm{D}_v(\varphi\otimes\psi) = \mathrm{D}_v\varphi\otimes\psi + \varphi\otimes\mathrm{D}_v\psi.$$

The last two results together prove that for any bilinear map $\Phi\colon W\times X\to Y$, we get Leibniz rule:

$$\mathrm{D}_v\Phi(\varphi,\psi) = \Phi(\mathrm{D}_v\varphi,\psi) + \Phi(\varphi,\mathrm{D}_v\psi)$$

(since $\Phi(\varphi,\psi) = F(\varphi\otimes\psi)$ with $F\colon W\otimes X\to Y$ the linear map corresponding to Φ). Special cases are the multiplication $\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ (this gives the usual Lebiniz rule), multiplication by scalars $V\times\mathbb{R}\to V$, inner product $V\times V\to\mathbb{R}$, matrix multiplication $\mathrm{Mat}_{k\times n}\times\mathrm{Mat}_{n\times m}\to\mathrm{Mat}_{k\times m}$ etc.

Lemma 1.8 (symmetry of second derivatives)**.** *Assuming that* φ *is* C^2 *in a neighbourhood of* x*, we have*

$$\mathrm{D}_u\mathrm{D}_v\varphi(x) = \mathrm{D}_v\mathrm{D}_u\varphi(x).$$

The second derivative is a mapping

$$\mathrm{D}^2\varphi\colon V \dashrightarrow \mathrm{Hom}(V,\mathrm{Hom}(V,W)) \cong \mathrm{Hom}(V\otimes V, W).$$

As a reformulation, the previous lemma says that $\mathrm{D}^2\varphi\colon V\dashrightarrow \mathrm{Hom}(S^2V,W)$.

Changing coordinates. We have seen (or we know) that, with respect to any given bases on V and W, $(\mathrm{D}\varphi)^i_j = \partial_j\varphi^i$. We will now study the effect of changing coordinates (non-linearly!), i.e. we compose φ with maps (changes of coordinates) on both sides: writing $x=\chi(y)$ and $z=\varphi(x)$, we have a chain rule

$$\mathrm{D}|_y(\psi\circ\varphi\circ\chi) = \mathrm{D}|_z\psi\circ\mathrm{D}|_x\varphi\circ\mathrm{D}|_y\chi.$$

In words, the linear map $\mathrm{D}|_x\varphi$ gets replaced by an equivalent one, i.e. one modified by linear coordinate changes of both domain and codomain.

The second derivative is more complicated and it is useful to rewrite the above formula first in coordinates:

$$\partial_j(\psi\circ\varphi\circ\chi)^i(y) = \partial_l\psi^i(\varphi\chi(y))\cdot\partial_m\varphi^l(\chi(y))\cdot\partial_j\chi^m(y).$$

Now we can differentiate once more to get

$$\begin{aligned}\partial_k\partial_j(\psi\circ\varphi\circ\chi)^i(y) &= \partial_{l'}\partial_l\psi^i(\varphi\chi(y))\cdot\partial_{m'}\varphi^{l'}(\chi(y))\cdot\partial_k\chi^{m'}(y)\cdot\partial_m\varphi^l(\chi(y))\cdot\partial_j\chi^m(y)\\ &\quad+\partial_l\psi^i(\varphi\chi(y))\cdot\partial_n\partial_m\varphi^l(\chi(y))\cdot\partial_k\chi^n(y)\cdot\partial_j\chi^m(y)\\ &\quad+\partial_l\psi^i(\varphi\chi(y))\cdot\partial_m\varphi^l(\chi(y))\cdot\partial_k\partial_j\chi^m(y).\end{aligned}$$

The second term is the standard way of transforming a bilinear map, the existence of the first and the third term however makes it clear that the second derivative depends on coordinates – it is relatively simple to come up with an example where $\mathrm{D}^2|_x\varphi$ is zero and $\mathrm{D}^2|_y(\psi\circ\varphi\circ\chi)$ is non-zero. The invariance under coordinate changes holds in two special cases: ψ and χ linear (so that the second derivatives vanish) and $\mathrm{D}|_x\varphi = 0$ (more generally, the first non-vanishing derivative is invariant). To get an invariant object, it is necessary to take the Taylor polynomial, i.e. to include also the first derivative (and the value), but this results in an object very different from a symmetric bilinear form.

2. Bump functions

Lemma 2.1. *Let* $\varepsilon>0$*. There exists a smooth function* $\rho\colon\mathbb{R}^n\to[0,1]$ *such that* $\rho(0)=1$ *and* $\rho(x)=0$ *for* $|x|\geq\varepsilon$*.*

Proof. In the case $n=1$, using the smooth function λ of the next lemma, we set

$$\rho(x) = \lambda(\varepsilon+x)\lambda(\varepsilon-x)/\lambda(\varepsilon)^2 = e^{-\frac{2x^2}{\varepsilon(\varepsilon^2-x^2)}}$$

on $(-\varepsilon,\varepsilon)$ and zero otherwise – the equality is a straightforward calculation, giving $\rho(x)\in[0,1]$. For general n, we use the constructed function for $n=1$ (possibly for smaller ε) in the following way: $\rho(x^1,\ldots,x^n) = \rho(x^1)\cdots\rho(x^n)$. $\square$

Lemma 2.2. *The function*

$$\lambda(x) = \begin{cases} e^{-1/x} & x > 0 \\ 0 & x \le 0 \end{cases}$$

is smooth.

Proof. For any positive integer n, it is known that $e^x > x^n$ for $x \gg 0$. Thus, $e^{1/x} > (1/x)^n$ for $x > 0$ small or, in other words, $e^{-1/x} < x^n$ for $x > 0$ small. It is easy to see that, in such a case, the right-sided derivatives of $e^{-1/x}$ at zero vanish up to order $n-1$. $\square$

3. Derivations

Let $U \subseteq V$ be an open subset of a vector space and $x_0 \in U$. We say that a mapping (an "operator") $A\colon C^\infty U \to \mathbb{R}$ is a *derivation at* x_0 if it is $\mathbb{R}$-linear and satisfies the following "Leibniz rule at x_0"

$$A(f \cdot g) = Af \cdot g(x_0) + f(x_0) \cdot Ag.$$

An example is the differentiation along any vector at x_0, i.e. $A = \mathrm{D}_v|_{x_0}$. We will now show that this is the only example.

Theorem 3.1. *Let $A\colon C^\infty U \to \mathbb{R}$ be a derivation at x_0. Then there exists a unique vector $v \in V$ such that $Af = D_v|_{x_0} f$. In other words, there is an isomorphism between V and the space of derivations at x_0.*

Proof. To make our formulas simpler, we assume $x_0 = 0$. First we make a simple observation. From $A(1 \cdot 1) = A1 \cdot 1 + 1 \cdot A1$ we see that $A1 = 0$. Consequently, $Ac = c \cdot A1 = 0$ for any constant function c.

In the first part, we prove the statement for $U = V$ or, more generally, for functions that extend smoothly to V. Let $f \in C^\infty V$ and write

$$f(x) - f(0) = [f(t \cdot x)]_{t=0}^1 = \int_0^1 \partial_t(f(t \cdot x))\ \mathrm{d}t = \int_0^1 \partial_i f(t \cdot x) \cdot x^i\ \mathrm{d}t = \int_0^1 \partial_i f(t \cdot x)\ \mathrm{d}t \cdot x^i.$$

Denoting $g_i(x) = \int_0^1 \partial_i f(t \cdot x)\ \mathrm{d}t$, a smooth function on V with $g_i(0) = \partial_i f(0)$, we get

$$f = f(0) + g_i \cdot x^i.$$

Now we apply A to obtain

$$Af = \underbrace{A(f(0))}_{0} + Ag_i \cdot \underbrace{x^i(0)}_{0} + \underbrace{g_i(0)}_{\partial_i f(0)} \cdot Ax^i = \partial_i f(0) \cdot Ax^i = \mathrm{D}_v|_{x_0} f,$$

the derivative of f at $x_0 = 0$ along $v = (Ax^1, \dots, Ax^n)$.

Now we proceed to general open subset $U \subseteq V$. Since the first part applies to functions extendable to V, it will be our goal now to show that, in some sense, every function $f\colon U \to \mathbb{R}$ extends to V. Let $\lambda\colon V \to [0,1]$ be such that $\lambda(x_0) = 1$ and such that $\operatorname{supp}\lambda \subseteq U$. Then $\lambda \cdot f$ is clearly a smooth function that extends to V (by declaring it zero on $V \smallsetminus \operatorname{supp}\lambda$). According to the first part, we get $A\lambda = \mathrm{D}_v|_{x_0}\lambda = 0$, since x_0 is a maximum point of λ, and therefore

$$A(\lambda \cdot f) = A\lambda \cdot f(x_0) + \lambda(x_0) \cdot Af = Af.$$

As a special case, we get (this can be also seen directly)

$$\mathrm{D}_v|_{x_0}(\lambda \cdot f) = \mathrm{D}_v|_{x_0} f.$$

Since the left hand sides of the two equalities agree by the first part (again, $\lambda \cdot f$ extends to V), so do the right hand sides, i.e. $Af = \mathrm{D}_v|_{x_0} f$. $\square$

We say that a mapping (an "operator") $X\colon C^\infty U \to C^\infty U$ is a *derivation* if it is $\mathbb{R}$-linear and satisfies the following "Leibniz rule"

$$X(f \cdot g) = Xf \cdot g + f \cdot Xg.$$

Theorem 3.2. *Let $X\colon C^\infty U \to C^\infty U$ be a derivation. Then there exists a unique smooth map $v\colon U \to V$ such that $Xf(x) = \mathrm{D}_{v(x)}|_x f$.*

The map v should be thought of as a smooth vector field on U. We will be more precise on this matter later.

Proof. For any $x \in M$, the composition of X with the evaluation map at x (i.e. $\mathrm{ev}_x(f) = f(x)$),

$$X|_x \colon C^\infty U \xrightarrow{X} C^\infty U \xrightarrow{\mathrm{ev}_x} \mathbb{R},$$

is a derivation at x and, thus, there is a unique vector $v(x)$ such that $Xf(x) = X|_x f = \mathrm{D}_{v(x)}|_x f$. It remains to show that $x \mapsto v(x)$ is a *smooth* vector field. However, we have seen that

$$v(x) = (X|_x x^1, \ldots, X|_x x^n) = (Xx^1(x), \ldots, Xx^n(x)),$$

i.e. $v = (Xx^1, \ldots, Xx^n)$, and as such is smooth. □

4. Tangent map

We have seen that vectors are in bijection with derivations at x. It will be useful to think of the corresponding vector v as "based at x" and, for emphasis, we will denote it as a pair $A = (x, v)$ and use this based vector A interchangeably as a derivation, i.e. $Af = D_v|_x f$. We denote by $T_x V = \{x\} \times V$ the set of all vectors based at x, clearly a vector space isomorphic to V. It is called the *tangent space* of V at x and its elements, i.e. vectors based at x, the *tangent vectors*.

We will now rephrase the last theorem in terms of based vectors. We denote by $TU = U \times V$ the set of all vectors based at all points of U and call it the *tangent bundle* of U. Then a *vector field* is a map $X \colon U \to TU$ with values $X|_x \in T_x U$, i.e. $X|_x = (x, v(x))$. We will use Xf to denote the function $Xf(x) = X|_x f = \mathrm{D}_{v(x)}|_x f$. An important special case is the vector field ∂_i with $\partial_i|_x = (x, e_i)$ that gives the partial differentiation.

Thus, any tangent vector $A = (x, v) \in T_x U$ with coordinates $v = A^i \cdot e_i$ can be written as $A = A^i \cdot \partial_i|_x$ and any vector field can be written as $X = X^i \cdot \partial_i$, where X^i are now smooth functions. In addition to being simply the coordinate expression, these formulas also suggest how to take a derivative along A:

$$Af = (A^i \cdot \partial_i|_x) f = A^i \cdot \partial_i|_x f,$$

i.e. one multiplies the row of partial derivatives by the column of coordinates of A (or coordinate functions of a vector field X).

It is easy to see that, for a smooth map $\varphi \colon V \dashrightarrow W$ and a tangent vector $A \in T_x V$, the association $f \mapsto A(f \circ \varphi)$ is a derivation at $\varphi(x)$ and is thus given by a tangent vector from $T_{\varphi(x)} W$. Let us derive a formula for this vector. Writing $A = (x, v)$, we have

$$A(f \circ \varphi) = \mathrm{D}|_x (f \circ \varphi)(v) = \mathrm{D}|_{\varphi(x)} f \circ \mathrm{D}|_x \varphi(v) = \mathrm{D}_{\mathrm{D}|_x \varphi(v)}|_{\varphi(x)} f.$$

Denoting $\varphi_* A = (\varphi(x), \mathrm{D}|_x \varphi(v))$, we thus arrive at a formula

$$A(f \circ \varphi) = (\varphi_* A) f.$$

Note that A is based at x and its image $\varphi_* A$ is based at $\varphi(x)$; this should be in correspondence with one's geometric intuition. There results a map $\varphi_* \colon TU \to TW$ or, again, $\varphi_* \colon TV \dashrightarrow TW$, called the *tangent map* of φ. The chain rule for this kind of "derivative" is particularly simple: $(\psi \circ \varphi)_* = \psi_* \circ \varphi_*$ (one can prove this either directly from the usual chain rule or formally from the above formula).

Proposition 4.1. *For a smooth map $\varphi \colon V \dashrightarrow W$, the tangent map $\varphi_* \colon TV \dashrightarrow TW$ is also smooth.*

Proof. This is clear from the definition, since

$$\varphi_*(x, v) = (\varphi(x), \mathrm{D}|_x \varphi \cdot v)$$

and $\varphi \colon V \dashrightarrow W$, $\mathrm{D}\varphi \colon V \dashrightarrow \mathrm{Hom}(V, W)$ and the evaluation $\mathrm{Hom}(V, W) \times V \to W$ (the matrix multiplication) are smooth maps. □

5. Implicit function theorem and its applications

Theorem 5.1 (Implicit function theorem)**.** *Let $F\colon \mathbb{R}^n \times \mathbb{R}^k \dashrightarrow \mathbb{R}^k$ be a smooth map such that $F(a,b) = 0$ and such that $\mathrm{D}|_{(a,b)}F \in \mathrm{Mat}_{k\times(n+k)}$ has the right $(k \times k)$-block invertible. Then there exist neighbourhoods $U \ni a$, $V \ni b$ such that for each $x \in U$ there is a unique $y = \varphi(x) \in V$ with $F(x,y) = 0$. Moreover, the resulting map $\varphi\colon U \to V$ is smooth.*

Now we will be concerned with applications of the theorem. We start with an inverse function theorem.

Theorem 5.2 (Inverse function theorem)**.** *Let $\varphi\colon \mathbb{R}^n \dashrightarrow \mathbb{R}^n$ be a smooth map such that $\mathrm{D}|_a\varphi$ is regular. Then there exist neighbourhoods $U \ni a$, $V \ni \varphi(a)$ such that the restriction $\varphi\colon U \to V$ is invertible with a smooth inverse.*

In particular, $\varphi(a)$ lies in the interior of the image.

Proof. We set $F(x,y) = \varphi(y) - x$. This satisfies the assumptions of the previous theorem and thus has a unique solution $y \in V$ for each $x \in U$. Clearly, this y equals $\varphi^{-1}(x)$ and, in particular, $\varphi\colon V \cap \varphi^{-1}(U) \longrightarrow U$ is a smooth bijection with an inverse smooth by the previous theorem. □

Now, we generalize the previous theorem in two ways to maps between spaces of different dimensions.

Theorem 5.3 (Submersion theorem)**.** *Let $\varphi\colon \mathbb{R}^n \dashrightarrow \mathbb{R}^k$ be a smooth map such that $\mathrm{D}|_a\varphi$ is surjective. Then there exist neighbourhoods $U \ni a$, $V \ni \varphi(a)$ and a diffeomorphism ψ such that*

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, dashed, "\varphi"] & \mathbb{R}^k \\ \mathbb{R}^n \arrow[u, dashed, "\psi"] \arrow[ur, "\mathrm{pr}"'] & \end{tikzcd}$$

where $\mathrm{pr}(x^1,\dots,x^n) = (x^1,\dots,x^k)$. In other words, in the coordinates given by ψ, the map φ is the projection map.

Proof. We assume for simplicity that $\mathrm{D}|_a\varphi = \begin{pmatrix} A & B \end{pmatrix}$ has the left $(k\times k)$-block A invertible. Let $\chi\colon \mathbb{R}^n \to \mathbb{R}^n$ be the map

$$\chi\begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = \begin{pmatrix} \varphi^1(x^1,\dots,x^n) \\ \vdots \\ \varphi^k(x^1,\dots,x^n) \\ x^{k+1} \\ \vdots \\ x^n \end{pmatrix}.$$

Clearly, $\mathrm{D}|_a\chi = \begin{pmatrix} A & B \\ 0 & E \end{pmatrix}$ and is thus invertible. We set $\psi = \chi^{-1}$ according to the inverse function theorem. Since the diagram

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, dashed, "\varphi"] \arrow[d, dashed, "\chi"'] & \mathbb{R}^k \\ \mathbb{R}^n \arrow[ur, "\mathrm{pr}"'] & \end{tikzcd}$$

clearly commutes, the same is true for the diagram from the statement. □

Theorem 5.4 (Immersion theorem)**.** *Let $\varphi\colon \mathbb{R}^k \dashrightarrow \mathbb{R}^n$ be a smooth map such that $\mathrm{D}|_b\varphi$ is injective. Then there exist neighbourhoods $V \ni b$, $U \ni \varphi(a)$ and a diffeomorphism ψ such that*

$$\begin{tikzcd} \mathbb{R}^k \arrow[r, dashed, "\varphi"] \arrow[dr, "\mathrm{in}"'] & \mathbb{R}^n \arrow[d, dashed, "\psi"] \\ & \mathbb{R}^n \end{tikzcd}$$

where $\mathrm{in}(x^1,\dots,x^k)=(x^1,\dots,x^k,0,\dots,0)$. *In other words, in the coordinates given by* ψ*, the map* φ *is the inclusion map.*

Proof. We assume for simplicity that $\mathrm{D}|_b\varphi=\begin{pmatrix}A\\B\end{pmatrix}$ has the top $(k\times k)$-block A invertible. Let $\chi\colon\mathbb{R}^n\to\mathbb{R}^n$ be the map

$$\chi\begin{pmatrix}x^1\\\vdots\\x^n\end{pmatrix}=\begin{pmatrix}\varphi^1(x^1,\dots,x^k)\\\vdots\\\varphi^k(x^1,\dots,x^k)\\x^{k+1}+\varphi^{k+1}(x^1,\dots,x^k)\\\vdots\\x^n+\varphi^n(x^1,\dots,x^k)\end{pmatrix}.$$

Clearly, $\mathrm{D}|_a\chi=\begin{pmatrix}A&0\\B&E\end{pmatrix}$ and is thus invertible. We set $\psi=\chi^{-1}$ according to the inverse function theorem. Since the diagram

$$\begin{tikzcd}\mathbb{R}^k \arrow[r, dashed, "\varphi"] \arrow[dr, "\mathrm{in}"'] & \mathbb{R}^n \\ & \mathbb{R}^n \arrow[u, dashed, "\chi"']\end{tikzcd}$$

clearly commutes, the same is true for the diagram from the statement. $\square$

6. Submanifolds of a vector space

Definition 6.1. A subset $M\subseteq\mathbb{R}^n$ is a *smooth submanifold* of *dimension* m if, for every $x\in M$, there is a diffeomorphism $\psi\colon\mathbb{R}^n\dashrightarrow\mathbb{R}^n$ defined near x such that

$$\psi(M\cap\operatorname{dom}\psi)=\mathbb{R}^m\cap\operatorname{im}\psi,$$

where we understand $\mathbb{R}^m\subseteq\mathbb{R}^n$.

