Riemannian Geometry

Ivan Izmestiev

University of Fribourg, 2016/17

Chapter 1

Smooth manifolds

1.0 Smooth maps between Euclidean domains

Let $F\colon \mathbb{R}^n \to \mathbb{R}^m$ be a map with continuous partial derivatives at the point $x \in \mathbb{R}^n$. Then the *differential of* F *at* x is the linear map

$$DF_x\colon \mathbb{R}^n \to \mathbb{R}^m$$

given by the matrix of the partial derivatives of F. That is, for every $v = (v^1, \ldots, v^n)^\top \in \mathbb{R}^n$ we have

$$DF_x(v) = \left(\sum_{i=1}^{n} \frac{\partial F^i}{\partial x^j} v^j \right)^\top \in \mathbb{R}^m$$

The chain rule for the composition of the maps $\mathbb{R}^n \xrightarrow{F} \mathbb{R}^m \xrightarrow{G} \mathbb{R}^k$ can be written as

$$D(G \circ F)_x = DG_{F(x)} \circ DF_x$$

Check that this formula says the same as the coordinate form of the chain rule

$$\frac{\partial (G \circ F)^i}{\partial x^j} =$$

1.1 Smooth manifolds and smooth maps

Definition 1.1. A *topological* n*-dimensional manifold* M is a Hausdorff, second countable topological space locally homeomorphic to $\mathbb{R}^n$. That is, every point $p \in M$ has an open neighborhood $U \subset M$ for which there is a map $\varphi\colon U \to \mathbb{R}^n$ sending U homeomorphically onto an open subset of $\mathbb{R}^n$.

- A pair (U, φ) is called a *chart*.
- A *chart* is also called a *coordinate map*. To every point $p \in U$ we can associate the coordinates of its image $\varphi(p) \in \mathbb{R}^n$.

- A collection of charts $(U_\alpha, \varphi_\alpha)$ such that $\cup_\alpha(U_\alpha) = M$ is called an *atlas* on M.

- If (U, φ) and (V, ψ) are two charts such that $U \cap V \neq \varnothing$, then the homeomorphism

$$\psi \circ \varphi^{-1} \colon \varphi(U \cap V) \to \psi(U \cap V)$$

  is called a *transition map* or *coordinate change.*

Definition 1.2. A *smooth n-dimensional manifold* is a topological n-dimensional manifold M together with an atlas $(U_\alpha, \varphi_\alpha)$ such that all transition maps $\varphi_\beta \circ \varphi_\alpha^{-1}$ are infinitely differentiable.

- These are C^∞-manifolds. If we require the transition maps to be k times continuously differentiable, then we have a so-called C^k-manifold. We will deal with C^∞-manifolds only, and the word "smooth" will always mean C^∞ for us.

- A chart (V, ψ) is called *admissible* for a smooth manifold $(M, (U_\alpha, \varphi_\alpha))$, if all transition maps $\psi \circ \varphi_\alpha^{-1}$ are smooth. An admissible chart can be added to the atlas, it does not change the "smooth structure". Often one defines a smooth manifold by means of a maximal atlas, which is obtained by adding all admissible charts to a given atlas.

Definition 1.3. A map $F \colon M \to N$ between two smooth manifolds is called *smooth* at the point $p \in M$ if for some charts (U, φ) around p and (V, ψ) around $F(p)$ the composition $\psi \circ F \circ \varphi^{-1}$ is smooth at p.

Exercise. Show that if for some φ and ψ the map $\psi \circ F \circ \varphi^{-1}$ is smooth at p, then for any other charts around p and $F(p)$ the map $\psi_\alpha \circ F \circ \varphi_\alpha^{-1}$ is also smooth at p.

- If F is smooth at every point of M, then we simply say "M is smooth".

- A composition of smooth maps is smooth. Why?

- A *diffeomorphism* is a smooth bijective map with a smooth inverse.

Important special cases of smooth maps:

- Smooth functions $f \colon M \to \mathbb{R}$. The space of all smooth functions on M is denoted by $C^\infty(M)$.