Clearly, we may replace $\mathbb{R}^m$ in the definition by any affine subspace of $\mathbb{R}^n$, since there is always an (affine) diffeomorphism that maps this subspace to $\mathbb{R}^m$ and we may compose the original diffeomorphism with the affine one and get a diffeomorphism from the definition.

Theorem 6.2. *Let* $F\colon\mathbb{R}^n\to\mathbb{R}^k$ *be a smooth map,* $b\in\mathbb{R}^k$ *and denote* $M=F^{-1}(b)$*. If* F *is a submersion at every point of* M *then* M *is a smooth submanifold of dimension* $n-k$*.*

Proof. Under the assumptions of the theorem, near every $x\in M$, there is a diffeomorphism $\varphi\colon\mathbb{R}^n\dashrightarrow\mathbb{R}^n$ such that $F=\mathrm{pr}\circ\varphi$. Therefore, $F^{-1}(b)=\varphi^{-1}(\mathrm{pr}^{-1}(b))$ and clearly

$$\mathrm{pr}^{-1}(b)=\{(b^1,\dots,b^k,x^{k+1},\dots,x^n)\mid x^{k+1},\dots,x^n\in\mathbb{R}\}$$

is an affine subspace of dimension $n-k$. $\square$

By definition, ψ restricts to a homeomorphism $\varphi\colon M\dashrightarrow\mathbb{R}^m$ and we think of this map as introducing local coordinates on M and thus call it local *coordinates* or a local *chart* on M. Its inverse $\varphi^{-1}\colon\mathbb{R}^m\dashrightarrow M$ is called a local *parametrization* of M.

Let φ_1, φ_2 be two local charts, i.e. restrictions of diffeomorphisms ψ_1, ψ_2 from the definition of a submanifold. Then $\varphi_{12}=\varphi_2\circ\varphi_1^{-1}$ is called the *transition map* for the coordinates. It is clearly a diffeomorphism as a restriction of the diffeomorphism $\psi_2\circ\psi_1^{-1}$.

The above serves as a motivation for the definition of an abstract manifold (i.e. not a submanifold of some vector space).

7. Smooth manifolds

Definition 7.1. A *topological manifold* is a topological space M that

- is Hausdorff,
- has a countable basis for topology, and
- is *locally euclidean*, i.e. each point $x \in M$ has an open neighbourhood $U \ni x$ that is homeomorphic to some open subset $V \subseteq \mathbb{R}^m$, i.e. there exists $\varphi\colon U \xrightarrow{\cong} V$.

We say that M has *dimension* m.

We remind the reader that M is Hausdorff if any pair of distinct points $x \neq y$ admits a pair of disjoint neighbourhoods $U \ni x$, $V \ni y$, $U \cap V = \emptyset$. Equivalently, any sequence[1] has at most one limit point ("$\Rightarrow$": if a sequence converged to both x and y, it would have to lie eventually in U and V, thus eventually in $U \cap V = \emptyset$, a contradiction; "$\Leftarrow$": if $B_{1/n}(x) \cap B_{1/n}(y)$ was non-empty, containing a point z_n, the sequence z_n would converge to both x and y, a contradiction).

Further, we remind that a basis for topology is a collection $\mathcal{U}$ of open sets such that any open set W is a union of some elements from this collection, $W = \bigcup_{U \in \mathcal{U}, U \subseteq W} U$. Any open subset $V \subseteq \mathbb{R}^n$ is second countable, generated by all open balls $B_\varepsilon(x^1, \dots, x^n) \subseteq V$ with all $x^1, \dots, x^n$ and ε rational.

For a pair of local charts φ_1, φ_2 on M, we again form the transition map $\varphi_{12} = \varphi_2 \circ \varphi_1^{-1}$. We say that the charts are *compatible* if the transition map is a diffeomorphism. An *atlas* is a collection of charts $\mathcal{A} = \{\varphi_i \mid i \in I\}$ whose domains cover M, i.e. $\bigcup_{i \in I} \operatorname{dom} \varphi_i = M$, and such that any pair is compatible.

Lemma 7.2. *Any pair of charts ψ, χ compatible with an atlas $\mathcal{A}$ is itself compatible.*

Proof. We consider

$$\begin{tikzcd} & M \arrow[dl, dashed, "\psi"'] \arrow[d, dashed, "\varphi_i"] \arrow[dr, dashed, "\chi"] & \\ \mathbb{R}^m \arrow[r, dashed, "\varphi_i \circ \psi^{-1}"'] & \mathbb{R}^m \arrow[r, dashed, "\chi \circ \varphi_i^{-1}"'] & \mathbb{R}^m \end{tikzcd}$$

with both horizontal maps smooth by assumption. Their composition $\chi \circ \psi^{-1}$ is then also smooth at points where this composition is defined. But for any $x \in \operatorname{dom} \psi \cap \operatorname{dom} \chi$, we may choose φ_i with $x \in \operatorname{dom} \varphi_i$ and then the horizontal composition is defined at $\psi(x)$. □

Corollary 7.3. *Let $\mathcal{A}$ be an atlas. Then there exists a unique maximal atlas containing $\mathcal{A}$, consisting of all charts compatible with $\mathcal{A}$.* □

Definition 7.4. A *smooth manifold* is a topological manifold M equipped with a maximal atlas.

When speaking of charts (or coordinates) on a smooth manifold, we will always mean a chart from the given maximal atlas.

Definition 7.5. A continuous map $F\colon M \to N$ between smooth manifolds is said to be *smooth* if, for every chart φ on M and every chart ψ on N, the composition $\psi \circ F \circ \varphi^{-1}\colon \mathbb{R}^m \to \mathbb{R}^n$ is smooth.

$$\begin{tikzcd} M \arrow[r, "F"] \arrow[d, dashed, "\varphi"'] & N \arrow[d, dashed, "\psi"] \\ \mathbb{R}^m \arrow[r, dashed, "\psi \circ F \circ \varphi^{-1}"'] & \mathbb{R}^n \end{tikzcd}$$

Similarly to the previous lemma (in fact, the lemma is a special case for $F = \mathrm{id}$), it is enough to check smoothness for some atlas on M and a collection of charts covering $\operatorname{im} F$.

Example 7.6. A smooth submanifold $M \subseteq \mathbb{R}^n$ is a smooth manifold and the inclusion ι is smooth.

---

[1] Since M is locally euclidean, every point has a countable basis of neighbourhoods and thus sequences suffice to capture topology. In general, one would replace them by nets or filters.

Example 7.7. A local chart is precisely a diffeomorphism $M \dashrightarrow \mathbb{R}^m$. Construct two local charts on $S^m \subseteq \mathbb{R}^{m+1}$ and study the transition map (it should be a disc inversion).

Remark. The charts are $(x^0, x) \mapsto \frac{1}{1 \pm x^0} \cdot x$ with inverses $x \mapsto \frac{1}{1+|x|^2}(\pm(1-|x|^2), 2 \cdot x)$.

8. Tangent bundle

Definition 8.1. We say that $A \colon C^\infty M \to \mathbb{R}$ is a *derivation at* x_0 if A is $\mathbb{R}$-linear and satisfies the *Leibniz rule at* x_0,

$$A(f \cdot g) = Af \cdot g(x_0) + f(x_0) \cdot Ag.$$

Definition 8.2. We define the *tangent space* T_xM of a smooth manifold M at a point $x \in M$ to be

$$T_xM = \{A \colon C^\infty M \to \mathbb{R} \mid A \text{ is derivation at } x\},$$

the set of all derivations at x.

It is clear that derivations at x are closed under addition and multiplication by real scalars and thus T_xM is a vector space over $\mathbb{R}$. Our main aim will now be to show that this vector space can be computed in coordinates and is thus isomorphic to $\mathbb{R}^m$ (in particular, it is finite dimensional, something that is not obvious from the definition).

In order to compare the tangent spaces of various manifolds, we define the *tangent map* φ_* of a smooth map $\varphi \colon M \to N$ in two steps. First, the precomposition with φ defines a map

$$\varphi^* \colon C^\infty N \to C^\infty M$$

given by $\varphi^* f = f \circ \varphi$ and it is clearly a homomorphism of algebras, e.g. $\varphi^*(f \cdot g) = \varphi^* f \cdot \varphi^* g$. Next comes:

Lemma 8.3. *The algebra homomorphism φ^* determines a map*

$$\varphi_* \colon \operatorname{Der}_x(C^\infty M, \mathbb{R}) \to \operatorname{Der}_{\varphi(x)}(C^\infty N, \mathbb{R}),$$

given by $\varphi_* A = A \circ \varphi^*$, i.e. by $(\varphi_* A)f = A(\varphi^* f) = A(f \circ \varphi)$.

Proof. We need to show that $\varphi_* A$ is indeed a derivation at $\varphi(x)$. The linearity of $\varphi_* A$ is clear and the Leibniz rule is

$$\begin{aligned}(\varphi_* A)(f \cdot g) &= A(\varphi^*(f \cdot g)) = A(\varphi^* f \cdot \varphi^* g) = A(\varphi^* f) \cdot \varphi^* g(x) + \varphi^* f(x) \cdot A(\varphi^* g) \\ &= (\varphi_* A)f \cdot g(\varphi(x)) + f(\varphi(x)) \cdot (\varphi_* A)g. \qquad \square\end{aligned}$$

By our definition of tangent spaces, φ_* is thus a map

$$\varphi_* \colon T_xM \to T_{\varphi(x)}N.$$

We define

$$TM = \bigsqcup_{x \in M} T_xM$$

and call it the *tangent bundle* of M. The various maps $\varphi_* \colon T_xM \to T_{\varphi(x)}N$ define together a map

$$\varphi_* \colon TM \to TN,$$

called again the *tangent map* of φ. We use the notation $\varphi_{*x} \colon T_xM \to T_{\varphi(x)}N$ if we want to stress that it is defined only on the tangent space at x (see e.g. the next theorem).

Theorem 8.4. *If $\varphi \colon M \dashrightarrow N$ is a local diffeomorphism at x then φ_{*x} is an isomorphism.*

Germs. In fact, we have not defined a tangent map to a partially defined map φ so far. This and the proof of the theorem will be achieved by passing to *germs* of functions at a point x. We consider smooth functions $f\colon M \dashrightarrow \mathbb{R}$ defined on an open neighbourhood $\operatorname{dom} f \ni x$ and an equivalence relation on such functions: $f \sim g$ if and only if $f = g$ in some neighbourhood of x. The equivalence class of f is called the *germ* of f at x and denoted $\operatorname{germ}_x f$. The set of all germs at x will be denoted $C_x^\infty M$. In particular, we obtain a map

$$\operatorname{germ}_x\colon C^\infty M \to C_x^\infty M$$

associating to each (globally defined) function its germ at x.

Lemma 8.5. *The above map* germ_x *is surjective.*

Proof. Let $f\colon U \to \mathbb{R}$ be a smooth function. We choose a function $\lambda\colon M \to [0,1]$ such that $\lambda = 1$ in a neighbourhood of x and such that $\operatorname{supp}\lambda \subseteq U$. Then $\operatorname{germ}_x f = \operatorname{germ}_x(\lambda\cdot f)$, since $\lambda = 1$ near x, and the product $\lambda \cdot f$ can be extended by zero to a smooth function of M, giving a preimage of $\operatorname{germ}_x f$. □

It is easy to define linear combinations and products of germs in terms of their representatives (more abstractly, one may use the above lemma to view $C_x^\infty M$ as the quotient algebra of $C^\infty M$). Thus, it makes sense to say that a map $C_x^\infty M \to \mathbb{R}$ is a derivation at x (on the other hand, it makes no sense to speak about derivations at other points, since the value of a germ at a point y different from x is ill-defined).

Lemma 8.6. *Every derivation* $A\colon C^\infty M \to \mathbb{R}$ *at* x *factors uniquely*

$$\begin{tikzcd} C^\infty M \arrow[r, "A"] \arrow[d, two heads, "\operatorname{germ}_x"'] & \mathbb{R} \\ C_x^\infty M \arrow[ur, dashed, "\exists!"'] & \end{tikzcd}$$

through a derivation $C_x^\infty M \to \mathbb{R}$*, i.e. the precomposition with* germ_x *gives an isomorphism*

$$\operatorname{Der}_x(C_x^\infty M, \mathbb{R}) \xrightarrow{\cong} \operatorname{Der}_x(C^\infty M, \mathbb{R}) = T_x M.$$

In words, the above lemma says that it is possible to define a unique derivation of germs if we impose that the derivation of a germ of a globally defined function $f \in C^\infty M$ is Af.

Proof. The factorization, if exists, is unique by the surjectivity of germ_x. On the other hand, a factorization exists if any only if $\operatorname{germ}_x f = 0 \Rightarrow Af = 0$. The condition $\operatorname{germ}_x f = 0$ means that $f = 0$ in some neighbourhood $U \ni x$. Let $\lambda\colon M \to [0,1]$ be such that $\lambda = 1$ in a neighbourhood of x and such that $\operatorname{supp}\lambda \subseteq U$. Set $\rho = 1 - \lambda$, so that $\rho = 0$ in a neighbourhood of x and $\rho = 1$ on $M \smallsetminus U$. Thus, $f = \rho \cdot f$ and

$$Af = A(\rho\cdot f) = A\rho \cdot \underbrace{f(x)}_{0} + \underbrace{\rho(x)}_{0} \cdot Af = 0. \qquad \square$$

Now we are ready to prove the theorem.

Proof of Theorem 8.4. Clearly, any smooth map $\varphi\colon M \dashrightarrow N$ defines a map

$$\varphi^*\colon C_{\varphi(x)}^\infty N \to C_x^\infty M.$$

In the case that φ is a local diffeomorphism at x, the map φ^* has an inverse $(\varphi^{-1})^*$ and is therefore an isomorphism. In particular, it induces an isomorphism

$$\varphi_*\colon T_x M = \operatorname{Der}_x(C_x^\infty M, \mathbb{R}) \xrightarrow{\cong} \operatorname{Der}_{\varphi(x)}(C_{\varphi(x)}^\infty N, \mathbb{R}) = T_{\varphi(x)} N. \qquad \square$$

Tangent bundle as a smooth manifold. We start with a simple observation. If $\mathcal{U}$ is a covering of a topological space X, then a subset $V \subseteq X$ is open if and only if $U \cap V$ is open for every $U \in \mathcal{U}$. Thus, given an atlas on M, a subset $V \subseteq M$ is open if and only if for every chart $\varphi\colon M \dashrightarrow \mathbb{R}^m$ the image $\varphi(V \cap \operatorname{dom}\varphi) \subseteq \mathbb{R}^m$ is open. Consequently, an atlas also determines the toplogy of M.

Now, we would like to equip TM with a structure of a smooth manifold. Its atlas will consist of the tangent maps to charts on M, i.e. for $\varphi\colon M \dashrightarrow \mathbb{R}^m$ we consider

$$\mathcal{A} = \{\varphi_*\colon TM \dashrightarrow T\mathbb{R}^m \cong \mathbb{R}^{2m}\}.$$

Theorem 8.7. *There is a structure of a smooth manifold of dimension $2m$ on TM, given by the above atlas.*

State the main ingredient formally.

Proof. A smooth structure is given by a countable collection of maps $M \dashrightarrow \mathbb{R}^m$ (charts), bijections from the domain onto the image, such that the transition maps are smooth maps between open subsets of $\mathbb{R}^m$ and such that any pair of points lies in a domain of a chart.

Now the above property for M implies the same for TM, since the transition maps φ_{ij*} are smooth. $\square$

9. Vector fields

There is a canonical projection $p\colon TM \to M$, associating to each $A \in T_xM$ its base $p(A) = x$.

Definition 9.1. A smooth *vector field* is a smooth map $X\colon M \to TM$ such that $X|_x \in T_xM$, i.e.

$$\begin{tikzcd} & TM \arrow[d, "p"] \\ M \arrow[ur, "X"] \arrow[r, "\mathrm{id}"'] & M \end{tikzcd}$$

The definition of a smooth structure on TM implies that a vector field is smooth if and only if its expression in coordinates is smooth, i.e. if $X = X^i\partial_i$ with coordinate functions X^i smooth. Thus, every smooth vector field X induces a derivation $X\colon C^\infty M \to C^\infty M$ (denoted by the same symbol), given by $Xf(x) = X|_x f$; locally $Xf = X^i \cdot \partial_i f$ and is clearly smooth. We have the following converse.

Theorem 9.2. *For a derivation $X\colon C^\infty M \to C^\infty M$ there exists a unique smooth vector field inducing it.*

Proof. The composition $\mathrm{ev}_x \circ X$ with the evaluation map $\mathrm{ev}_x\colon C^\infty M \to \mathbb{R}$ gives a derivation at x and is thus given by a unique vector $X|_x \in T_xM$,

$$Xf(x) = X|_x f.$$

It remains to show that X is smooth. In local coordinates, $X^i = Xx^i$ which almost gives smoothness except $x^i \notin C^\infty M$ so that Xx^i does not make sense. This is corrected by passing, at each point, to germs and the extension of $X|_x$ to germs at x. We choose, in a coordinate neighbourhood $U \ni x_0$, a smooth function λ with $\lambda = 1$ in a neighbourhood $V \ni x_0$ and with $\operatorname{supp}\lambda \subseteq U$; then for $x \in V$

$$X^i(x) = X|_x x^i = X|_x(\lambda x^i) = X(\lambda x^i)(x)$$

since $\mathrm{germ}_x\, x^i = \mathrm{germ}_x\, \lambda x^i$. As λx^i is a smooth function on M, so is its image $X(\lambda x^i)$ and X^i is smooth on V. $\square$

Let $\gamma\colon \mathbb{R} \dashrightarrow M$ be a path, i.e. we assume that $\operatorname{dom}\gamma$ is an interval. We define the *tangent vector* to γ at time t_0 to be $\gamma'|_{t_0} = \gamma_*(\partial_t|_{t_0})$. The chain rule then easily gives

$$\varphi_*(\gamma'|_t) = (\varphi \circ \gamma)'|_t,$$

i.e. the image of a tangent vector to a path γ under the tangent map φ_* is the tangent vector to the image of the path under φ.

Definition 9.3. We say that γ is an *integral curve* of a vector field X if

$$\gamma'|_t = X|_{\gamma(t)}$$

for each $t \in \operatorname{dom}\gamma$.

In local coordinates this reads

$$\partial_t\gamma^i(t) = X^i(\gamma^1(t), \ldots, \gamma^m(t)),$$

i.e. the tuple of coordinate functions γ^i forms a solution of a system of ordinary differential equations. Since the involved functions X^i are smooth, there exists a solution γ_x with any given initial value $\gamma_x(0) = x$. Together these form a map

$$\mathrm{Fl}^X\colon \mathbb{R} \times M \dashrightarrow M, \quad \mathrm{Fl}^X(t,x) = \gamma_x(t),$$

defined in a neighbourhood of $\{0\} \times M$ and smooth. This map is called the *flow* of X.

Theorem 9.4. $\mathrm{Fl}^X(t, \mathrm{Fl}^X(s,x)) = \mathrm{Fl}^X(t+s,x)$.

Proof. We have to show that the right hand side $\gamma(t) = \mathrm{Fl}^X(t+s,x)$ is an integral curve of X. However, since the translation by s clearly takes $\partial_t|_{t_0}$ to $\partial_t|_{t_0+s}$, we get

$$\gamma'|_{t_0} = (\mathrm{Fl}^X(-,x))'|_{t_0+s} = X|_{\mathrm{Fl}^X(t_0+s,x)} = X|_{\gamma(t_0)}.$$

Thus, it is indeed an integral curve; the initial value is also correct, $\gamma(0) = \mathrm{Fl}^X(s,x)$. □

We say that X is *complete* if the flow is defined on $\mathbb{R} \times M$. We define a *support* of a vector field X, denoted $\operatorname{supp} X$, to be the closure of the set $\{x \in M \mid X|_x \neq 0\}$.

Theorem 9.5. *A compactly supported vector field is complete. In particular, any vector field on a compact manifold is complete.*

Proof. For every $x \in \operatorname{supp} X$ there is a neighbourhood U_x and $\varepsilon_x > 0$ such that Fl^X is defined on $(-\varepsilon_x, \varepsilon_x) \times U_x$. Since the support is compact, we have $\operatorname{supp} X \subseteq U_{x_1} \cup \cdots \cup U_{x_k}$. Taking $\varepsilon = \min\{\varepsilon_{x_1}, \ldots, \varepsilon_{x_k}\}$, the flow of X is defined on $(-\varepsilon, \varepsilon) \times M$ (at points not in the support, the integral curve through that point is constant and thus defined on $\mathbb{R}$). By the previous theorem, we may write, for any $t \in \mathbb{R}$ and any $x \in M$,

$$\mathrm{Fl}^X(t,x) = \mathrm{Fl}^X(t/T, \cdots \mathrm{Fl}^X(t/T, x) \cdots)$$

and for $T \gg 0$ we have $t/T \in (-\varepsilon, \varepsilon)$ so that the right hand side is defined. □

Definition 9.6. Let $\varphi\colon M \to N$ be a smooth map, X a vector field on M and Y a vector field on N. We say that X and Y are *φ-related* if

$$\varphi_* X|_x = Y|_{\varphi(x)},$$

i.e. if the following diagram commutes:

$$\begin{tikzcd} TM \arrow[r, "\varphi_*"] & TN \\ M \arrow[u, "X"] \arrow[r, "\varphi"'] & N \arrow[u, "Y"'] \end{tikzcd}$$

We will occasionally denote this by $X \sim_\varphi Y$.

A special case is that of an inclusion of a submanifold, that we will denote $\mathrm{in}\colon M \subseteq N$. Since each in_{*x} is injective, we may and will think of it as an inclusion of a subspace; then for each vector field $Y \in \mathfrak{X}N$, a vector field $X \in \mathfrak{X}M$ with $X \sim_\varphi Y$ exists if and only if for each $x \in M$, the value $Y|_x$ lies in $T_xM \subseteq T_xN$. We say that Y is *tangent* to N.

We have the following two characterizations using the induced derivation of functions and using the flows.

Lemma 9.7. *Vector fields X and Y are φ-related if and only if*

$$(Yf) \circ \varphi = X(f \circ \varphi).$$

Proof. This is clear upon unfolding the value of the right hand side at x:

$$X(f \circ \varphi)(x) = X|_x(f \circ \varphi) = (\varphi_* X|_x)f$$

while the value of the left hand side is simply $Y|_{\varphi(x)}f$. □

Lemma 9.8. *Vector fields X and Y are φ-related if and only if*

$$\varphi(\mathrm{Fl}^X(t,x)) = \mathrm{Fl}^Y(t,\varphi(x))$$

In other words φ maps the flow lines of X to the flow lines of Y. We will use this property quite often.

Proof. Taking the tangent vectors to the paths in the above equality, we get

$$\varphi_* X|_x = (\varphi(\mathrm{Fl}^X(-,x)))'|_0 = (\mathrm{Fl}^Y(-,\varphi(x)))'|_0 = Y|_{\varphi(x)},$$

which is precisely the definition of φ-relatedness.

In the opposite direction, given that X and Y are φ-related, we wish to prove the equality from the statement, i.e. we want to prove that $\gamma(t) = \varphi(\mathrm{Fl}^X(t,x))$ is an integral curve of Y through $\varphi(x)$. Since the initial value $\gamma(0) = \varphi(x)$ is correct, we need only check that it satisfies the differential equation of an integral curve:

$$\gamma'(t_0) = (\varphi(\mathrm{Fl}^X(-,x)))'|_{t_0} = \varphi_* X|_{\mathrm{Fl}^X(t_0,x)} = Y|_{\varphi(\mathrm{Fl}^X(t_0,x))} = Y|_{\gamma(t_0)}$$

(the third equality uses the φ-relatedness). □

10. Lie bracket

We define the Lie bracket through derivations.

Definition 10.1. Let X and Y be two vector fields on M. Then it is easy to see that $f \mapsto XYf - YXf$ is a derivation and the corresponding vector field is denoted $[X,Y]$ and called the *Lie bracket* of the vector fields X and Y.

We will now derive a coordinate formula

$$\begin{aligned} XYf - YXf &= X^j \cdot \partial_j(Y^i \cdot \partial_i f) - Y^j \cdot \partial_j(X^i \cdot \partial_i f) \\ &= X^j \cdot (\partial_j Y^i \cdot \partial_i f + Y^i \cdot \partial_j \partial_i f) - Y^j \cdot (\partial_j X^i \cdot \partial_i f + X^i \cdot \partial_j \partial_i f) \\ &= (X^j \cdot \partial_j Y^i - Y^j \cdot \partial_j X^i) \cdot \partial_i f \end{aligned}$$

so that $[X,Y] = (X^j \cdot \partial_j Y^i - Y^j \cdot \partial_j X^i) \cdot \partial_i$.

Proposition 10.2. *The Lie bracket has the following properties:*

- $[X, Y_1 + Y_2] = [X, Y_1] + [X, Y_2]$,
- $[X, f \cdot Y] = Xf \cdot Y + f \cdot [X, Y]$,
- $[X, Y] = -[Y, X]$,
- $[X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]]$.
- $X \sim_\varphi Z,\ Y \sim_\varphi W \Rightarrow [X, Y] \sim_\varphi [Z, W]$.

Proof. All points are rather straightforward, we explain the most interesting one – the second:

$$\begin{aligned} [X, f \cdot Y]g &= X(f \cdot Yg) - f \cdot Y(Xg) \\ &= Xf \cdot Yg + f \cdot X(Yg) - f \cdot Y(Xg) \\ &= Xf \cdot Yg + f \cdot ([X,Y]g) \\ &= (Xf \cdot Y + f \cdot [X,Y])g. \end{aligned}$$

The last point is also interesting:

$$\begin{aligned} ([Z,W]f) \circ \varphi &= Z(Wf) \circ \varphi - W(Zf) \circ \varphi \\ &= X(Wf \circ \varphi) - Y(Zf \circ \varphi) \\ &= X(Y(f \circ \varphi)) - Y(X(f \circ \varphi)) \\ &= [X,Y](f \circ \varphi). \end{aligned}$$

□

Denoting $\mathcal{L}_X f = Xf$ and $\mathcal{L}_X Y = [X,Y]$ (the Lie derivatives of f and Y along X), the second point becomes $\mathcal{L}_X(f \cdot Y) = \mathcal{L}_X f \cdot Y + f \cdot \mathcal{L}_X Y$ and the fourth becomes $\mathcal{L}_X[Y,Z] = [\mathcal{L}_X Y, Z] + [Y, \mathcal{L}_X Z]$, i.e. both are some forms of the Leibniz rule.

Corollary 10.3. *Let $M \subseteq N$ be a submanifold. If $X, Y \in \mathfrak{X}N$ are tangent to M, so is $[X,Y]$.*

Proof. This is just the last point applied to the inclusion $\varphi = \mathrm{in}$. $\square$

Definition 10.4. Let X, Y be two vector fields on a manifold M. Then we denote

$$(\mathrm{Fl}_t^X)^* Y(x) = (\mathrm{Fl}_{-t}^X)_* Y(\mathrm{Fl}_t^X(x)) \in T_x M$$

the *pullback* of Y along the flow Fl_t^X of X. For each $x \in M$ it is defined for t small. (In other words, the pullback $(\mathrm{Fl}_t^X)^* Y$ is the unique vector field that is Fl_t^X-related to Y.)

The *Lie derivative* of Y along X is

$$\mathcal{L}_X Y(x) = \partial_t|_0((\mathrm{Fl}_t^X)^* Y(x)).$$

Theorem 10.5. *Let X be a vector field and $x \in M$ a point. If $X|_x \neq 0$ then, in a neighbourhood of x, there exists a coordinate chart in which $X = \partial_1$.*

Proof. We set $X_1 = X$ and choose vector fields $X_2, \dots X_m$ so that $X_1, \dots, X_m$ form a basis in a neighbourhood of x. We define a map

$$\varphi\colon \mathbb{R}^m \dashrightarrow M, \quad (t^1, \dots, t^m) \mapsto \mathrm{Fl}_{t^1}^{X_1} \cdots \mathrm{Fl}_{t^m}^{X_m}(x)$$

(it is defined in a neighbourhood of 0). The image of the coordinate vector field

$$\partial_i|_0 = \partial_t|_0(0, \dots, t, \dots, 0)$$

at the origin then equals

$$\varphi_* \partial_i|_0 = \partial_t|_0 \mathrm{Fl}_t^{X_i}(x) = X_i|_x$$

and φ is a local diffeomorphism at 0. We may thus use its inverse φ^{-1} as a coordinate chart on M.

Now, for $i = 1$, we get more generally

$$\varphi_* \partial_1|_{t_0} = \partial_t|_{t_0^1} \mathrm{Fl}_t^{X_1} \mathrm{Fl}_{t_0^2}^{X_2} \cdots \mathrm{Fl}_{t_0^m}^{X_m}(x) = X_1(\mathrm{Fl}_{t_0^1}^{X_1} \mathrm{Fl}_{t_0^2}^{X_2} \cdots \mathrm{Fl}_{t_0^m}^{X_m}(x)) = X_1|_{\varphi(t_0)}$$

and this shows that, in the coordinates given by φ^{-1}, we have $\partial_1 = X_1$. $\square$

Proposition 10.6. *The following holds*

$$\mathcal{L}_X Y|_x = [X,Y]|_x.$$

More generally, $\partial_t|_{t_0}(\mathrm{Fl}_t^X)^* Y|_x = (\mathrm{Fl}_{t_0}^X)^*[X,Y]|_x$.

Proof. First assume that $t_0 = 0$. Let $x \in M$ be such that $X|_x \neq 0$. Then, by Theorem 10.5, there is a coordinate chart in which $X = \partial_1$ near x. Then $\mathrm{Fl}_t^X(x) = x + t \cdot \partial_1|_x$ is the translation by the t-multiple of the coordinate vector $\partial_1|_x$, its derivative is then the identity. Consequently, $(\mathrm{Fl}_{-t}^X)_* \partial_i|_{x + t\cdot\partial_1} = \partial_i|_x$, so that, for $Y = Y^i \cdot \partial_i$, we get

$$(\mathrm{Fl}_t^X)^* Y|_x = (\mathrm{Fl}_{-t}^X)_* Y(\mathrm{Fl}_t^X(x)) = Y^i(x + t \cdot \partial_1) \cdot \partial_i|_x$$

and finally

$$\mathcal{L}_X Y|_x = \partial_t|_0 (\mathrm{Fl}_t^X)^* Y|_x = \partial_1 Y^i(x) \cdot \partial_i.$$

This equals $[X,Y]|_x$ by the coordinate formula for the Lie bracket, proving the claim in this case.

By continuity, the same holds for points in the closure of the set $\{x \in M \mid X|_x \neq 0\}$, i.e. on the support of X. On the other hand, if $x \notin \operatorname{supp} X$, then $X = 0$ in a neighbourhood of x and then both sides equal zero.

For a general t_0, we have $(\mathrm{Fl}_t^X)^* Y|_x = (\mathrm{Fl}_{t_0}^X)^*(\mathrm{Fl}_{t-t_0}^X)^* Y|_x$. Since $(\mathrm{Fl}_{t_0}^X)^*$ is a linear map we can interchange with ∂_t. $\square$

Remark. We will need a useful property for the proof of the next proposition. It is based on an observation that for a function $\varphi(s,t)$ of two variables, with values in a vector space, we have

$$\partial_t|_{t_0}\varphi(t,t) = \partial_t|_{t_0}\varphi(t,t_0) + \partial_t|_{t_0}\varphi(t_0,t).$$

Now let X be a time-dependent vector field, i.e. a map $X\colon \mathbb{R}\times M \to TM$ such that $X(t,x)\in T_xM$ (**not true** in the second application of $(*)$, but also not necessary there). Let $f\colon \mathbb{R}\times M\to\mathbb{R}$ be a "time-dependent" function on M (just a function on $\mathbb{R}\times M$). Write X_t and f_t for the vector field and function obtained by plugging in a specific value of t. Then we may form the directional derivative X_tf_t and

$$\partial_t|_{t_0}X_tf_t = \partial_t|_{t_0}(X_tf_{t_0}) + \partial_t|_{t_0}(X_{t_0}f_t). \tag{$*$}$$

(Locally, we have $X_tf_t(x) = X(t,x)^i\cdot\partial_i f(t,pX(t,x))$ and we apply the previous observation.)

Proof. First assume that $t_0=0$ and let $f: M\to\mathbb{R}$ be a smooth function. We differentiate f in the direction of the left hand side:

$$\begin{aligned}
\Big(\partial_t|_0(\mathrm{Fl}_t^X)^*Y(x)\Big)f &\overset{(*)}{=} \partial_t|_0\Big((\mathrm{Fl}_t^X)^*Y(x)f\Big)\\
&= \partial_t|_0\Big((\mathrm{Fl}_{-t}^X)_*Y(\mathrm{Fl}_t^X(x))f\Big)\\
&= \partial_t|_0\Big(Y(\mathrm{Fl}_t^X(x))(f\circ\mathrm{Fl}_{-t}^X)\Big)\\
&\overset{(*)}{=} \partial_t|_0\Big(Y(x)(f\circ\mathrm{Fl}_{-t}^X)\Big) + \partial_t|_0\Big(Y(\mathrm{Fl}_t^X(x))(f)\Big)\\
&\overset{(*)}{=} Y(x)\Big(\partial_t|_0(f\circ\mathrm{Fl}_{-t}^X)\Big) + \partial_t|_0\Big((Yf)(\mathrm{Fl}_t^X(x))\Big)\\
&= Y(x)(-Xf) + X(x)(Yf)\\
&= -(YXf)(x) + (XYf)(x) = \Big([X,Y](x)\Big)f
\end{aligned}$$

(the steps labeled by $(*)$ involve the observation made before the proposition, of which the first in the opposite direction).

For a general t_0, we have $(\mathrm{Fl}_t^X)^*Y(x) = (\mathrm{Fl}_{t_0}^X)^*(\mathrm{Fl}_{t-t_0}^X)^*Y(x)$. Since $(\mathrm{Fl}_{t_0}^X)^*$ is a linear map we can interchange with ∂_t.

Corollary 10.7. *The following conditions are equivalent:*

- $[X,Y]=0$,
- $(\mathrm{Fl}_t^X)^*Y = Y$*, i.e. Y is Fl_t^X-related with itself for all t,*
- $\mathrm{Fl}_t^X\,\mathrm{Fl}_s^Y(x) = \mathrm{Fl}_s^Y\,\mathrm{Fl}_t^X(x)$*, i.e. the flow lines commute.*

In general we have $\mathrm{Fl}_{-s}^Y\,\mathrm{Fl}_{-t}^X\,\mathrm{Fl}_s^Y\,\mathrm{Fl}_t^X(x) = x + st[X,Y](x) + o(s,t)^2$.

Proof. The equivalence of the three conditions follows immediately from the previous proposition – the second condition states that $(\mathrm{Fl}_t^X)^*Y$ is a constant function of t, i.e. that the derivative is zero and this is clearly equivalent to the first condition. At the same time, the second condition is equivalent to Y being Fl_t^X-related to itself and this is equivalent to Fl_t^X preserving the integral curves of Y, which is precisely the third condition.

Differentiating the commutator of the flows twice, we get

$$\partial_t|_0\partial_s|_0\,\mathrm{Fl}_{-s}^Y\,\mathrm{Fl}_{-t}^X\,\mathrm{Fl}_s^Y\,\mathrm{Fl}_t^X(x) = \partial_s|_0\Big(-Y(x) + (\mathrm{Fl}_t^X)^*Y(x)\Big) = [X,Y](x)$$

The remaining derivatives of order at most two are clearly zero. □

We will need the following generalization of Theorem 10.5.

Theorem 10.8. *If vector fields $X_1,\dots,X_k$ are linearly independent and satisfy $[X_i,X_j]=0$ then, in a neighbourhood of any point x, there exists a coordinate chart in which $X_i=\partial_i$.*

Proof. We choose vector fields $X_{k+1},\dots X_m$ so that $X_1,\dots,X_m$ form a basis in a neighbourhood of x. We define a map

$$\varphi\colon \mathbb{R}^m \dashrightarrow M,\quad (t^1,\dots,t^m)\mapsto \mathrm{Fl}_{t^1}^{X_1}\cdots\mathrm{Fl}_{t^m}^{X_m}(x)$$

(it is defined in a neighbourhood of 0). The image of the coordinate vector field

$$\partial_i|_0 = \partial_t|_0(0,\dots,t,\dots,0)$$

at the origin then equals

$$\varphi_*\partial_i|_0 = \partial_t|_0\,\mathrm{Fl}_t^{X_i}(x) = X_i|_x$$

and φ is a local diffeomorphism at 0. We may thus use its inverse φ^{-1} as a coordinate chart on M.

Now, for $i \leq k$, we study the image of the coordinate vector field $\partial_i|_{t_0}$ at a general point t_0. Since we may interchange the flows (since $[X_i, X_j] = 0$), we get

$$\varphi_* \partial_i|_{t_0} = \partial_t|_{t_0^i} \operatorname{Fl}_t^{X_i} \operatorname{Fl}_{t_0^1}^{X_1} \cdots \widehat{\operatorname{Fl}_{t_0^i}^{X_i}} \cdots \operatorname{Fl}_{t_0^m}^{X_m}(x) = X_i(\operatorname{Fl}_{t_0^i}^{X_i} \operatorname{Fl}_{t_0^1}^{X_1} \cdots \widehat{\operatorname{Fl}_{t_0^i}^{X_i}} \cdots \operatorname{Fl}_{t_0^m}^{X_m}(x)) = X_i|_{\varphi(t_0)}$$

and this shows that, in the coordinates given by φ^{-1}, we have $\partial_i = X_i$, for $i \leq k$. □

11. Distributions

Definition 11.1. A (non-smooth) *distribution* $\mathcal{S}$ of dimension k is a mapping $x \mapsto \mathcal{S}(x)$ that associates to each point $x \in M$ a k-dimensional vector subspace $\mathcal{S}(x) \subseteq T_x M$.