- Smooth paths $\gamma \colon \mathbb{R} \to M$, or $\gamma \colon I \to M$, where $I \subset \mathbb{R}$ is an open interval.

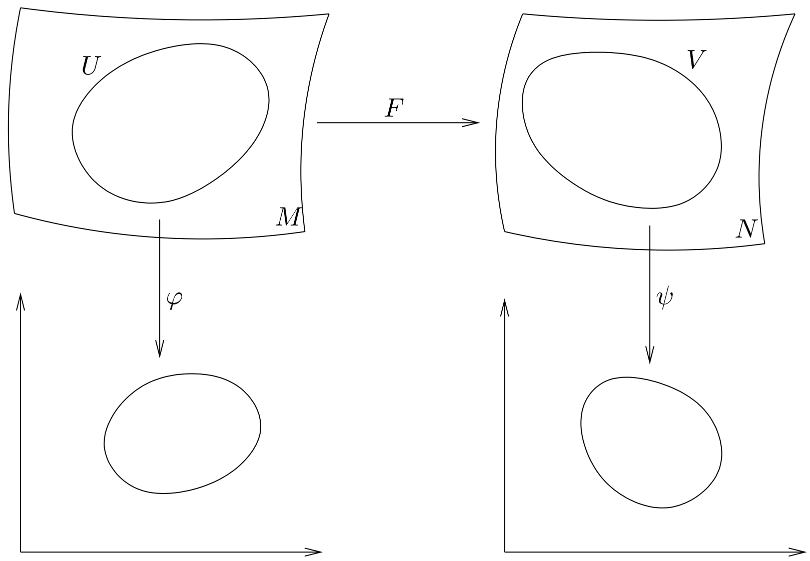

Figure 1.1: To the definition of a smooth map between manifolds.

1.2 Tangent vectors

Definition 1.4 (a)**.** A *tangent vector* to a smooth manifold M at a point p is an equivalence class of pairs

$$\{(\varphi, v) \mid \varphi\colon U \to \mathbb{R}^n \text{ a chart, } p \in U, v \in \mathbb{R}^n\}$$

under the equivalence relation

$$(\varphi, v) \sim (\psi, w) \Leftrightarrow w = D(\psi \circ \varphi^{-1})_{\varphi(p)}(v)$$

The set of all tangent vectors at p is denoted by T_pM.

- This is indeed an equivalence relation. What should be used to prove its transitivity?

- For a given φ, every equivalence class has a unique representative of the form (φ, v). This identifies T_pM with $\mathbb{R}^n$ and puts a linear structure on T_pM. This map $T_pM \to \mathbb{R}^n$ depends on φ, but the linear structure on T_pM is independent of the chart. Why?

When we consider coordinate changes, we always "go up" with φ^{-1} and then "go down" with ψ. One can simplify things by choosing a preferred chart (U, φ) and identifying U with $\varphi(U) \subset \mathbb{R}^n$. Then a tangent vector at $p \in \mathbb{R}^n$ is an n-tuple $(v^1, \ldots, v^n)^\top$. Under a coordinate change $F\colon \mathbb{R}^n \to \mathbb{R}^n$

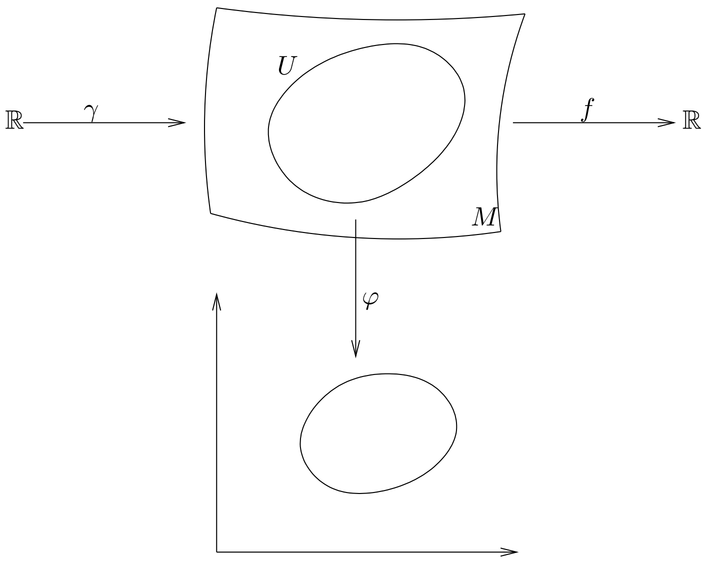

Figure 1.2: Smooth paths and smooth functions on a manifold.