A distribution $\mathcal{S}$ is smooth if, for each point $x_0 \in M$, there exist a neighbourhood U and local vector fields $X_1, \ldots, X_k \in \mathfrak{X}U$ such that $X_1|_x, \ldots, X_k|_x$ form a basis of $\mathcal{S}(x)$ for $x \in U$.

From now on, all our distributions will be smooth.

Definition 11.2. An submanifold $N \subseteq M$ is said to be an *integral manifold* of a distribution $\mathcal{S}$ if, for each $x \in N$, one hase $\mathcal{S}(x) = T_x N$.

A distribution $\mathcal{S}$ on M is called *integrable* if, for each $x \in M$, there exists an integral manifold passing through x.

We say that a vector field X lies in $\mathcal{S}$ if, for each x, we have $X|_x \in \mathcal{S}(x)$. Integrable distributions have a special property of being involutive. In fact, the converse also holds, as we will see shortly.

Definition 11.3. A distribution $\mathcal{S}$ on M is called *involutive* if, for every pair of vector fields X, Y lying in $\mathcal{S}$, their bracket $[X, Y]$ also lies in $\mathcal{S}$.

Theorem 11.4. *Every integrable distribution is involutive.*

Proof. Let X and Y be vector fields lying in $\mathcal{S}$, let $x \in M$ be an arbitrary point and let N be an integral manifold passing through x. Then both X and Y are tangent to N and, thus, so is their Lie bracket $[X, Y]$. In particular, $[X, Y]|_x \in \mathcal{S}(x)$. Since x was arbitrary, $[X, Y]$ indeed lies in $\mathcal{S}$. □

Now we are ready to prove the converse, in fact it proves a stronger version of integrability, since it also describes how the integral manifolds vary locally – they form a so-called *foliation*.

Theorem 11.5 (Frobenius theorem)**.** *If $\mathcal{S}$ is involutive then for every $x_0 \in M$ there exists a local coordinate system in a neighbourhood $U \ni x_0$ such that the vector fields $\partial_1, \ldots, \partial_k$ form a basis of the distribution $\mathcal{S}$ on U. In particular, $\mathcal{S}$ is integrable.*

Proof. Let $X_1, \ldots, X_k$ be vector fields defined in a neighbourhood of x_0 that form a basis of $\mathcal{S}$. By composing the local chart $M \dashrightarrow \mathbb{R}^m$ with a suitable projection map $\mathbb{R}^m \to \mathbb{R}^k$, we get a map $\operatorname{pr}\colon M \dashrightarrow \mathbb{R}^k$ for which $\operatorname{pr}_{*x}\colon \mathcal{S}(x) \subseteq T_x\mathbb{R}^m \to T_{\operatorname{pr}(x)}\mathbb{R}^k$ is an isomorphism for $x = x_0$ and consequently also for x in a neighbourhood of x_0. Since $\partial_1, \ldots, \partial_k$ form a basis, we may write

$$\operatorname{pr}_*(X_1|_x, \ldots, X_k|_x) = (\operatorname{pr}_* X_1|_x, \ldots, \operatorname{pr}_* X_k|_x) = (\partial_1|_{\operatorname{pr}(x)}, \ldots, \partial_k|_{\operatorname{pr}(x)}) \cdot A(x)$$

for an invertible matrix $A(x) \in \mathrm{GL}(k)$ that depends smoothly on x. Since the map $\mathrm{GL}(k) \to \mathrm{GL}(k)$, $M \mapsto M^{-1}$ is smooth (it is given by a rational map using the determinant of the matrix and its minors), the matrix $A^{-1}(x)$ also depends smoothly on x. Defining

$$(Y_1|_x, \ldots, Y_k|_x) = (X_1|_x, \ldots, X_k|_x) \cdot A^{-1}(x),$$

clearly vector fields lying in $\mathcal{S}$ and also giving a basis of $\mathcal{S}$ in a neighbourhood of x_0, we easily get

$$\operatorname{pr}_*(Y_1|_x, \ldots, Y_k|_x) = \operatorname{pr}_*(X_1|_x, \ldots, X_k|_x) \cdot A^{-1}(x) = (\partial_1|_{\operatorname{pr}(x)}, \ldots, \partial_k|_{\operatorname{pr}(x)})$$

so that the Y_i are pr-related to the ∂_i. But then $[Y_i, Y_j]$ is pr-related to $[\partial_i, \partial_j] = 0$. Since $[Y_i, Y_j]$ lies in $\mathcal{S}$ by involutivity and pr_{*x} is an isomorphism on $\mathcal{S}(x)$, we get $[Y_i, Y_j] = 0$. Thus, Theorem 10.8 applies. □

Proof. Let $X_1, \dots, X_k$ be local vector fields which span the distribution $\mathcal{S}$ near x and choose vector fields $X_{k+1}, \dots, X_n$ so that $(X_1, \dots, X_n)$ form a basis near x. We then define a map

$$\begin{aligned} \varphi \colon \mathbb{R}^n \supseteq U &\longrightarrow M \\ (t^1, \dots, t^n) &\longmapsto \mathrm{Fl}^{X_1}_{t^1} \cdots \mathrm{Fl}^{X_n}_{t^n}(x) \end{aligned}$$

The partial derivatives at the origin clearly consist of the vectors $X_i(x)$ and thus φ is a local diffeomorphism – its inverse will form our coordinate system.

Let us compute the partial derivative with respect to t^i for $i \le k$ at a general point.

$$\partial_i \varphi(t^1, \dots, t^n) = \left(\mathrm{Fl}^{X_1}_{t^1}\right)_* \cdots \left(\mathrm{Fl}^{X_{i-1}}_{t^{i-1}}\right)_* X_i \left(\mathrm{Fl}^{X_i}_{t^i} \cdots \mathrm{Fl}^{X_n}_{t^n}(x)\right)$$

To conclude the proof it is therefore enough to show that for any Y belonging to $\mathcal{S}$ the pullbacks $(\mathrm{Fl}^Y_t)^* X_i$ also belong to $\mathcal{S}$ (then the same will be true for pullbacks $(\mathrm{Fl}^Y_t)^* X$ with $X \in \mathcal{S}$ by linearity, and we apply the claim to X_i, $\left(\mathrm{Fl}^{X_{i-1}}_{t^{i-1}}\right)_* X_i$, etc.) Denote this pullback by

$$Y_i(t) = (\mathrm{Fl}^Y_t)^* X_i(x) \in T_x M$$

and write $[Y, X_i] = a^j_i X_j$. By Lemma 10.6 the paths $Y_i(t)$ satisfy the following system of differential equations

$$\tfrac{\mathrm{d}}{\mathrm{d}t} Y_i(t) = (\mathrm{Fl}^Y_t)^* [Y, X_i] = a^j_i(\mathrm{Fl}^Y_t(x)) Y_j(t)$$

We have $Y_i(0) = X_i(x) \in \mathcal{S}(x)$ and, since the system is linear, we must have $Y_i(t) \in \mathcal{S}(x)$ for all t (namely, there exists a solution of the system $\frac{\mathrm{d}}{\mathrm{d}t} Z_i(t) = a^j_i(\mathrm{Fl}^Y_t(x)) Z_j(t)$ with $Z_i \in \mathcal{S}(x)$ and with $Z_i(0) = X_i(x)$. By uniqueness, we must have $Y_i(t) = Z_i(t)$ and, thus, $Y_i(t) \in \mathcal{S}(x)$.)

Theorem 11.6 (Frobenius theorem through 1-forms)**.** *Let $\omega \colon TM \to V$ be a smooth map that is linear on each $T_x M$ (we say that ω is a V-valued* 1*-form) and surjective. Then $\ker \omega$ is a distribution. It is integrable if and only if $\omega(X) = 0$, $\omega(Y) = 0 \Rightarrow \mathrm{d}\omega(X, Y) = 0$.*

This uses the exterior differential of the next section.

Proof. In local coordinates on M and in a basis of V, the 1-form ω is given by a matrix of maximal rank. We may assume that the left most square block is regular in a neighbourhood of a given point and use the Gauss elimination to make this matrix $(E \mid A)$. Then $\ker \omega$ is given by $(\xi^{n-k+1}, \dots, \xi^n)^T = -A(\xi^1, \dots, \xi^{n-k})$, proving that it is a (smooth) distribution. Now $\mathrm{d}\omega(X, Y) = X\omega(Y) - Y\omega(X) - \omega([X, Y])$ gives easily the result. ∎

12. Cotangent bundle

Every smooth function $f \in C^\infty M$ defines a mapping $\mathrm{d}f \colon TM \to \mathbb{R}$, called the *differential* of f at x, given by

$$\mathrm{d}f(A) = Af$$

i.e. its values are the various directional derivatives of f along various tangent vectors. The restriction to each tangent space $T_x M$ is then a linear form $\mathrm{d}f|_x \in (T_x M)^*$, i.e. an element of the cotangent space $T^*_x M = (T_x M)^*$.

We denote $T^* M = \bigsqcup_{x \in M} T^*_x M$ and call it the *cotangent bundle* of M. Its elements are called *cotangent vectors*. We will soon equip the cotangent bundle $T^* M$ with a structure of a smooth manifold. At this point, we want to rephrase the differential: it can be viewed as a mapping $\mathrm{d}f \colon M \to T^* M$, $x \mapsto \mathrm{d}f|_x \in T^*_x M$. Again, it is a field in the sense that it maps each point to an object of the appropriate vector space.

In coordinates, for $A = A^i \partial_i|_x$, we get

$$\mathrm{d}f|_x(A) = Af = (A^i \partial_i|_x) f = A^i \cdot \partial_i|_x f,$$

(the row of partial derivatives of f, as expected). As a special case, we may apply this to the coordinate functions x^i to get

$$\mathrm{d}x^i|_x(A) = A^j \cdot \partial_j|_x x^i = A^i$$

so that we may rewrite the formula for $\mathrm{d}f$ as $\mathrm{d}f|_x(A) = \mathrm{d}x^i|_x(A) \cdot \partial_i|_x f$, or simply

$$\mathrm{d}f = \partial_i f \cdot \mathrm{d}x^i.$$

This formula is well known from calculus, but now it has an exact meaning, the differentials are sections of cotangent bundles, partial derivatives are functions and both the product and the (implicit) sum make sense.

We will now study how cotangent vectors transform along a smooth map $\varphi\colon M \to N$. Namely, for $x \in M$, the tangent map is $\varphi_{*x}\colon T_xM \to T_{\varphi(x)}N$ and thus induces a dual linear map

$$(\varphi_{*x})^*\colon T^*_{\varphi(x)}N \to T^*_xM.$$

When φ is a local diffeomorphism its inverse is then a linear map $\varphi_{*x}\colon T^*_xM \to T^*_{\varphi(x)}N$. For varying $x \in M$, these form a map

$$\varphi_*\colon T^*M \to T^*N.$$

We will now show that this map is smooth for M and N open subsets of $\mathbb{R}^m$. First, we have

$$(\varphi_{*x})^*(\varphi(x), \eta) = (x, \eta \circ \mathrm{D}|_x\varphi)$$

and thus

$$\varphi_*(x, \theta) = (\varphi(x), \eta \circ (\mathrm{D}|_x\varphi)^{-1}).$$

Since the composition (matrix multiplication) is smooth and so is the map $\mathrm{GL}(m) \to \mathrm{GL}(m)$, $A \mapsto A^{-1}$, the result follows.

Since the transition maps are smooth, we may use the induced maps $\varphi_*\colon T^*M \dashrightarrow T^*\mathbb{R}^m$ to give a smooth atlas for T^*M.

Definition 12.1. A 1-form on M is a smooth map $\omega\colon M \to T^*M$ such that $\omega|_x \in T^*_xM$. The set of all 1-forms on M will be denoted Ω^1M.

In local coordinates, we have $\omega = \omega_i \cdot \mathrm{d}x^i$ and the smoothness is equivalent to the functions ω_i being smooth, since in the charts φ and φ_* the coordinate expression is $\omega(x) = (x, (\omega_1(x), \cdots, \omega_m(x)))$. In particular, for every smooth function f, its differential $\mathrm{d}f$ is a 1-form.

An important feature of forms is that they pull back along smooth maps: for a 1-form ω on N and a smooth map $\varphi\colon M \to N$, we get a 1-form $\varphi^*\omega$ on M given by

$$(\varphi^*\omega)|_x = (\varphi_{*x})^*(\omega|_{\varphi(x)}) = \omega|_{\varphi(x)} \circ \varphi_{*x}.$$

It will be useful to give a local formula for the pull back: if $\omega = \omega_j \cdot \mathrm{d}y^j$, we obtain from the following lemma that

$$\varphi^*\omega = \varphi^*\omega_j \cdot \mathrm{d}(\underbrace{\varphi^*y^j}_{\varphi^j}) = \omega_j \circ \varphi \cdot \partial_i\varphi^j \cdot \mathrm{d}x^i.$$

(writing $y^j = \varphi^j$ the essential part of the formula becomes very intuitive: $\mathrm{d}y^j = \partial y^j/\partial x^i \cdot \mathrm{d}x^i$.)

Lemma 12.2. *$\varphi^*\mathrm{d}f = \mathrm{d}(\varphi^*f)$ and $\varphi^*(f \cdot \omega) = \varphi^*f \cdot \varphi^*\omega$.*

Proof. The first equality is just the definition of the push forward,

$$(\varphi_*A)f = A(f \circ \varphi),$$

rewritten in terms of the differential – the left hand side is $\mathrm{d}f(\varphi_*A) = (\varphi^*\mathrm{d}f)(A)$ and the right hand side is $\mathrm{d}(f \circ \varphi)(A) = \mathrm{d}(\varphi^*f)(A)$.

The second equality is straightforward using the linearity of the dual map:

$$(\varphi^*(f \cdot \omega))|_x = (\varphi_{*x})^*(f(\varphi(x)) \cdot \omega|_{\varphi(x)}) = f(\varphi(x)) \cdot (\varphi_{*x})^*(\omega|_{\varphi(x)}) = (\varphi^*f \cdot \varphi^*\omega)|_x \qquad \square$$

Remark. We have the following relation of $\mathrm{d}f(A) = Af$ to $f_*A \in T_{f(x)}\mathbb{R}$

$$Af = A(\mathrm{id} \circ f) = (f_*A)\,\mathrm{id}$$

where, clearly, $(t, \tau)\,\mathrm{id} = \tau$, i.e. taking the derivative of the identity makes from a based vector the corresponding free vector.

13. Tensor fields

A common generalization of vector fields and 1-forms are the so called tensor fields. They are associations

$$x \mapsto \omega|_x \in \bigotimes^r T_xM \otimes \bigotimes^s T_x^*M$$

that are smooth in a sense similar to that of vector fields and 1-forms.

This is again achieved by defining a smooth manifold

$$\bigotimes^r TM \otimes \bigotimes^s T^*M = \bigsqcup_{x\in M} \bigotimes^r T_xM \otimes \bigotimes^s T_x^*M$$

called the *tensor bundle*. Again, this is done via coordinate charts $\varphi\colon M \dashrightarrow \mathbb{R}^m$ on M by using their induced maps on tensor bundles

$$\varphi_{*x} = \bigotimes^r \varphi_{*x} \otimes \bigotimes^s \varphi_{*x}$$

(the two maps φ_{*x} are different – the first is for the tangent bundle and the second for the cotangent bundle). Again, it is easy to see that these are smooth for open subsets $U \subseteq \mathbb{R}^m$ where

$$\bigotimes^r TU \otimes \bigotimes^s T^*U = U \times \bigotimes^r \mathbb{R}^m \otimes \bigotimes^s (\mathbb{R}^m)^*.$$

The corresponding field is then called a *tensor field* of type (r,s). Of special importance are tensor fields of type $(0,k)$, i.e. those with values in the tensor power of the cotangent bundle, since they again pull back along smooth maps: if $\varphi\colon M \to N$ is a smooth map and $\omega\colon N \to \bigotimes^k T^*N$ a tensor field then

$$(\varphi^*\omega)|_x = \omega|_{\varphi(x)} \circ (\varphi_{*x})^{\otimes k}$$

Since $(T_x^*M)^{\otimes k}$ is naturally isomorphic to the vector space $\mathrm{Lin}_k(T_xM,\dots,T_xM;\mathbb{R})$ of k-linear forms on T_xM, every tensor field ω of type $(0,k)$ can be also seen as a collection of k-linear forms $\omega|_x$ on T_xM and we will use $\omega(A_1,\dots,A_k)$ to denote the values on a k-tuple of tangent vectors, necessarily in a single tangent space T_xM. We may then rewrite the above definition of the pull back of ω as

$$(\varphi^*\omega)(A_1,\dots,A_k) = \omega(\varphi_*A_1,\dots,\varphi_*A_k).$$

We will now explain in more detail how we view a tensor product of 1-forms as a k-linear form:

$$\omega^1 \otimes \cdots \otimes \omega^k(A_1,\dots,A_k) = \omega^1(A_1)\cdots\omega^k(A_k).$$

In particular, we may write every tensor field of type $(0,k)$ locally as

$$\omega = \omega_{i_1\cdots i_k} \cdot \mathrm{d}x^{i_1} \otimes \cdots \otimes \mathrm{d}x^{i_k}$$

and then we get

$$\omega(A_1,\dots A_k) = \omega_{i_1\cdots i_k}(A_1)^{i_1}\cdots(A_k)^{i_k}.$$

We clearly have $\varphi^*(\omega\otimes\theta) = \varphi^*\omega \otimes \varphi^*\theta$ and thus the pull back $\varphi^*\omega$ has the form

$$\varphi^*\omega = (\omega_{i_1\cdots i_k}\circ\varphi)\cdot\partial_{j_1}\varphi^{i_1}\cdots\partial_{j_k}\varphi^{i_k}\cdot \mathrm{d}y^{j_1}\otimes\cdots\otimes\mathrm{d}y^{j_k}.$$

An important special case are the antisymmetric tensor fields of type $(0,k)$, also called (exterior) k-forms. We adopt the convention for $\omega^i \in \Omega^{k_i}M$:

$$\omega^1 \wedge \cdots \wedge \omega^r = \tfrac{(k_1+\cdots+k_r)!}{k_1!\cdots k_r!}\cdot \mathrm{Alt}(\omega^1\otimes\cdots\otimes\omega^r).$$

It is not difficult to check that this wedge product is associative. The advantage of the factor stems from the fact that it eliminates the appearance of such factors in many subsequent formulas (there are also some more subtle advantages that we will not discuss here). Also, for $k_i = 1$, i.e. for 1-forms ω^i, the formula becomes

$$\omega^1\wedge\cdots\wedge\omega^r = \sum_{\sigma\in\Sigma_r}\mathrm{sign}(\sigma)\cdot\omega^{\sigma(1)}\otimes\cdots\otimes\omega^{\sigma(r)}.$$

The evaluation at a k-tuple of vectors is thus

$$\omega^1\wedge\cdots\wedge\omega^k(A_1,\dots,A_k) = \sum_{\sigma\in\Sigma_k}\mathrm{sign}(\sigma)\cdot\omega^1(A_{\sigma(1)})\cdots\omega^k(A_{\sigma(k)}).$$

Applying these formulas to the 1-forms $\mathrm{d}x^i$ results, for

$$\omega = \omega_{i_1\cdots i_k} \cdot \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k},$$

in the following

$$\omega(A_1, \ldots, A_k) = \sum_{\sigma\in\Sigma_k} \operatorname{sign}(\sigma) \cdot \omega_{i_1\cdots i_k} \cdot (A_{\sigma(1)})^{i_1} \cdots (A_{\sigma(k)})^{i_k}.$$

14. Exterior differential

14.1. **Differentiation of tensor fields on a vector space.** A tensor field of type (r, s) on an open subset of a vector space $V = \mathbb{R}^m$ may be interpreted as a map

$$\omega \colon V \dashrightarrow \bigotimes^r V \otimes \bigotimes^s V^*$$

and as such may be differentiated along any vector field X, giving another tensor field $D_X\omega$ of type (r, s). The total derivative $D\omega$ is then a map

$$D\omega \colon V \dashrightarrow \operatorname{Hom}(V, \bigotimes^r V \otimes \bigotimes^s V^*) \cong \bigotimes^r V \otimes \bigotimes^{s+1} V^*,$$

i.e. a tensor field of type $(r, s + 1)$. In coordinates, this consists simply of differentiating the coordinate functions of the tensor field, i.e. for

$$\omega = \omega^{j_1\cdots j_r}_{i_1\cdots i_s} \cdot \partial_{j_1} \otimes \cdots \otimes \partial_{j_r} \otimes \mathrm{d}x^{i_1} \otimes \cdots \otimes \mathrm{d}x^{i_s}$$

we obtain

$$(D_X\omega)^{j_1\cdots j_r}_{i_1\cdots i_s} = \partial_{i_0}\omega^{j_1\cdots j_r}_{i_1\cdots i_s} \cdot X^{i_0},$$

so that we obtain the final formula

$$(D\omega)^{j_1\cdots j_r}_{i_0 i_1\cdots i_s} = \partial_{i_0}\omega^{j_1\cdots j_r}_{i_1\cdots i_s}.$$

14.2. **Differentiating vector fields along paths.** We concentrate in this paragraph on the case of vector fields Y, since this is the only case needed later. We observe that $D_X Y|_x$ depends only on the value $X|_x$ and on the values of Y “in the direction of $X|_x$”. More concretely, if $X|_x = \gamma'(t_0)$ and if we denote $Z = Y \circ \gamma$, then $D_X Y$ depends only on Z:

$$D_{\gamma'(t_0)}Y = D_{\gamma'(t_0)}Y^j \cdot \partial_j = \partial_t|_{t_0}(Y^j \circ \gamma) \cdot \partial_j = \partial_t|_{t_0} Z^j \cdot \partial_j = D_t Z|_{t_0}$$

or more concisely $D_{\gamma'} Y = D_t Z$.

More generally, a *vector field along* γ is a smooth map $Z \colon \mathbb{R} \dashrightarrow TM$ such that $Z|_t \in T_{\gamma(t)}M$. Then the formula

$$D_t Z|_{t_0} = \partial_t|_{t_0} Z^j \cdot \partial_j$$

as above defines, for an arbitrary vector field Z along γ, another vector field $D_t Z$ along γ. In particular, the tangent vectors $\gamma' = \partial_t\gamma^i \cdot \partial_i$ form a vector field along γ and we get

$$D_t\gamma' = \gamma'' = \partial^2_{tt}\gamma^i \cdot \partial_i.$$

14.3. **Non-invariance of** D**.** Let us study the invariance of D under the change of coordinates, i.e. let ω be a tensor field of type $(0, k)$ with components $\omega_{i_1\cdots i_k}$,

$$\omega = \omega_{i_1\cdots i_k} \cdot \mathrm{d}x^{i_1} \otimes \cdots \otimes \mathrm{d}x^{i_k},$$

and recall that $(D\omega)_{i_0 i_1\cdots i_k} = \partial_{i_0}\omega_{i_1\cdots i_k}$. Now we apply the change of coordinates φ to get

$$(\varphi^*(D\omega))_{j_0 j_1\cdots j_k} = (\partial_{i_0}\omega_{i_1\cdots i_k} \circ \varphi) \cdot \partial_{j_0}\varphi^{i_0} \cdot \partial_{j_1}\varphi^{i_1} \cdot \cdots \cdot \partial_{j_k}\varphi^{i_k}.$$

On the other hand $D(\varphi^*\omega)$ equals

$$\begin{aligned}
(D(\varphi^*\omega))_{j_0 j_1\cdots j_k} &= \partial_{j_0}((\omega_{i_1\cdots i_k} \circ \varphi) \cdot \partial_{j_1}\varphi^{i_1} \cdot \cdots \cdot \partial_{j_k}\varphi^{i_k}) \\
&= \partial_{j_0}((\omega_{i_1\cdots i_k} \circ \varphi)) \cdot \partial_{j_1}\varphi^{i_1} \cdot \cdots \cdot \partial_{j_k}\varphi^{i_k} \\
&\quad + \sum_r (\omega_{i_1\cdots i_k} \circ \varphi) \cdot \partial_{j_1}\varphi^{i_1} \cdot \cdots \cdot \partial^2_{j_0 j_r}\varphi^{i_r} \cdot \cdots \cdot \partial_{j_k}\varphi^{i_k}
\end{aligned}$$

The first term equals $\varphi^*(D\omega)_{j_0 j_1\cdots j_k}$ by the chain rule. Now the point is that in order to get $\varphi^*(D\omega) = D(\varphi^*\omega)$, i.e. to get a differential that does not depend on coordinates as we will see shortly, we have to get rid of the second term involving the second derivative $\partial^2_{j_0 j_r}\varphi^{i_r}$. It clearly disappears after antisymmetrization. Some representation theory would be required to get that no "other part" is invariant and we will not attempt to do this here.

Thus, we get an invariant differentiation operator – the *exterior differential* – on antisymmetric forms by antisymmetrizing $D_{X_0}\omega(X_1,\dots,X_k)$; for technical reasons, we multiply the antisymmetrization by $\frac{(k+1)!}{1!k!}$, since the form already was antisymmetric in the variables $X_1,\dots,X_k$ and obtain

$$\mathrm{d}\omega(X_0,\dots,X_k)=\sum_i(-1)^i D_{X_i}\omega(X_0,\dots,\widehat{X_i},\dots,X_k)$$

Starting from the formula for D,

$$\mathrm{D}\omega=\partial_{i_0}\omega_{i_1\cdots i_k}\mathrm{d}x^{i_0}\otimes\mathrm{d}x^{i_1}\wedge\cdots\wedge\mathrm{d}x^{i_k},$$

we obtain a formula for d by antisymmetrization, i.e.

$$\mathrm{d}\omega=\partial_{i_0}\omega_{i_1\cdots i_k}\mathrm{d}x^{i_0}\wedge\mathrm{d}x^{i_1}\wedge\cdots\wedge\mathrm{d}x^{i_k}.$$

We will now explain a useful formalism. We denote, for a k-tuple $I=(i_1,\dots,i_k)$,

$$\mathrm{d}x^I=\mathrm{d}x^{i_1}\wedge\cdots\wedge\mathrm{d}x^{i_k}$$

and also $\omega_I=\omega_{i_1\cdots i_k}$ so that we may write

$$\omega=\omega_I\cdot\mathrm{d}x^I$$

where the implicit summation occurs over all ordered k-tuples I. We assume ω_I antisymmetric in I, as usual. We then get a very simple formula for all operations:

$$(\omega_I\cdot\mathrm{d}x^I)\wedge(\theta_J\cdot\mathrm{d}x^J)=\omega_I\theta_J\cdot\mathrm{d}x^{IJ}$$

and most importantly

$$\mathrm{d}\omega=\underbrace{\partial_k\omega_I\cdot\mathrm{d}x^k}_{\mathrm{d}\omega_I}\wedge\mathrm{d}x^I=\mathrm{d}\omega_I\wedge\mathrm{d}x^I.$$

Proposition 14.1. *The exterior differential satisfies $\varphi^*(\mathrm{d}\omega)=\mathrm{d}(\varphi^*\omega)$ and, therefore, it induces an operator $\mathrm{d}\colon\Omega^kM\to\Omega^{k+1}M$ on any smooth manifold.*

Proof. The invariance was proved earlier. This enables to define $\mathrm{d}\omega$ in the domain of a chart φ_i by the formula $\mathrm{d}\omega=(\varphi_i)^*\mathrm{d}((\varphi_i^{-1})^*\omega)$, i.e. ω is translated to the chart, differentiated there and then translated back to M. For any other chart φ_j, we have $\varphi_j=\varphi_{ij}\circ\varphi_i$ and then

$$(\varphi_j)^*\mathrm{d}((\varphi_j^{-1})^*\omega)=(\varphi_i)^*\underbrace{(\varphi_{ij})^*\mathrm{d}((\varphi_{ij}^{-1})^*}_{\mathrm{d}}(\varphi_i^{-1})^*\omega)=(\varphi_i)^*\mathrm{d}((\varphi_i^{-1})^*\omega)$$

by the invariance with respect to the transition map φ_{ij}, wherever both sides are defined. □

Theorem 14.2. *The exterior differential satisfies the following properties:*

(1) *the exterior differential of a 0-form, i.e. a function f, is the usual differential,*
(2) $\mathrm{d}(\omega+\theta)=\mathrm{d}\omega+\mathrm{d}\theta$,
(3) $\mathrm{d}(\omega\wedge\theta)=\mathrm{d}\omega\wedge\theta+(-1)^{|\omega|}\cdot\omega\wedge\mathrm{d}\theta$,
(4) $\mathrm{d}(\mathrm{d}\omega)=0$.

It is the unique invariant operator satisfying these properties.

Proof. The first point is clear from the definition and so is the second. For the third point, write $\omega=\omega_I\cdot\mathrm{d}x^I$, $\theta=\theta_J\cdot\mathrm{d}x^J$ and thus we get

$$\begin{aligned}\mathrm{d}(\omega\wedge\theta)&=\mathrm{d}(\omega_I\theta_J\cdot\mathrm{d}x^I\wedge\mathrm{d}x^J)\\&=(\partial_k\omega_I\cdot\theta_J+\omega_I\cdot\partial_k\theta_J)\cdot\mathrm{d}x^k\wedge\mathrm{d}x^I\wedge\mathrm{d}x^J\\&=(\partial_k\omega_I\cdot\mathrm{d}x^k\wedge\mathrm{d}x^I)\wedge(\theta_J\cdot\mathrm{d}x^J)+(-1)^{|\omega|}\cdot(\omega_I\cdot\mathrm{d}x^I)\wedge(\partial_k\theta_J\cdot\mathrm{d}x^k\wedge\mathrm{d}x^J),\end{aligned}$$

where the sign comes from interchanging $\mathrm{d}x^k$ with $\mathrm{d}x^I$.

First observe

$$\mathrm{d}(\mathrm{d}f)=\mathrm{d}(\partial_i f\cdot\mathrm{d}x^i)=\partial_j\partial_i f\cdot\mathrm{d}x^j\wedge\mathrm{d}x^i$$

and, for each i and j, the term $\partial_{ji}f \cdot \mathrm{d}x^{ji}$ cancels with the term $\partial_{ij}f \cdot \mathrm{d}x^{ij}$ since the partial derivatives are symmetric, while $\mathrm{d}x^{ij} = -\mathrm{d}x^{ji}$. By the Leibniz rule, we then get

$$\begin{aligned}\mathrm{d}(\mathrm{d}x^I) &= \mathrm{d}(\mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}) \\ &= \sum (-1)^{r-1} \cdot \mathrm{d}x^{i_1} \wedge \cdots \wedge \underbrace{\mathrm{d}(\mathrm{d}x^{i_r})}_{0} \wedge \cdots \wedge \mathrm{d}x^{i_k} = 0.\end{aligned}$$

Consequently, for general $\omega = \omega_I \cdot \mathrm{d}x^I$, we get

$$\mathrm{d}(\mathrm{d}\omega) = \mathrm{d}(\mathrm{d}\omega_I \cdot \mathrm{d}x^I) = \mathrm{d}(\mathrm{d}\omega_I) \wedge \mathrm{d}x^I - \mathrm{d}\omega_I \wedge \mathrm{d}(\mathrm{d}x^I) = 0 - 0 = 0. \qquad \square$$

14.4. **Coordinate-free formula for the exterior differential.** The differentiation operator D_{X_0} satisfies the Leibniz rule

$$\begin{aligned}D_{X_0}(\omega_{i_1\cdots i_k}(X_1)^{i_1}\cdots(X_k)^{i_k}) = D_{X_0}\omega_{i_1\cdots i_k} \cdot (X_1)^{i_1}\cdots(X_k)^{i_k} \\ + \sum \omega_{i_1\cdots i_k}(X_1)^{i_1}\cdots D_{X_0}(X_j)^{i_j}\cdots(X_k)^{i_k}\end{aligned}$$

which (after subtracting the sum from the right hand side) translates to

$$(D_{X_0}\omega)(X_1, \ldots, X_k) = D_{X_0}(\omega(X_1, \ldots, X_k)) - \sum \omega(X_1, \ldots, X_{j-1}, D_{X_0}X_j, X_{j+1}, \ldots, X_k).$$

Here the first D_{X_0} on the right is the directional derivative of the function $\omega(X_1, \ldots, X_k)$. The second appearance is, however, very different and we have $[X, Y] = D_XY - D_YX$. The exterior differential then equals

$$\begin{aligned}\mathrm{d}\omega(X_0, \ldots, X_k) &= \sum_i (-1)^i (D_{X_i}\omega)(X_0, \ldots, \widehat{X_i}, \ldots, X_k) \\ &= \sum_i (-1)^i X_i\omega(X_0, \ldots, \widehat{X_i}, \ldots, X_k) \\ &\quad - \sum_{i<j} (-1)^i \omega(X_0, \ldots, \widehat{X_i}, \ldots, X_{j-1}, D_{X_i}X_j, X_{j+1}, \ldots, X_k) \\ &\quad - \sum_{i>j} (-1)^i \omega(X_0, \ldots, X_{j-1}, D_{X_i}X_j, X_{j+1}, \ldots, \widehat{X_i}, \ldots, X_k)\end{aligned}$$

where we split the antisymmetrization of the second term according to whether $i < j$ or $i > j$. Next we move the term $D_{X_i}X_j$ onto the first spot (here the sign differs for the two possibilities):

$$\begin{aligned}&= \sum_i (-1)^i X_i\omega(X_0, \ldots, \widehat{X_i}, \ldots, X_k) \\ &\quad + \sum_{i<j} (-1)^{i+j} \omega(D_{X_i}X_j, X_0, \ldots, \widehat{X_i}, \ldots, \widehat{X_j}, \ldots, X_k) \\ &\quad - \sum_{i>j} (-1)^{i+j} \omega(D_{X_i}X_j, X_0, \ldots, \widehat{X_j}, \ldots, \widehat{X_i}, \ldots, X_k)\end{aligned}$$

and finally swap the indices i, j in the last sum and subtract, using $D_{X_i}X_j - D_{X_j}X_i = [X_i, X_j]$, to obtain the final formula

$$\begin{aligned}&= \sum_i (-1)^i X_i\omega(X_0, \ldots, \widehat{X_i}, \ldots, X_k) \\ &\quad + \sum_{i<j} (-1)^{i+j} \omega([X_i, X_j], X_0, \ldots, \widehat{X_i}, \ldots, \widehat{X_j}, \ldots, X_k)\end{aligned}$$

15. Integration of forms

15.1. **Partitions of unity.**

Definition 15.1. Let $\mathcal{U}$ be an open cover of a manifold M. A *(smooth) partition of unity subordinate to* $\mathcal{U}$ is a collection of functions $\lambda_U \colon M \to [0, 1]$, for $U \in \mathcal{U}$, such that $\operatorname{supp} \lambda_U \subseteq U$, such that, in a neighbourhood of every point $x \in M$, there is only a finite number of non-zero λ_U and such that $\sum_{U \in \mathcal{U}} \lambda_U = 1$.

For simplicity, we will assume M to be compact, it is however not necessary. Then the local finiteness of the λ_U in the definition of the partition of a unity is translated to the finiteness, i.e. only a finite number of the λ_U is nonzero.

Theorem 15.2. *Let M be compact. There exist a partition of unity subordinate to any open cover of M.*

Proof. For each $x \in M$, choose $U_x \in \mathcal{U}$ with $x \in U_x$ and a function $\lambda_x\colon M \to [0,1]$ with $\lambda_x(x) > 0$ and $\operatorname{supp} \lambda_x \subseteq U_x$. Clearly the open sets $V_x = \{y \in M \mid \lambda_x(y) > 0\}$ cover M (since $V_x \ni x$) and thus, by compactness, there is a finite subcover $M = V_{x_1} \cup \cdots \cup V_{x_k}$. This means that $\lambda = \lambda_{x_1} + \cdots + \lambda_{x_k} > 0$ on M. Take $\lambda_U = (\sum_{U_{x_i} = U} \lambda_{x_i})/\lambda$. □

Corollary 15.3. *There exists a Riemannian metric on every manifold.*

Proof. We choose a Riemannian metric g_U arbitrarily in every coordinate chart U. Using a partition of unity λ_U (assuming that M is compact or using the non-compact version of the previous theorem), we set $g = \sum_{U \in \mathcal{U}} \lambda_U g_U$. □

15.2. **Orientation.** An *orientation* of a manifold is an orientation of each tangent space T_xM that is "smooth" in the following sense: for every m-tuple $\alpha = (X_1, \ldots, X_m)$ of local vector fields that form a basis of T_xM where defined (the so-called local *frame*), the function $\operatorname{sign} \alpha$ is smooth. Since it takes values in $\{\pm 1\}$, it must in fact be locally constant.

Let $\varphi\colon M \to N$ be a local diffeomorphism. Then, for each $x \in M$, the tangent map $\varphi_{*x}\colon T_xM \to T_{\varphi(x)}$ is an isomorphism and we define $\operatorname{sign} \varphi_{*x}$ to be $+1$ if φ_{*x} preserves orientation and -1 if it reverses orientation. In formula,

$$\operatorname{sign}(\varphi_{*x}\alpha|_x) = \operatorname{sign} \varphi_{*x} \cdot \operatorname{sign} \alpha|_x$$

For each (local) frame β on N, there is a (local) frame $\alpha = \varphi^*\beta$ on M (consisting of the pullbacks $(\varphi^*Y)|_x = (\varphi_{*x})^{-1}Y|_{\varphi(x)}$) so that the two frames are φ-related. Thus,

$$\operatorname{sign} \beta|_{\varphi(x)} = \operatorname{sign}(\varphi_{*x}\alpha|_x) = \operatorname{sign} \varphi_{*x} \cdot \operatorname{sign} \alpha|_x$$

and since both $\operatorname{sign} \alpha|_x$ and $\operatorname{sign} \beta|_{\varphi(x)}$ are locally constant, so is $\operatorname{sign} \varphi_{*x}$. In particular, if M is connected, then φ either preserves orientation at every point or it reverses orientation at every point.