vector components are transformed according to

$$w^i = \sum_{j=1}^{n} \frac{\partial F^i}{\partial x^j} v^j$$

(the local diffeomorphism F is the same as $\psi \circ \varphi^{-1}$).

Definition 1.4 (b)**.** A *tangent vector* to a smooth manifold M at a point p is an equivalence class of curves

$$\gamma \colon \mathbb{R} \to M \text{ such that } \gamma(0) = p$$

under the equivalence relation

$$\gamma_1 \sim \gamma_2 \Leftrightarrow \left.\frac{d}{dt}\right|_{t=0} (\varphi \circ \gamma_1) = \left.\frac{d}{dt}\right|_{t=0} (\varphi \circ \gamma_2)$$

for some chart φ.

- If $\left.\frac{d}{dt}\right|_{t=0} (\varphi \circ \gamma_1) = \left.\frac{d}{dt}\right|_{t=0} (\varphi \circ \gamma_2)$ for some φ, then

$$\begin{aligned}\left.\frac{d}{dt}\right|_{t=0} (\psi \circ \gamma_1) &= D(\psi \circ \varphi^{-1})_{\varphi(p)} \left(\left.\frac{d}{dt}\right|_{t=0} (\varphi \circ \gamma_1) \right) \\ &= D(\psi \circ \varphi^{-1})_{\varphi(p)} \left(\left.\frac{d}{dt}\right|_{t=0} (\varphi \circ \gamma_2) \right) = \left.\frac{d}{dt}\right|_{t=0} (\psi \circ \gamma_2)\end{aligned}$$

  for any other chart ψ.

- For a given φ, the map

$$(\varphi, v) \mapsto \left\{ \gamma \mid \frac{d}{dt}\bigg|_{t=0} (\varphi \circ \gamma) = v \right\} \tag{1.1}$$

defines a bijection between the tangent vectors in the sense of Definitions 1.4(a) and 1.4(b).

Exercise. Show that the bijection (1.1) is independent of the local chart: a vector $(\psi, w) \sim (\varphi, v)$ is sent to the same equivalence class of paths.

Definition 1.4 (c)**.** A map $\ell \colon C^\infty(M) \to \mathbb{R}$ is called a *derivation at* p, if

1. ℓ is $\mathbb{R}$-linear:

$$\ell(f+g) = \ell(f) + \ell(g), \quad \ell(\lambda f) = \lambda \ell(f) \text{ for all } \lambda \in \mathbb{R}$$

2. ℓ satisfies the Leibniz rule at p:

$$\ell(fg) = f(p)\ell(g) + \ell(f)g(p)$$

The set of all derivations at p has a structure of a vector space: to add two derivations we add their values, to multiply a derivation with a scalar, we multiply its values with this scalar.

Lemma 1.5. *The vector space of derivations at* p *is canonically isomorphic to* T_pM.

Proof. First define an isomorphism in terms of a chart, then prove that it is independent of the choice of a chart.

For a given φ define

$$\ell_{(\varphi,v)}(f) = D(f \circ \varphi^{-1})_{\varphi(p)}(v) = \sum_{i=1}^n v^i \frac{\partial(f \circ \varphi^{-1})}{\partial x^i}$$

This is clearly a linear map. It is independent of the choice of a chart since $(\psi, w) \sim (\varphi, v)$ implies

$$\begin{aligned} \ell_{(\psi,w)}(f) = D(f \circ \psi^{-1})_{\psi(p)}(w) &= D(f \circ \psi^{-1})_{\psi(p)} \left(D(\psi \circ \varphi^{-1})_{\varphi(p)}(v) \right) \\ &= D(f \circ \varphi^{-1})_{\varphi(p)}(v) = \ell_{(\varphi,v)}(f) \end{aligned}$$

Let us show that the map $(\varphi, v) \mapsto \ell_{(\varphi,v)}$ is injective. Assume that $\ell_{(\varphi,v)}(f) = 0$ for all functions f. Take f such that $f(\varphi^{-1}(x)) = x^i$, the i-th coordinate of a point $x \in \mathbb{R}^n$. Denote the corresponding function by φ^i. Then

$$\ell_{(\varphi,v)}(\varphi^i) = v^i$$

Therefore $\ell_{(\varphi,v)}(\varphi^i) = 0$ for all i implies $v = 0$.

The proof of surjectivity of the map $(\varphi, v) \mapsto \ell_{(\varphi,v)}$ is more difficult, see e. g. [Spi79, Theorem 3.3]. □

Exercise. Let $\gamma\colon \mathbb{R} \to M$ be a path such that $\gamma(0) = p$. Describe the derivation at p that corresponds to γ.

Definition 1.6. Let $F\colon M \to N$ be a smooth map between smooth manifolds. The *differential of* F *at the point* $p \in M$ is a linear map $DF_p\colon T_pM \to T_{F(p)}N$ defined as

$$DF_p(\varphi, v) = \big(\psi, D(\psi F \varphi^{-1})_{\varphi(p)}(v)\big)$$

where φ is a chart on M around p, and ψ is a chart on N around $F(p)$.

Once again, the chain rule ensures that DF_p is canonical, that is independent of the choice of charts.

Exercise. How does the differential of a smooth map acts on the equivalence classes of curves? How does it act on the derivations?

In view of the Definition 1.4(c), for a vector $X \in T_pM$ we will denote by $X(f)$ the image of the function f under the corresponding derivation at p.

1.3 The commutator of vector fields

Definition 1.7. The *tangent bundle* of a smooth manifold M is the union of the tangent spaces at all of its points:

$$TM = \bigcup_{p \in M} T_pM$$

Definition 1.8 (a)**.** A *vector field* on M is a choice of a tangent vector at every point of M, smoothly depending on the point. That is, a vector field is a map

$$X\colon M \to TM, \quad X(p) \in T_pM$$

A smooth dependence on the point means that if we represent X in a chart φ as $X(p) = (\varphi, v(p))$, then the map $v \circ \varphi^{-1}\colon \varphi(U) \to \mathbb{R}^n$ is of class C^∞.

Definition 1.9 (b)**.** A *vector field* on M is a map $X\colon C^\infty(M) \to C^\infty(M)$ that is $\mathbb{R}$-linear and satisfies the Leibniz rule:

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

The way from (a) to (b): given a map $X\colon M \to TM$, define a map $X\colon C^\infty(M) \to C^\infty(M)$ by

$$X(f)(p) = X(p)(f)$$

The way from (b) to (a): evaluating the function $X(f)$ at a point p yields a derivation $C^\infty(M) \to \mathbb{R}$ at p, that is an element of T_pM. One also needs to check that X is smooth in the sence of (a) if and only if it maps (in the sense of (b)) every smooth function to a smooth function.

Often we will consider vector fields defined only on some subset of M.

Definition 1.10. The *coordinate vector fields* with respect to a chart (U, φ) are the vector fields $(\varphi, e_i)_{i=1}^n$, where $(e_1, \ldots, e_n)$ is the standard basis of $\mathbb{R}^n$. Notation:

$$(\varphi, e_i) = \frac{\partial}{\partial x^i} \text{ or } \partial_i$$

- The notation $\frac{\partial}{\partial x^i}$ is due to the fact that the derivation corresponding to the i-th basis vector field sends a function f to $\frac{\partial(f\circ\varphi^{-1})}{\partial x^i}$.
- Coordinate vector fields form a basis at every point $p \in M$. Therefore every vector field on U can be written as

$$X = \sum_{i=1}^n v^i \frac{\partial}{\partial x^i}, \quad v^i \in C^\infty(U) \text{ for all } i$$

Problem 1.11. *Let $(X_1, \ldots, X_n)$ be a frame field on M, that is a collection of vector fields such that at every $p \in M$ the vectors $X_1(p), \ldots, X_n(p)$ form a basis of T_pM.*

The coordinate vector fields form a frame field. Is the converse true: does every frame field consist of the coordinate vector fields of some chart? If not, what are the necessary and sufficient conditions for this being the case?