In particular, every connected chart either preserves orientation or its composition with a reflection $\mathbb{R}^m \to \mathbb{R}^m$ preserves orientation. Consequently, the collection of all orientation preserving charts forms an atlas – the maximal oriented atlas on M. For this atlas, the transition maps also preserve orientation. (There is a special case $m = 0$, where no reflection exists and, thus, it is not always possible to get an oriented atlas.)

15.3. **Integral.** Let ω be an m-form on an open subset $V \subseteq \mathbb{R}^m$ and we assume for simplicity that it has compact support. Writing

$$\omega = a \cdot \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^m,$$

i.e. $a = \omega(\partial_1, \ldots, \partial_m)$, we define

$$\int_V \omega = \int \!\cdots\! \int_V a(x)\ \mathrm{d}x^1 \cdots \mathrm{d}x^m.$$

Clearly, the above defined integral is additive in ω.

For a smooth map $\varphi\colon \mathbb{R}^m \dashrightarrow \mathbb{R}^m$, we get

$$\begin{aligned} \varphi^*\omega(y) &= a \circ \varphi(y) \cdot \varphi_y^*\mathrm{d}x^1 \wedge \cdots \wedge \varphi_y^*\mathrm{d}x^m \\ &= a \circ \varphi(y) \cdot \det \varphi_{*y} \cdot \mathrm{d}y^1 \wedge \cdots \wedge \mathrm{d}y^m. \end{aligned}$$

with $\det \varphi_{*y}$ the Jacobian; thus, for a diffeomorphism $\varphi\colon W \to V$ with *positive Jacobian*, we get

$$\int_W \varphi^*\omega = \int_V \omega.$$

Let M be an oriented manifold of dimension m. We assume for simplicity that M is compact. Consider the maximal *oriented* atlas on M and choose a partition of unity λ_i so that $\operatorname{supp}\lambda_i$ is a subset of a domain U_i of a chart $\varphi_i\colon U_i \to V_i$. Let ω be an m-form on M. Then we define

$$\int_M \omega = \sum_i \int_{V_i} (\varphi_i^{-1})^*(\lambda_i\omega)$$

We note that $\lambda_i\omega$ has a compact support inside U_i and, thus, the pullback $(\varphi_i^{-1})^*(\lambda_i\omega)$ has a compact support inside V_i so that the integral exists and is finite. It remains to show that it does not depend on the choice of the partition λ_i.

Thus, let μ_i be another partition. Then we get

$$\sum_i \int_{V_i} (\varphi_i^{-1})^*(\lambda_i\omega) = \sum_{i,j} \int_{V_i} (\varphi_i^{-1})^*(\lambda_i\mu_j\omega)$$

Denoting $\theta = \lambda_i\mu_j\omega$, a compactly supported m-form inside $U_i \cap U_j$, and further $V_{ij} = \varphi_i(U_i \cap U_j)$ and $V_{ji} = \varphi_j(U_i \cap U_j)$, we have

$$\int_{V_i} (\varphi_i^{-1})^*\theta = \int_{V_{ij}} (\varphi_i^{-1})^*\theta = \int_{V_{ij}} \varphi_{ij}^*((\varphi_j^{-1})^*\theta) = \int_{V_{ji}} (\varphi_j^{-1})^*\theta = \int_{V_j} (\varphi_j^{-1})^*\theta$$

by the invariance of the integral with respect to the diffeomorphism $\varphi_{ij} = \varphi_j \circ \varphi_i^{-1}\colon V_{ij} \to V_{ji}$.

By the additivity of the integral, we may reformulate the procedure: write the m-form ω as a finite sum $\omega = \sum \omega^i$ of m-forms with each ω^i concentrated in an oriented coordinate patch. Then $\int \omega = \sum \int \omega^i$ and each of the integrals is computed in coordinates,

$$\int \omega^i = \int \cdots \int \omega^i(\partial_1, \dots, \partial_m)\, \mathrm{d}x^1 \cdots \mathrm{d}x^m.$$

15.4. **Manifolds with boundary.** The main idea here is: in exactly the same manner in which manifolds are built from the Euclidean space $\mathbb{R}^m$, manifolds with boundary are built from the Euclidean halfspace $\mathbb{H}^{m+1} = \{(x^0, x^1, \dots, x^m) \in \mathbb{R}^{m+1} \mid x^0 \leq 0\}$. It is however important that we allow tangent vectors at the boundary hyperplane to be all vectors from $\mathbb{R}^{m+1}$, i.e.

$$T\mathbb{H}^{m+1} = \mathbb{H}^{m+1} \times \mathbb{R}^{m+1}.$$

Thus, the geometric definition using paths is inappropriate. Derivations work well if we interpret $\partial_0 f(x)$ for a boundary point x to be the one-sided partial derivative.

Formally, a map between open subsets of the half-spaces is said to be smooth, if all partial derivatives exist (one-sided where needed) and are continuous. A diffeomorphism between open subsets of $\mathbb{H}^{m+1}$ preserves the boundary points, since at an interior point, any (local) diffeomorphism has a local inverse and as such maps to an interior point.

With this notion, we define a (smooth) *manifold with boundary* W as a topological space, Hausdorff and with countable basis of topology, equipped with a maximal atlas consisting of homeomorphisms $\varphi\colon U \to V$ with V an open subset of $\mathbb{H}^{m+1}$ and with all change of coordinate maps smooth in the above sense. We define the *boundary* of W to be the set ∂W of points that correspond to the boundary points in a chart (equivalently, in all charts).

The standard bases $(e_0, e_1, \dots, e_m)$ of $\mathbb{H}^{m+1}$ and $(e_1, \dots, e_m)$ of $\mathbb{R}^n$ are considered positive. We say that $\partial\mathbb{H}^{m+1}$ is oriented via its outward normal: The outward normal is by definition $u_0 = e_0$ (or any combination $u_0 = x^0e_0 + x^1e_1 + \cdots x^me_m$ with $x^0 > 0$) and a basis $(u_1, \dots, u_m)$ is then a positive basis of $\partial\mathbb{H}^{m+1}$ according to this principle if and only if $(u_0, u_1, \dots, u_m)$ is a positive basis of $\mathbb{H}^{m+1}$. This gives a way of orienting the boundary ∂W of any oriented manifold with boundary W. We will always consider ∂W with this induced orientation.

15.5. **Stokes' theorem.**

Theorem 15.4. *For a compact manifold with boundary W of dimension $m+1$ and an m-form ω on W, we have*

$$\int_{\partial W} \omega = \int_W \mathrm{d}\omega.$$

(The left hand side is really the integral of the pullback $j^*\omega$ along the inclusion $j\colon \partial W \to W$.)

Proof. We may write ω as a sum of m-forms supported in a coordinate chart and thus reduce to a local situation, i.e. we may assume that $W = \mathbb{H}^{m+1}$. Since ω is an m-form on $\mathbb{H}^{m+1}$, we may write

$$\omega = \sum_i a_i \cdot \mathrm{d}x^0 \wedge \cdots \wedge \widehat{\mathrm{d}x^i} \wedge \cdots \wedge \mathrm{d}x^m.$$

Since $j(x^1, \ldots, x^m) = (0, x^1, \ldots, x^m)$, we get $j^*\mathrm{d}x^0 = 0$ and $j^*\mathrm{d}x^i = \mathrm{d}x^i$, for $i > 0$. Thus,

$$j^*\omega = a_0 \circ j \cdot \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^m$$

and the integral on the left is

$$\int_{\partial\mathbb{H}^{m+1}} j^*\omega = \int \cdots \int_{\mathbb{R}^m} a_0(0, x^1, \ldots, x^m)\ \mathrm{d}x^1 \cdots \mathrm{d}x^m.$$

Now we simplify the integral on the right, i.e. we compute

$$\begin{aligned} \mathrm{d}\omega &= \sum_i \partial_i a_i \cdot \mathrm{d}x^i \wedge \mathrm{d}x^0 \wedge \cdots \wedge \widehat{\mathrm{d}x^i} \wedge \cdots \wedge \mathrm{d}x^m \\ &= \sum_i (-1)^i \partial_i a_i \cdot \mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^m. \end{aligned}$$

Now the integral simplifies to

$$\begin{aligned} \int_{\mathbb{H}^{m+1}} \mathrm{d}\omega &= \sum_i (-1)^i \int \cdots \int_{\mathbb{H}^{m+1}} \partial_i a_i\ \mathrm{d}x^0 \cdots \mathrm{d}x^m \\ &= \sum_i (-1)^i \int \cdots \int_{\mathbb{H}^{m+1}} \partial_i a_i\ \mathrm{d}x^i \mathrm{d}x^0 \cdots \widehat{\mathrm{d}x^i} \cdots \mathrm{d}x^m \end{aligned}$$

For $i > 0$, we get $\int_{-\infty}^{\infty} \partial_i a_i \cdot \mathrm{d}x^i = a_i|_{x^i=\infty} - a_i|_{x^i=-\infty} = 0 - 0 = 0$, since ω is assumed compactly supproted, while $\int_{-\infty}^{0} \partial_i a_i \cdot \mathrm{d}x^0 = a_0|_{x^0=0} - a_0|_{x^0=-\infty} = a_0(0, x^1, \ldots, x^m)$. Thus, the integral also equals

$$\int_{\mathbb{H}^{m+1}} \mathrm{d}\omega = \int \cdots \int_{\mathbb{R}^m} a_0(0, x^1, \ldots, x^m)\ \mathrm{d}x^1 \cdots \mathrm{d}x^m. \qquad \square$$

Remark. It is interesting to see what we would get if we integrated over a cube instead. Then the "boundary conditions" $a_i|_{x^i=\pm\infty} = 0$ would be replaced by the non-zero restrictions to the faces of the cube and the resulting formula would be

$$\int_{I^{m+1}} \mathrm{d}\omega = \sum_i (-1)^i \int_{\partial_i^+ I^{m+1}} \omega - \sum_i (-1)^i \int_{\partial_i^- I^{m+1}} \omega$$

where the $\partial_i^\varepsilon I^{m+1}$ denotes the subset $\{(x^0, \ldots, x^m) \in I^{m+1} \mid x^i = \varepsilon\}$. The signs reflect the orientations of these faces so that the right hand side actually equals $\int_{\partial I^{m+1}} \omega$ when ∂I^{m+1} is interpreted correctly.

15.6. **An interpretation of Stoke's theorem.** First we prove that a k-form ω is uniquely determined by integrals $\int_{D^k} \iota^*\omega$ for arbitrary embeddings $\iota\colon D^k \hookrightarrow M$. Here, D^k is a k-dimensional unit ball. Alternatively, the same holds for cubes. To prove this claim, observe that

$$\int_{D^k} \iota^*\omega = \int \cdots \int_{D^k} \omega(\iota_*\partial_1, \ldots, \iota_*\partial_k)\ \mathrm{d}t^1 \cdots \mathrm{d}t^k$$

where the integrand is the function obtained by evaluating ω at the images of the canonical vector fields on $D^k \subseteq \mathbb{R}^k$. Clearly, this equals roughly

$$\int_{D^k} \iota^*\omega \approx \mathrm{Vol}(D^k) \cdot \omega(\iota_{*0}\partial_1, \ldots, \iota_{*0}\partial_k)$$

Restricting to an ε-ball εD^k, we obtain an equality in limit

$$\lim_{\varepsilon \to 0} \left(\frac{1}{\varepsilon^k} \int_{\varepsilon D^k} \iota^*\omega \right) = \mathrm{Vol}(D^k) \cdot \omega(\iota_{*0}\partial_1, \ldots, \iota_{*0}\partial_k)$$

Thus, if ω integrates to 0 over any embedded k-dimensional ball, then $\omega = 0$. As an immediate consequence, we see that $\omega = 0$ if and only if ω integrates to 0 over any k-dimensional ∂-submanifold $W \subseteq M$.

Now ω is closed if and only if $0 = \int_W \mathrm{d}\omega = \int_{\partial W} \omega$, i.e. if and only if ω integrates to 0 over any k-dimensional boundary. Similarly, if ω is exact, say $\omega = \mathrm{d}\theta$, then for any submanifold $N \subseteq M$ (without boundary!) $\int_N \omega = \int_N \mathrm{d}\theta = \int_{\partial N} \theta = 0$, since $\partial N = \emptyset$. Thus, when $\int_N \omega \neq 0$ for ω closed and N submanifold without boundary, we conclude that ω is not exact and that N is not a boundary (of a compact ∂-submanifold). In particular, $H^k M \neq 0$. (This, in general, is far from an equivalence.)

15.7. **Cohomology in top dimension.** In order to distinguish compact manifolds without boundary from those with boundary, we call them *closed*.

Theorem 15.5. *For any closed oriented Riemannian manifold M of dimension m, $H^m(M) \neq 0$.*

Proof. Since every m-form on M is closed, it is enough to find one that is not exact. We know that oriented Euclidean spaces admit a canonical volume form specified by the requirement $\mathrm{Vol}(e_1, \ldots, e_m) = 1$ for any positive orthonormal basis $(e_1, \ldots, e_m)$. In this way, we obtain a volume form $\mathrm{Vol} \in \Omega^m M$. In any chart compatible with the orientation,

$$\mathrm{Vol} = a \cdot \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^m$$

with $a = \mathrm{Vol}(\partial_1, \ldots, \partial_m) > 0$. Thus, $\int_M \mathrm{Vol}$ is an integral of a positive function and as such must also be positive. Thus, Vol cannot be exact. □

15.8. **Homotopy invariance.** We would like to show that $H^k\mathbb{R}^n = 0$ for $k > 0$. This will follow from the following "homotopy invariance" property.

Theorem 15.6. *Let $h\colon [-1, 1] \times M \to N$ be a smooth map and denote $h_t = h(t, -)$. Then for any closed k-form ω, we get $[h_{-1}^*\omega] = [h_1^*\omega] \in H^k M$.*

Proof. Employing, for $X \in \mathfrak{X}M$ and $\omega \in \Omega^{k+1}M$, the notation

$$X \llcorner \omega(A_1, \ldots, A_k) = \omega(X, A_1, \ldots, A_k),$$

we define a homotopy operator $K\colon \Omega^{k+1}([-1, 1] \times M) \to \Omega^k M$ via

$$K\omega = \int_{-1}^{1} j_t^*(\partial_t \llcorner \omega)\ \mathrm{d}t,$$

where $j_t\colon M \to [-1, 1] \times M$ is the map $x \mapsto (t, x)$. Writing in local coordinates

$$\omega = \omega_{i_1 \cdots i_k}\ \mathrm{d}t \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k} + \omega_{i_0 \cdots i_k}\ \mathrm{d}x^{i_0} \wedge \cdots \wedge \mathrm{d}x^{i_k},$$

we get a formula

$$K\omega = \int_{-1}^{1} j_t^* \omega_{i_1 \cdots i_k}\ \mathrm{d}t \cdot \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}.$$

(the coordinate functions corresponding to terms that involve $\mathrm{d}t$ get integrated along t, the remaining terms disappear). Now we compute

$$\mathrm{d}K\omega = \int_{-1}^{1} j_t^* \partial_{i_0} \omega_{i_1 \cdots i_k}\ \mathrm{d}t \cdot \mathrm{d}x^{i_0} \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}$$

while

$$\begin{aligned} \mathrm{d}\omega = \partial_{i_0} \omega_{i_1 \cdots i_k}\ & \underbrace{\mathrm{d}x^{i_0} \wedge \mathrm{d}t \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}}_{-\mathrm{d}t \wedge \mathrm{d}x^{i_0} \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}} + \partial_t \omega_{i_0 \cdots i_k}\ \mathrm{d}t \wedge \mathrm{d}x^{i_0} \wedge \cdots \wedge \mathrm{d}x^{i_k} \\ & + \text{terms not involving } \mathrm{d}t \end{aligned}$$

and consequently

$$\begin{aligned} K\mathrm{d}\omega &= -\mathrm{d}K\omega + \int_{-1}^{1} j_t^* \partial_t \omega\ \mathrm{d}t \cdot \mathrm{d}x^{i_0} \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k} \\ &= -\mathrm{d}K\omega + j_1^*\omega - j_{-1}^*\omega. \end{aligned}$$

This implies rather easily the result, since, for ω closed, the term on the left vanishes and, thus, the difference $j_1^*\omega - j_{-1}^*\omega = \mathrm{d}K\omega$ is exact, i.e. the two terms represent the same cohomology class; now $h_\varepsilon^*\omega = h^* j_\varepsilon^*\omega$. $\square$

Proof. The idea of the proof is simple. Any k-form is determined by its integrals along k-dimensional cubes embedded in M. This is so because any embedding $[-1,1]^k \to M$ that maps $\partial_i(0)$ to $A_i \in T_xM$ restricts to the cube $[-\varepsilon,\varepsilon]^k$ to an embedding i_ε such that $\int (i_\varepsilon)^*\omega \sim (2\varepsilon)^k\omega(A_1,\dots,A_k)$ (equality holds in $\lim_{\varepsilon\to 0}$).

Now for an embedding $i\colon [-1,1]^k \to M$, we get an associated embedding $\mathrm{id}\times i\colon [-1,1]^{k+1} \to [-1,1]\times M$. Denote by $j_t\colon [-1,1]^k \to [-1,1]^{k+1}$ the embedding given by $j_t(t^1,\dots,t^k) = (t,t^1,\dots,t^k)$. Then $h_t^*\omega = j_t^*h^*\omega$ and both $j_{\pm 1}$ are embeddings as part of the boundary. Thus, the Stokes' theorem relates

$$\begin{aligned}\int_{[-1,1]\times[-1,1]^k} \mathrm{d}(h^*\omega) &= \int_{\partial([-1,1]\times[-1,1]^k)} h^*\omega \\ &= \int_{[-1,1]^k} h_1^*\omega - \int_{[-1,1]^k} h_{-1}^*\omega - \int_{[-1,1]\times\partial[-1,1]^k} h^*\omega \qquad (*)\end{aligned}$$

(the first two terms correspond to $\partial[-1,1]\times[-1,1]^k$). Writing

$$\mathrm{d}(h^*\omega) = a\cdot \mathrm{d}t\wedge \mathrm{d}t^1\wedge\cdots\wedge \mathrm{d}t^k,$$

the integral on the left can be computed using Fubini's theorem as

$$\int_{[-1,1]\times[-1,1]^k} \mathrm{d}(h^*\omega) = \int_{[-1,1]^k}\left(\int_{[-1,1]} a(t,t^1,\dots,t^k)\ \mathrm{d}t\right)\mathrm{d}t^1\cdots\mathrm{d}t^k.$$

This can be rephrased in terms of an operator $K\colon \Omega^{k+1}([-1,1]\times M)\to\Omega^kM$, given by the integral

$$K(\eta)_x(A_1,\dots,A_k) = \begin{cases}\int_{[-1,1]}\theta_{(t,x)}(A_1,\dots,A_k)\ \mathrm{d}t & \eta = \mathrm{d}t\wedge\theta \\ 0 & \eta(\partial_t,-,\dots,-)=0\end{cases}$$

(formally, there is an isomorphism $\Lambda^k(V\oplus W)\cong\bigoplus_{i+j=k}\Lambda^iV\otimes\Lambda^jW$; apply this to the canonical decomposition $T_{(t,x)}([-1,1]\times M)\cong T_t[-1,1]\oplus T_xM$; then K is defined by projecting to $\Lambda^1T_t^*[-1,1]\otimes\Lambda^kT_x^*M$, writing the image uniquely as $\mathrm{d}t\otimes\theta$ and then integrating θ as above) as

$$\int_{[-1,1]\times[-1,1]^k}\mathrm{d}(h^*\omega) = \int_{[-1,1]^k} K\mathrm{d}(h^*\omega) = \int_{[-1,1]^k} K(h^*\mathrm{d}\omega).$$

The remaining boundary term in $(*)$ is then

$$\int_{[-1,1]\times\partial[-1,1]^k} h^*\omega = \int_{\partial[-1,1]^k} K(h^*\omega) = \int_{[-1,1]^k}\mathrm{d}K(h^*\omega),$$

again by the Stokes' theorem. Thus, we have finally obtained

$$\int_{[-1,1]^k} h_1^*\omega - \int_{[-1,1]^k} h_{-1}^*\omega = \int_{[-1,1]^k}(\mathrm{d}K(h^*\omega)+K(h^*\mathrm{d}\omega)))$$

or, in other words, $h_1^*\omega - h_{-1}^*\omega = \mathrm{d}K(h^*\omega)+K(h^*\mathrm{d}\omega)$. This implies rather easily the result, since, for ω closed, the first term on the right vanishes and, thus, the difference on the left is exact, i.e. the two terms represent the same cohomology class.