This problem will be solved on the next pages.

Theorem-Definition 1.12. *Let X and Y be two vector fields on M. Then the map*

$$f \mapsto X(Y(f)) - Y(X(f))$$

from $C^\infty(M)$ to $C^\infty(M)$ is $\mathbb{R}$-linear and satisfies the Leibniz rule.

The corresponding vector field is denoted by $[X, Y]$ and is called the commutator *of X and Y or the* Lie bracket *of X and Y.*

Proof. **A coordinate-free proof.** We have

$$(X(Y(fg)) - Y(X(fg))) =$$

$$= f \cdot (X(Y(g)) - Y(X(g))) + (X(Y(f)) - Y(X(f)))g$$

Thus $X \circ Y - Y \circ X$ satisfies the Leibniz rule. The $\mathbb{R}$-linearity is obvious.
A proof in local coordinates. Let $X = \sum_{i=1}^n v^i \frac{\partial}{\partial x^i}$ and $Y = \sum_{i=1}^n w^i \frac{\partial}{\partial x^i}$. Then we have

$$Y(f) = \sum_{j=1}^n w^j \frac{\partial f}{\partial x^j}$$

$$X(Y(f)) = \sum_{i=1}^{n} v^i \frac{\partial}{\partial x^i} \left(\sum_{j=1}^{n} w^j \frac{\partial f}{\partial x^j} \right) = \sum_{i,j} v^i \frac{\partial w^j}{\partial x^i} \frac{\partial f}{\partial x^j} + \sum_{i,j} v^i w^j \frac{\partial^2 f}{\partial x^i \partial x^j}$$

Subtracting the similar expression for $Y(X(f))$ we see that the terms involving the second derivatives of f cancel and we obtain

$$X(Y(f)) - Y(X(f)) = \sum_{i,j} \left(v^i \frac{\partial w^j}{\partial x^i} - w^i \frac{\partial v^j}{\partial x^i} \right) \frac{\partial f}{\partial x^j}$$

The right hand side is the derivative of f with respect to a vector field. □

From the second proof we obtain a formula for the components of $[X, Y]$ in the local coordinates:

$$[X, Y]^j = \sum_{i=1}^{n} \left(v^i \frac{\partial w^j}{\partial x^i} - w^i \frac{\partial v^j}{\partial x^i} \right) \tag{1.2}$$

Lemma 1.13. *For every chart the corresponding coordinate vector fields commute.*

Proof. Just the commutation of partial derivatives or the formula (1.2). □

Exercise. Give an example of two non-commuting vector fields in $\mathbb{R}^2$.

The vector fields from your example cannot be coordinate vector fields of a chart. Thus the answer to the first question in Problem 1.11 is negative.

Lemma 1.14. *The commutator has the following properties.*

1. *It is $\mathbb{R}$-bilinear:*

$$[X_1 + X_2, Y] = [X_1, Y] + [X_2, Y] \ \textit{etc.}$$

2. *It is antisymmetric:*

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

3. *It is a derivation in the second and an "antiderivation" in the first argument:*

$$[X, gY] = X(g)Y + g[X, Y]$$
$$[fX, Y] = -Y(f)X + f[X, Y]$$

4. *It satisfies the Jacobi identity:*

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

Proof. □

1.4 Differential 1-forms

Definition 1.15. The *cotangent space at the point* p of M is the dual of the tangent space at p. Notation:

$$T_p^* M = (T_p M)^*$$

Every element of $T_p^* M$ is a linear functional on the vector space $T_p M$.

The *cotangent bundle of* M is the union of the cotangent spaces at all points:

$$T^* M = \bigcup_{p \in M} T_p^* M$$

Definition 1.16 (a)**.** A *differential* 1*-form* on M is a choice of a linear functional on each of the tangent spaces $T_p M$, smoothly depending on p.

Denote by $\mathcal{X}(M)$ the space of all vector fields, and by $\Omega^1(M)$ the space of all differential 1-forms on M. The pointwise pairing between dual vector spaces defines a map

$$\Omega^1(M) \times \mathcal{X}(M) \to C^\infty(M), \quad (X, \omega)(p) = \omega(p)\,(X(p))$$

In other words, every 1-form defines a map $\mathcal{X}(M) \to C^\infty(M)$. This provides a basis for an equivalent definition.

Definition 1.16 (b)**.** A *differential* 1*-form* on M is a $C^\infty(M)$-linear map $\mathcal{X}(M) \to C^\infty(M)$.

Lemma 1.17. *Definitions 1.16(a) and 1.16(b) are equivalent.*

Proof. Let $\omega = \omega(p)$ be a family of linear functionals on $T_p M$. Define a map

$$\ell_\omega \colon \mathcal{X}(M) \to C^\infty(M), \quad \ell_\omega(X)(p) = \omega(p)(X(p))$$

Check the $C^\infty(M)$-linearity of ℓ_ω:

$$\ell_\omega(fX)(p) =$$

$$= (f\ell_\omega)(X)(p)$$

In the opposite direction, let $\ell \colon \mathcal{X}(M) \to C^\infty(M)$ be a $C^\infty(M)$-linear map. Define a linear functional $\omega(p) \in T_p^* M$ as follows. For every tangent vector $v \in T_p M$ take a vector field X such that $X(p) = v$ (why does such a field exist?) and put

$$\omega(p)(v) = \ell(X)(p)$$

One has to check that this is well-defined, that is for any two vector fields X and Y with $X(p) = Y(p)$ we have $\ell(X)(p) = \ell(Y)(p)$. Due to the additivity of ℓ, this is equivalent to showing that $X(p) = 0$ implies $\ell(X)(p) = 0$.

Let's make a simplifying assumption that $X = 0$ outside of a coordinate neighborhood U. Then $X = \sum_{i=1}^n X^i \partial_i$, where each $X^i \in C^\infty(M)$ has support in U and satisfies $X^i(p) = 0$. Due to $C^\infty(M)$-linearity of ℓ we have

$$\ell(X) = \sum_{i=1}^n X^i \ell(\partial_i)$$

In particular, $\ell(X)(p) = \sum_{i=1}^n X^i(p)\ell(\partial_i)(p) = 0$. □

Restore the missing details yourself or find them in [Spi79, Theorem 4.2].

Definition 1.18. Let $f \in C^\infty(M)$. The *differential* of the function f is a 1-form on M determined by

$$df(X)(p) = X(f)(p)$$

Problem 1.19. *Is every* 1*-form on* M *the differential of a function?*

We will see that Problems 1.11 and 1.19 are related.

As in the case of vector fields, often we will consider 1-forms defined only on an open subset of M.

Definition 1.20. The *coordinate* 1*-forms* or *basis* 1*-forms* with respect to a chart (U, φ) are the differentials of the coordinate functions φ^i.

- By abuse of notation, the function φ^i is also denoted by x^i (although $x^i\colon \mathbb{R}^n \to \mathbb{R}$, but we are implicitly identifying U with its image $\varphi(U)$). Hence the notation for the basis 1-forms: dx^i, $i = 1, \dots, n$.

- At every point p, the basis $(dx^1, \dots, dx^n)$ of T_p^*M is dual to the basis $\left(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}\right)$ of T_pM:
$$dx^i(\partial_j) = \delta^i_j$$

- Every 1-form on U can be written in local coordinates as
$$\omega = \sum_{i=1}^n a_i dx^i$$
Besides, $a_i = \omega\left(\frac{\partial}{\partial x^i}\right)$.