In the situation from the above proof, we say that two chain maps (maps that commute with differentials, such as pullback maps j_ε^*) are *chain homotopic* if there exists a collection of maps η such that

$$g - f = \mathrm{d}\eta + \eta\mathrm{d}.$$

Then, f and g induce the same map in cohomology.

Corollary 15.7. *$H^k\mathbb{R}^m = 0$ for $k>0$.*

Proof. There is a homotopy $\mathrm{id}\sim 0$ between the identity and the constant map onto the zero. Then for any closed k-form ω we have $[\omega] = [\mathrm{id}^*\omega] = [0^*\omega] = [0]$. $\square$

Remark. The case $k=1$ gives the following: a 1-form $\omega = g_i\cdot\mathrm{d}x^i$ is a differential of a function f, i.e. we have $g_i = \partial_if$, if and only if $\mathrm{d}\omega = \sum_{i<j}(\partial_ig_j - \partial_jg_i)\mathrm{d}x^i\wedge\mathrm{d}x^j = 0$, i.e. $\partial_ig_j = \partial_jg_i$. Tracing the proof, we get the function f as $f(x) = \int_0^1 \mathrm{d}f(\gamma'(t))\ \mathrm{d}t$ where γ is a path from 0 to x, e.g. the straight line. The Stokes' theorem gives independence on the choice of the path. At the same time, travelling in the direction of the axes one at a time and finishing in the direction of x^i easily gives the required $\partial_if = g_i$.

Special cases of Stokes theorem. We treat some special cases of Stokes theorem that are the most classical instances. In particular, we interpret the integrals of the forms as classical line and surfaces integrals.

Dimension 1. We start with 1-forms on $\mathbb{R}^m$ or any oriented Riemannian manifold for that matter. The Riemannian metric gives an isomorphism $\mathfrak{X}\mathbb{R}^m \xrightarrow{\cong} \Omega^1\mathbb{R}^m$, given by $X \mapsto X \llcorner g = \langle X, - \rangle$. In coordinates,

$$X^i \cdot \partial_i \mapsto X^1 \cdot \mathrm{d}x^1 + \cdots + X^m \mathrm{d}x^m.$$

(The Einstein summation notation does not apply on the right hand side; more naturally, it would be $g_{ij} X^i \cdot \mathrm{d}x^j$ where g_{ij} are the coordinates of the Riemannian metric, i.e. $g_{ij} = \delta_{ij}$, the Kronecker delta.) Thus, the differential $\mathrm{d}\colon \Omega^0\mathbb{R}^m \to \Omega^1\mathbb{R}^m$ is identified with

$$\operatorname{grad}\colon C^\infty\mathbb{R}^m \to \mathfrak{X}\mathbb{R}^m.$$

given by $\operatorname{grad} f = \partial_1 f \cdot \partial_1 + \cdots + \partial_m f \cdot \mathrm{d}x^m$. Thus, the Stokes theorem in this case gives

$$\int_C \operatorname{grad} f = f(b) - f(a)$$

where a and b are the endpoints of the curve C and the line integral on the left is defined through the corresponding 1-form, i.e. for a parametrization γ, we have

$$\begin{aligned} \int_C X = \int_C (X \llcorner g) &= \int_{\operatorname{dom}\gamma} (X \llcorner g)(\gamma'(t))\ \mathrm{d}t \\ &= \int_{\operatorname{dom}\gamma} \langle X|_{\gamma(t)}, \gamma'(t) \rangle\ \mathrm{d}t \end{aligned}$$

Codimension 1. We proceed with $(m-1)$-forms on $\mathbb{R}^m$ or any oriented Riemannian manifold for that matter. The multiplication by the volume form gives an isomorphism $C^\infty\mathbb{R}^m \xrightarrow{\cong} \Omega^m\mathbb{R}^m$. Further, we have an isomorphism

$$\mathfrak{X}\mathbb{R}^m \xrightarrow{\cong} \Omega^{m-1}\mathbb{R}^m$$

given by $X \mapsto X \llcorner \mathrm{Vol} = \mathrm{Vol}(X, -, \dots, -)$, in coordinates

$$X^i \cdot \partial_i \mapsto \sum_i (-1)^{i-1} X^i \cdot \mathrm{d}x^1 \wedge \cdots \widehat{\mathrm{d}x^i} \cdots \wedge \mathrm{d}x^m.$$

In the special case $m = 3$, this becomes

$$A \cdot \partial_x + B \cdot \partial_y + C \cdot \partial_z \mapsto A \cdot \mathrm{d}y \wedge \mathrm{d}z + B \cdot \mathrm{d}z \wedge \mathrm{d}x + C \cdot \mathrm{d}x \wedge \mathrm{d}y.$$

Thus, the differential $\mathrm{d}\colon \Omega^{m-1}\mathbb{R}^m \to \Omega^m\mathbb{R}^m$ is identified with

$$\operatorname{div}\colon \mathfrak{X}\mathbb{R}^m \xrightarrow{\mathrm{d}} C^\infty\mathbb{R}^m$$

given by $\operatorname{div}(X^i \cdot \partial_i) = \partial_i X^i$. Thus, the Stokes theorem in this case gives

$$\int_D \partial_i X^i = \int_{\partial D} X^i \cdot \partial_i$$

where the hypersurface integral is defined through the $(m-1)$-form, i.e.

$$\begin{aligned} \int_S X = \int_S (X \llcorner \mathrm{Vol}) &= \int \cdots \int_{\operatorname{dom}\sigma} (X \llcorner \mathrm{Vol})(\partial_1\sigma(t), \dots, \partial_{m-1}\sigma(t))\ \mathrm{d}t^1 \cdots \mathrm{d}t^{m-1} \\ &= \int \cdots \int_{\operatorname{dom}\gamma} \mathrm{Vol}(X|_{\sigma(t)}, \partial_1\sigma(t), \dots, \partial_{m-1}\sigma(t))\ \mathrm{d}t^1 \cdots \mathrm{d}t^{m-1} \\ &= \int \cdots \int_{\operatorname{dom}\gamma} \langle X|_{\sigma(t)}, n|_t \rangle \cdot \mathrm{Vol}(\partial_1\sigma(t), \dots, \partial_{m-1}\sigma(t))\ \mathrm{d}t^1 \cdots \mathrm{d}t^{m-1} \end{aligned}$$

where n is a unit normal vector field along S such that $(n|_t, \partial_1\sigma(t), \dots, \partial_{m-1}\sigma(t))$ is positive. When $S = \partial D$, including orientation, and the parametrization σ agrees with the orientation of S then n is the outward unit normal field.

Classical Stokes theorem. Finally, we treat $\mathrm{d}\colon \Omega^1\mathbb{R}^3 \to \Omega^2\mathbb{R}^3$ that under the above identification becomes

$$A\cdot\partial_x + B\cdot\partial_y + C\cdot\partial_z \mapsto (\partial_y C - \partial_z B)\cdot\partial_x + (\partial_z A - \partial_x C)\cdot\partial_y + (\partial_x B - \partial_y A)\cdot\partial_z$$

and the Stokes theorem becomes

$$\int_S (\partial_y C - \partial_z B)\cdot\partial_x + (\partial_z A - \partial_x C)\cdot\partial_y + (\partial_x B - \partial_y A)\cdot\partial_z = \int_{\partial S} A\cdot\partial_x + B\cdot\partial_y + C\cdot\partial_z.$$

16. Riemannian geometry

16.1. **Preliminary results.**

Lemma 16.1. *For every map*

$$F\colon \mathfrak{X}M \times \cdots \times \mathfrak{X}M \to C^\infty M$$

that is $C^\infty M$-linear in each variable there exists a unique tensor field ω of type $(0,k)$ such that $F(X_1,\ldots,X_k) = \omega(X_1,\ldots,X_k)$.

Proof. We first prove that $F_x = \mathrm{ev}_x F$ is local; we will assume $k=1$ here for simplicity. Thus, let $X \in \mathfrak{X}M$ be zero in a neighbourhood of $x \in M$. Then there exists a function λ such that λ is zero near x and $X = \lambda\cdot X$. Then $F_x(X) = F_x(\lambda\cdot X) = \lambda(x)\cdot F_x(X) = 0$. This allows one to define F_x on germs of vector fields and, consequently, a local version $F_x\colon \mathfrak{X}_x M \to \mathbb{R}$ as in the case of derivations. One may then assemble these into $F_U\colon \mathfrak{X}U \to C^\infty U$, $F(X)(x) = F_x(\mathrm{germ}_x X)$. Again, this is smooth, because, near any point of U, the vector field X is equal to a vector field $\bar{X}$ that extends to M and for which $F_x(\mathrm{germ}_x \bar{X}) = F(\bar{X})(x)$ is thus smooth in x.

Now, for general k, the map F is $C^\infty M$-linear in each variable and thus local in each variable. Since locally $X_j = X^{i_j}\partial_{i_j}$, we obtain

$$F(X_1,\ldots,X_k)(x) = F_x(X_1,\ldots,X_k) = F_x(\partial_{i_1},\ldots,\partial_{i_k})X^{i_1}(x)\cdots X^{i_k}(x),$$

i.e. we have to take $\omega = F(\partial_{i_1},\ldots,\partial_{i_k})\,\mathrm{d}x^{i_1}\otimes\cdots\otimes\mathrm{d}x^{i_k}$. □

A similar result holds for maps $F\colon \mathfrak{X}M\times\cdots\times\mathfrak{X}M \to \mathfrak{X}M$ – such maps are given by tensor fields of type $(1,k)$; the proof is the same.

A slight generalization of the first part of the proof of the previous lemma is the following (for simplicity, we state only unary version):

Lemma 16.2. *Suppose that $F\colon \mathfrak{X}M \to C^\infty M$ is $\mathbb{R}$-linear and satisfies $F_x(f\cdot X) = 0$ for each f that is zero in a neighbourhood of x. Then there exists a unique map $F\colon \mathfrak{X}U \to C^\infty U$ that passes to the same map $F_x\colon \mathfrak{X}_x M \to C^\infty_x M$.*

16.2. **Covariant derivative for submanifolds of Euclidean spaces.**

We start with the following situation. Let $M \subseteq E$ be a submanifold. Then we have the following concepts available in M: parallel transport and covariant derivative. We start with the parallel transport which we find more intuitive. Let $\gamma\colon \mathbb{R}\to M$ be a path and $Z\colon \mathbb{R}\to TM$ be a vector field along γ, i.e. we assume $Z(t) \in T_{\gamma(t)}M$. We say that Z *transports parallelly* along γ in M if $D_t Z|_t$ is perpendicular to $T_{\gamma(t)}M$. Denoting by P_x the orthogonal projection $T_xE \to T_xM$, this means $P_{\gamma(t)}(D_t Z|_t) = 0$ or simply $P(D_t Z) = 0$.

Denoting $\nabla_X Y = P(D_X Y)$, the condition of the parallel transport is thus $\nabla_{\gamma'} X = 0$. Since we have $D_X Y - D_Y X = [X,Y]$ and $[X,Y]$ is tangent to M if both X and Y are (so that $[X,Y]$ is preserved by P), we obtain

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

We say that the covariant derivative ∇ is symmetric.

The second property follows from the observation

$$D_X\langle Y,Z\rangle = \underbrace{(D_X g)(Y,Z)}_{0} + \langle D_X Y, Z\rangle + \langle Y, D_X Z\rangle$$

which is easily implied by the shape of the scalar product $\langle Y,Z\rangle = g_{ij}Y^iZ^j$ and the constantness of g, i.e. $D_X g = 0$. The left hand side is the usual derivative of a function along a vector field, i.e. $X\langle Y, Z\rangle$. For notational convenience, we will denote it also by $\nabla_X\langle Y, Z\rangle$. Since Z is tangent to M, its product with $D_X Y$ is the same as with $\nabla_X Y = P(D_X Y)$ and, thus, the above can be rewritten as

$$\nabla_X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle$$

(it is some form of Leibniz rule). We say that ∇ is metric.

16.3. **Covariant derivative of vector fields along paths.** We observe that $D_X Y|_x$ depends only on the value $X|_x$ and on the values of Y in the direction of $X|_x$. More concretely, if $X|_x = \gamma'(t_0)$ then only values of Y along γ are important, i.e. only the composition $Z = Y \circ \gamma$, a vector field along γ:

$$D_{\gamma'(t_0)}Y = D_{\gamma'(t_0)}Y^j \cdot \partial_j = \partial_t|_{t_0}(Y^j \circ \gamma) \cdot \partial_j = \partial_t|_{t_0} Z^j \cdot \partial_j = D_t Z|_{t_0}$$

or more concisely $D_{\gamma'}Y = D_t Z$. This formula defines, for an arbitrary vector field Z along γ, another vector field $D_t Z$ along γ. In particular, the tangent vectors $\gamma' = \partial_t \gamma^i \cdot \partial_i$ form a vector field along γ and we get

$$D_t \gamma' = \gamma'' = \partial_t^2 \gamma^i \cdot \partial_i.$$

16.4. **Riemannian manifolds and linear connections.**

Definition 16.3. A *Riemannian metric* on a smooth manifold M is a choice of a scalar product on each $T_x M$ that depends smoothly on $x \in M$. In detail, it is a tensor field of type $(0,2)$, i.e. a smooth map $g\colon M \to (T^*M)^{\otimes 2}$, that is symmetric and positive definite at each point (i.e. each $g_x \in (T^*M)^{\otimes 2}$ should be symmetric and positive definite).

A *Riemannian manifold* is a manifold equipped with a Riemannian metric.

Example 16.4. The Euclidean space with the constant field g. Any submanifold $M \subseteq E$ of a Euclidean space E with the restriction of the scalar product on E to M (formally a pullback along the inclusion).

We consider a mapping

$$\nabla\colon \mathfrak{X}M \times \mathfrak{X}M \to \mathfrak{X}M$$

and denote its values $\nabla_X Y$. To make the following formulas more symmetric, we also denote $\nabla_X f = Xf$, i.e. the usual directional derivative.

Definition 16.5. A *linear connection* is an operator ∇ as above satisfying

$$\begin{aligned} \nabla_X(Y_1 + Y_2) &= \nabla_X Y_1 + \nabla_X Y_2, \\ \nabla_X(f \cdot Y) &= \nabla_X f \cdot Y + f \cdot \nabla_X Y, \\ \nabla_{X_1+X_2} Y &= \nabla_{X_1} Y + \nabla_{X_2} Y, \\ \nabla_{f\cdot X} Y &= f \cdot \nabla_X Y. \end{aligned}$$

Example 16.6. The local covariant derivative D on an open subset of a vector space (all properties are trivial). The covariant derivative $\nabla_X Y = P(D_X Y)$ on a submanifold of a Euclidean space – the only non-trivial axiom is the second one (apply the projection P to the equality $D_X(f \cdot Y) = D_X f \cdot Y + f \cdot D_X Y$ and observe that the first term on the right belongs to TM so that it is preserved by P).

Definition 16.7. A connection ∇ is *symmetric* if $\nabla_X Y - \nabla_Y X = [X, Y]$.

Remark. The difference $T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]$ is called the *torsion* of ∇. Thus, a connection is symmetric if and only if it is *torsion-free.*

Definition 16.8. A connection ∇ on a Riemannian manifold M is *metric* if

$$\nabla_X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle.$$

The conditions of a connection imply that it is local (i.e. $\nabla_X Y|_x$ depends only on the germs of X and Y at x). Let now $D_X Y$ again denote the usual local covariant derivative in a given coordinate system on $U \subseteq M$, *in contrast* to the previous setup of a submanifold $M \subseteq E$, where it denoted the covariant derivative in the ambient euclidean space E! Then $\nabla_X Y - D_X Y$ is $C^\infty U$-linear in both X and Y and, thus, given by a tensor field Γ of type $(1,2)$ on U, i.e.

$$\nabla_X Y = D_X Y + \Gamma(X,Y).$$

We stress that Γ depends significantly on the coordinate system.

The symmetry of ∇ translates to $\Gamma(X,Y) = \Gamma(Y,X)$, i.e. the symmetry of Γ, and the metricity of ∇ translates to

$$\begin{aligned}
\nabla_X \langle Y, Z\rangle &= \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle \\
D_X \langle Y, Z\rangle &= \langle D_X Y + \Gamma(X,Y), Z\rangle + \langle Y, D_X Z + \Gamma(X,Z)\rangle \\
(D_X g)(Y,Z) + \langle D_X Y, Z\rangle + \langle Y, D_X Z\rangle &= \langle \Gamma(X,Y), Z\rangle + \langle Y, \Gamma(X,Z)\rangle + \langle D_X Y, Z\rangle + \langle Y, D_X Z\rangle \\
(D_X g)(Y,Z) &= \langle \Gamma(X,Y), Z\rangle + \langle \Gamma(X,Z), Y\rangle.
\end{aligned}$$

(The left hand side can be also denoted $Dg(X,Y,Z)$.) According to the following lemma, this determines $\langle \Gamma(X,Y), Z\rangle$ uniquely and, thus, also $\Gamma(X,Y)$, so that we obtain the following theorem.

Theorem 16.9. *There exists a unique symmetric and metric connection on a given Riemannian manifold – it is called the* Levi-Civita connection*.* □

Lemma 16.10. *The map* $\mathrm{sym}_{23}\colon (S^2V \otimes V)^* \to (V \otimes S^2V)^*$*, given by* $\mathrm{sym}_{23}\,\omega(X,Y,Z) = \omega(X,Y,Z) + \omega(X,Z,Y)$ *is an isomorphism.*

Proof. The spaces have the same dimensions; thus, it is enough to show that the kernel is zero. But any $\omega \in \ker \mathrm{sym}_{23}$ is symmetric in the first two and antisymmetric in the last two variables, hence zero. □

Remark. In fact, it is not difficult to show that the inverse is given by

$$((\mathrm{sym}_{23})^{-1}\theta)(X,Y,Z) = \tfrac{1}{2}(\theta(X,Y,Z) + \theta(Y,X,Z) - \theta(Z,X,Y)).$$

However, we will not make use of this formula.

16.5. **Parallel transport, geodesics.**

Definition 16.11. We say that a vector field Z along a path γ *transports parallelly* if $\nabla_t Z = 0$.

The equation for the parallel transport is

$$0 = \nabla_t Z = D_t Z + \Gamma(\gamma', Z),$$

i.e. $D_t Z = -\Gamma(\gamma', Z)$. This is an ordinary differential equation and, locally, a unique solution exists through each choice of $Z|_0$. However, since the solution exists globally for the zero vector, it must exist for any small vector and then for any vector since the parallel transport is clearly linear – any linear combination (with constant coefficients) of parallel vector fields is also parallel.

Another observation is that if both X and Y transport parallelly along γ then

$$\nabla_t \langle X, Y\rangle = \langle \nabla_t X, Y\rangle + \langle X, \nabla_t Y\rangle = 0$$

and the scalar product $\langle X, Y\rangle$ is constant along γ – we say that the parallel transport preserves the scalar product (in fact, this is equivalent to the metricity of ∇).

We denote by Pt^γ_t the map $T_{\gamma(t_0)}M \to T_{\gamma(t_0+t)}M$ obtained by transporting vectors parallelly along γ. We have thus proved that each Pt^γ_t is orthogonal.

Definition 16.12. A *geodesic path* is a path γ such that γ' transports parallelly along γ.

In coordinate, this becomes

$$\gamma'' = D_t \gamma' = -\Gamma(\gamma', \gamma')$$

and is a differential equation of second order. Again, locally, a unique solution exists with any given $A = \gamma'(0) \in T_{\gamma(0)}M$. We will temporarily denote it γ_A. Then it is pretty much clear that

$$\gamma_{sA}(t) = \gamma_A(st).$$

Thus, denoting $\exp A = \gamma_A(1)$, we obtain $\gamma_A(t) = \gamma_{tA}(1) = \exp tA$. The map $\exp\colon TM \dashrightarrow M$ is not defined globally; however, it is defined in a neighbourhood of the zero section of TM, since $\exp 0_x = x$. Each $\exp_x\colon T_xM \dashrightarrow M$ is a local diffeomorphism at 0_x (since $\exp_{x*0} = \mathrm{id}$ is the identity: $A \in T_0(T_xM)$ is tangent to the linear path tA and, thus, $\exp_{x*0} A$ is the tangent vector to $\exp_x(tA) = \gamma_A(t)$, that is A by definition).

Describing covariant derivative using parallel transport. We will now show how the covariant derivative can be reconstructed from the parallel transport. Let e_i be an orthonormal frame at x and transport it parallelly along a path γ through x. Then we get vector fields E_i along γ and they will still be orthonormal since parallel transport preserves scalar product. Let X be a vector field along γ and express it in this orthonormal frame as $X = f^i E_i$. Then

$$\nabla_t X = \nabla_t \left(f^i E_i\right) = \nabla_t f^i \cdot E_i + f^i \cdot \underbrace{\nabla_t E_i}_{0} = \nabla_t f^i \cdot E_i.$$

In other words, expressing the vector field in a parallel orthonormal frame makes it into a function $f\colon \mathbb{R} \to \mathbb{R}^n$ and then the covariant derivative is simply the usual derivative $\nabla_t f = D_t f$.

It will be important in the next section that $f^i(t)e_i = \mathrm{Pt}^{\gamma}_{-t} X(t)$ so that $\nabla_t|_0 X = D_t|_0 \mathrm{Pt}^{\gamma}_{-t} X(t)$. In plain words, transporting the vector field X along γ to $\gamma(0)$ produces a path of vectors in $T_{\gamma(0)}M$ and $\nabla_t|_0 X$ is then the usual derivative at zero of this function $\mathbb{R} \to T_xM$ with values in a vector space.

16.6. **Curvature.** The curvature is defined to be $R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z$. It is a tensor of type $(1,3)$ that is clearly antisymmetric in the first two variables. We have seen that the curvature of the Euclidean covariant derivative is zero. In fact, this characterizes the Euclidean connection, as the following theorem says.

Remark. We prove $C^\infty M$-linearity in Z, the same properties for X and the symmetric Y are simpler:

$$\begin{aligned}\nabla_X \nabla_Y (f \cdot Z) &= \nabla_X (Yf \cdot Z + f \cdot \nabla_Y Z)\\ &= XYf \cdot Z + Yf \cdot \nabla_X Z + Xf \cdot \nabla_Y Z + f \cdot \nabla_X \nabla_Y Z\end{aligned}$$

so that

$$\nabla_X \nabla_Y (f \cdot Z) - \nabla_Y \nabla_X (f \cdot Z) = [X,Y]f \cdot Z + f \cdot (\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z)$$

(the middle terms cancel out) and upon subtracting

$$\nabla_{[X,Y]}(f \cdot Z) = [X,Y]f \cdot Z + f \cdot \nabla_{[X,Y]} Z,$$

we obtain the required linearity $R(X,Y)(f \cdot Z) = f \cdot R(X,Y)Z$.

Remark. Expanding

$$\nabla_X \nabla_Y Z = D_X(D_Y Z + \Gamma(Y,Z)) + \Gamma(X, D_Y Z + \Gamma(Y,Z))$$

and the remaining two terms, we get

$$\begin{aligned}\nabla_X \nabla_Y Z &= D_X D_Y Z + (D_X\Gamma)(Y,Z) + \Gamma(D_X Y, Z) + \Gamma(Y, D_X Z) + \Gamma(X, D_Y Z) + \Gamma(X, \Gamma(Y,Z))\\ \nabla_Y \nabla_X Z &= D_Y D_X Z + (D_Y\Gamma)(X,Z) + \Gamma(D_Y X, Z) + \Gamma(X, D_Y Z) + \Gamma(Y, D_X Z) + \Gamma(Y, \Gamma(X,Z))\\ \nabla_{[X,Y]} Z &= D_X D_Y Z - D_Y D_X Z + \Gamma(D_X Y, Z) - \Gamma(D_Y X, Z)\end{aligned}$$

(since $D_{[X,Y]} Z = D_X D_Y Z - D_Y D_X Z$) and we finally obtain

$$R(X,Y)Z = D\Gamma(X,Y,Z) - D\Gamma(Y,X,Z) + \Gamma(X,\Gamma(Y,Z)) - \Gamma(Y,\Gamma(X,Z))$$

that is clearly a tensor field of type $(1,3)$.

Before going into the proof, we give a geometric meaning to the curvature. Let X, Y be vector fields that commute, i.e. such that $[X,Y] = 0$. Then $\nabla_X \nabla_Y Z(x)$ is obtained as the mixed partial derivative $\frac{\partial^2}{\partial s \partial t} A(0,0)$ of the vector valued function $A(s,t) \in T_xM$ given by transporting parallelly $Z(\mathrm{Fl}^Y_t(\mathrm{Fl}^X_s(x)))$ along the flow line of Y back to $\mathrm{Fl}^X_s(x)$ and then along the flow line of X back to x. A similar formula holds for the second term. We may however define Z by first transporting $Z(0,0) \in T_xM$ along the flow line of Y and then along the flow lines of X so that the second term actually becomes zero. Thus, we finally obtain

$$R(X,Y)Z = \frac{\partial^2}{\partial s \partial t}\Big|_{(s,t)=(0,0)} \mathrm{Pt}^{\mathrm{Fl}^X}_{-s}\, \mathrm{Pt}^{\mathrm{Fl}^Y}_{-t}\, \mathrm{Pt}^{\mathrm{Fl}^X}_{s}\, \mathrm{Pt}^{\mathrm{Fl}^Y}_{t}\, Z$$

(it would look slightly nicer if s, t were changed to their opposites – then this becomes a commutator).

Continuing the notation of the above proof, we will show that for $R(X,Y)Z = 0$ and $[X,Y] = 0$, the parallel transports along the flow lines of X and Y commute: $0 = R(X,Y)Z = \nabla_X \nabla_Y Z$, so that $\nabla_Y Z$ transports parallelly along the flow lines of X. Since $\nabla_Y Z = 0$ for $s = 0$, it must be zero everywhere, i.e. Z also transports parallelly along the flow lines of Y. In particular, we obtain $Z(s,t)$ also by transporting $Z(0,0)$ first parallelly along the flow line of X to get $Z(0,t)$ and then parallelly along the flow line of Y (this is what we have just proved).

Theorem 16.13. *The following conditions are equivalent.*

1. *The curvature is zero.*
2. *The parallel transport does not locally depend on the path.*
3. *There is an atlas in which all the Γ are zero.*
4. *There is an atlas consisting of isometries.*

Proof. We will prove (1)⇒(2)⇒(3)⇒(4)⇒(1).

(1)⇒(2): We use the fact that the parallel transports along vector fields X, Y commute when $[X,Y] = 0$. Start with a vector $Z_0 \in T_x M$ and transport it parallelly along the local vector fields $\partial_1, \ldots, \partial_n$ to obtain a vector field Z with $Z(x) = Z_0$. Since Z was obtained by parallel transport along ∂_i (any one could have been used the last), we have $\nabla_{\partial_i} Z = 0$. This holds for any i and, thus, $\partial_X Z = 0$ for any X. In particular, Z transports parallelly along any path, implying that the parallel transport of $Z(x)$ along a path from x to y is always $Z(y)$.

(2)⇒(3): Suppose that the parallel transport does not locally depend on the path. Start with a basis (e_i) of $T_x M$ and transport it locally to a neighbourhood to obtain vector fields E_i. Then $[E_i, E_j] = \nabla_{E_i} E_j - \nabla_{E_j} E_i = 0$ and there exists a chart with $E_i = \partial_i$. In particular, $\Gamma(E_i, E_j) = \nabla_{E_i} E_j - D_{E_i} E_j = 0$ and $\Gamma = 0$.

(3)⇒(4): Clearly, to obtain a local isometry, it is enough to have $D_X g = 0$ for all X (then g is constant and we may modify the chart by a linear isomorphism). But

$$0 = (\nabla_X g)(Y,Z) = (D_X g)(Y,Z) - \langle \Gamma(X,Y), Z \rangle - \langle Y, \Gamma(X,Z) \rangle = (D_X g)(Y,Z).$$

(4)⇒(1) is clear since we have $\nabla_X Y = D_X Y$ in a Euclidean space and the curvature is zero. □

17. Spaces of constant curvature

17.1. **Sectional curvature.** First we observe that $\langle R(X,Y)Z, W \rangle$ is also anti-symmetric in the variables Z, W. This follows from

$$\begin{aligned} 0 &= (\nabla^2_{\mathrm{alt}} \langle Z, W \rangle)(X,Y) = \nabla_X \nabla_Y \langle Z, W \rangle - \nabla_Y \nabla_X \langle Z, W \rangle - \nabla_{\nabla_X Y - \nabla_Y X} \langle Z, W \rangle \\ &= \langle R(X,Y)Z, W \rangle + \langle Z, R(X,Y)W \rangle \end{aligned}$$

(the terms where each Z, W receives one of the ∇_X, ∇_Y cancel out).

Let p be a 2-dimensional vector subspace of $T_x M$ with orthonormal basis (e_1, e_2). We define the *sectional curvature* $K(p) = -\langle R(e_1, e_2)e_1, e_2 \rangle$. Denoting $R(X,Y,Z,W) = \langle R(X,Y)Z, W \rangle$, we see that this depends only on $X \wedge Y$ and $Z \wedge W$. Thus, when replacing (X,Y) and (Z,W) by linear combinations thereof, the whole expression gets multiplied by the product of the determinants of the transformation matrices. In the case of $K(p)$, this means the square of an orthogonal matrix and so the value does not change.

17.2. **Sphere.**

This is better using Theorema egregium.

We compute the sectional curvature of a sphere $S^n \subseteq \mathbb{R}^{n+1}$. Thus, let $e_1, e_2 \in T_{x_0} S^n$ be two orthonormal vectors. We extend them to vector fields on S^n in the following way: think of the e_i as a constant vector field on $\mathbb{R}^{n+1}$ and project it orthogonally to obtain a vector field E_i on TS^n; at a point x, this equals

$$E_i(x) = e_i - \langle x, e_i \rangle x$$

In fact, this formula prescribes a vector field on $\mathbb{R}^{n+1}$ – this is useful since we want to apply the covariant derivative of $\mathbb{R}^{n+1}$:

$$D_A E_i = -\langle A, e_i\rangle x - \langle x, e_i\rangle A$$

Projecting to TS^n, the first term becomes zero and the second term remains unchanged (since A is now assumed tangent to S^n), i.e.

$$\nabla_A E_i = -\langle x, e_i\rangle A$$

This leads to

$$[E_i, E_j] = \nabla_{E_i} E_j - \nabla_{E_j} E_i = \langle x, e_i\rangle E_j - \langle x, e_j\rangle E_i$$

and finally

$$R(E_i, E_j)E_i = \nabla_{E_i} \underbrace{\nabla_{E_j} E_i}_{-\langle x,e_i\rangle E_j} - \nabla_{E_j} \underbrace{\nabla_{E_i} E_i}_{-\langle x,e_i\rangle E_i} - \nabla_{[E_i,E_j]} E_i$$

Now $A\langle x, e_i\rangle = \langle A, e_i\rangle$, since the function is linear in x. Thus,

$$\begin{aligned} R(E_i, E_j)E_i &= -\langle x, e_i\rangle \cdot \underbrace{\nabla_{E_i} E_j}_{-\langle x,e_j\rangle E_i} - \langle E_i, e_i\rangle E_j \\ &\quad + \langle x, e_i\rangle \cdot \underbrace{\nabla_{E_j} E_i}_{-\langle x,e_i\rangle E_j} + \langle E_j, e_i\rangle E_i \\ &\quad + \langle x, e_i\rangle(\langle x, e_i\rangle E_j - \langle x, e_j\rangle E_i) \\ &= \langle E_j, e_i\rangle E_i - \langle E_i, e_i\rangle E_j \end{aligned}$$

so that $-\langle R(E_i, E_j)E_i, E_j\rangle(x_0) = -\langle e_j, e_i\rangle\langle e_i, e_j\rangle + \langle e_i, e_i\rangle\langle e_j, e_j\rangle = 1 \cdot \mathrm{Vol}(e_i, e_j)^2$.

17.3. **Hyperbolic space.** We compute the sectional curvature of a hyperbolic space $H^n \subseteq \mathbb{R}^{n+1}$ equipped with a metric $g = -\mathrm{d}x^0 \otimes \mathrm{d}x^0 + \mathrm{d}x^1 \otimes \mathrm{d}x^1 + \cdots + \mathrm{d}x^n \otimes \mathrm{d}x^n$, where

$$H^n = \{x \in \mathbb{R}^{n+1} \mid g(x, x) = -1\}.$$

Since each $x \in H^n$ generates a 1-dimensional subspace where the metric is negative definite, it is easy to see that on its orthogonal complement – and this is again $T_x H^n$ – the metric must be positive definite (the inertia theorem). Thus, let $e_1, e_2 \in T_{x_0} H^n$ be two orthonormal vectors. We extend them to vector fields on H^n, this times the formula is

$$E_i(x) = e_i + \langle x, e_i\rangle x$$

(because $\langle x, x\rangle = -1$). Again, this prescribes a vector field on $\mathbb{R}^{n+1}$ and:

$$D_A E_i = \langle A, e_i\rangle x + \langle x, e_i\rangle A$$

Projecting to TS^n, we get

$$\nabla_A E_i = \langle x, e_i\rangle A$$

This leads to the same formula for $R(E_i, E_j)E_j$ as above, only with different signs. Since the surviving terms in $\langle R(E_i, E_j)E_j, E_i\rangle$ contain exactly one scalar product from $R(E_i, E_j)E_j$, we get

$$\langle R(E_i, E_j)E_j, E_i\rangle(x_0) = \langle e_i, e_j\rangle\langle e_j, e_i\rangle - \langle e_i, e_i\rangle\langle e_j, e_j\rangle = -1 \cdot \mathrm{Vol}(e_i, e_j)^2.$$

17.4. **Theorema Egregium.** We denote $D_X Y = \nabla_X Y + h(X, Y) \cdot n$ where n is a choice of a unit normal vector (such exists locally; globally, n, and thus also h, is well defined up to a sign). As a difference of two covariant derivatives, h is a tensor field of type $(0, 2)$. Moreover, it is symmetric:

$$(h(X, Y) - h(Y, X)) \cdot n = (D_X Y - D_Y X) - (\nabla_X Y - \nabla_Y X) = [X, Y] - [X, Y] = 0$$

Theorem 17.1 (Gauss formula)**.** *For a hypersurface $M \subseteq \mathbb{R}^{m+1}$ it holds*

$$-\langle R(X, Y)Z, U\rangle = h(X, Z)h(Y, U) - h(Y, Z)h(X, U)$$

Proof. By the metricity of the connection

$$\begin{aligned}\langle \nabla_X \nabla_Y Z, U\rangle &= X\langle \nabla_Y Z, U\rangle - \langle \nabla_Y Z, \nabla_X U\rangle \\ &= X\langle D_Y Z, U\rangle - \langle D_Y Z, D_X U\rangle + h(Y,Z)h(X,U) \\ &= \langle D_X D_Y Z, U\rangle + h(Y,Z)h(X,U)\end{aligned}$$

and similarly $\langle \nabla_{[X,Y]} Z, U\rangle = \langle D_{[X,Y]} Z, U\rangle$. Therefore,

$$\langle R(X,Y)Z, U\rangle = \langle R^{\mathrm{euc}}(X,Y)Z, U\rangle + h(Y,Z)h(X,U) - h(X,Z)h(Y,U)$$

with the first term zero since the curvature of $\mathbb{R}^{m+1}$ vanishes. □

Thus, we get

$$-\langle R(X,Y)X, Y\rangle = \det \begin{pmatrix} h(X,X) & h(X,Y) \\ h(Y,X) & h(Y,Y)\end{pmatrix} = K(p)\cdot\det\begin{pmatrix}\langle X,X\rangle & \langle X,Y\rangle \\ \langle Y,X\rangle & \langle Y,Y\rangle\end{pmatrix} = K(p)\cdot \mathrm{Vol}(X,Y)^2,$$

where $K(p)$ depends only on the plane p spanned by X and Y (both sides depend quadratically on $X \wedge Y$; since $X' \wedge Y' = X \wedge Y \cdot \det T$, where T is the transformation matrix $(X', Y') = (X, Y) \cdot T$, the quotient $K(p)$ is independent of the basis (X, Y) of p). It is called the *sectional curvature* in the direction of p. Clearly, it is the product of the eigenvalues of h on p. The corresponding eigenspaces are called principal directions (at least in the case of surfaces).