- A coordinate presentation of the differential of a function is
$$df = \sum_{i=1}^n \frac{\partial f}{\partial x^i} dx^i$$

Perform a coordinate change $y = F(x)$. Let us use for the entries of the Jacobi matrix of F the notation $\frac{\partial F^i}{\partial x^j} = \frac{\partial y^i}{\partial x^j}$. The transformation law for the vector components can be then written as

$$w^i = \sum_{j=1}^{n} \frac{\partial y^i}{\partial x^j} v^j \tag{1.3}$$

We will now find the transformation law for the components of a 1-form.

Lemma 1.21. *Let $y = y(x)$ be a local coordinate change on a smooth manifold, and let*

$$\omega = \sum_{i=1}^{n} a_i dx^i = \sum_{i=1}^{n} b_i dy^i$$

be a 1-form. Then we have

$$b_i = \sum_{j=1}^{n} a_j \frac{\partial x^j}{\partial y^i} \tag{1.4}$$

where $x = x(y)$ is the inverse function to $y = y(x)$.

All of the quantities in the above lemma are functions of a point $x \in \mathbb{R}^n$ or of a point $y \in \mathbb{R}^n$. For example, the last equation means

$$b_i(y) = \sum_{j=1}^{n} a_j(x(y)) \frac{\partial x^j}{\partial y^i}(y)$$

Proof. For every vector $X = \sum_{i=1}^{n} v^i \frac{\partial}{\partial x^i} = \sum_{i=1}^{n} w^i \frac{\partial}{\partial y^i}$ we have

$$\sum_{i=1}^{n} a_i v^i = \omega(X) = \sum_{i=1}^{n} b_i w^i$$

Substituting (1.3), we get

$$\sum_{i=1}^{n} a_i v^i = \sum_{i,j=1}^{n} b_i \frac{\partial y^i}{\partial x^j} v^j$$

Since this must hold for every $v \in \mathbb{R}^n$, the coefficients at each v^i on both sides must coincide. Thus we have

$$a_i = \sum_{j=1}^{n} b_j \frac{\partial y^j}{\partial x^i}$$

or $a^\top = b^\top DF$, where DF is the Jacobi matrix of $y = y(x)$. Multiplying on the left with $(DF)^{-1} = D(F^{-1})$ we get $b^\top = a^\top D(F^{-1})$ or

$$b_i = \sum_{j=1}^{n} a_j \frac{\partial x^j}{\partial y^i}$$

$\square$

The fact that the formulas (1.3) and (1.4) are different implies that there is no canonical correspondence between the tangent vectors and 1-forms. If a vector and a form have the same components in one basis, they will have different components after (almost every) coordinate change.

1.5 Differential k-forms and the exterior derivative

Recall that an *multilinear form* on a vector space V is a map

$$\underbrace{V \times \cdots \times V}_{k} \to \mathbb{R}$$

that is linear in each argument. A multilinear form is called *skew-symmetric* or *alternating* if its value changes the sign for every transposition of two arguments.

Denote by $\otimes^k(V^*)$ the space of all k-linear forms, and by $\wedge^k(V^*)$ the space of all alternating k-linear forms. Both of them are vector spaces.

Let us recall some facts.

- There is a tensor product operation $\otimes^k(V^*) \times \otimes^l(V^*) \to \otimes^{k+l}(V^*)$ sending (S, T) to a multilinear form $S \otimes T$ that acts as

  $$(S \otimes T)(v_1, \ldots, v_{k+l}) = S(v_1, \ldots, v_k) T(v_{k+1}, \ldots, v_{k+l})$$

- If $(\theta^1, \ldots, \theta^n)$ is a basis of V^*, then the forms

  $$\theta^{i_1} \otimes \cdots \otimes \theta^{i_k}, \quad 1 \le i_1, \ldots, i_k \le n$$

  constitute a basis for $\otimes^k(V^*)$.

- There is a linear map $\mathrm{Alt} : \otimes^k(V^*) \to \wedge^k(V^*)$ defined by

  $$(\mathrm{Alt}\, T)(v_1, \ldots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sgn} \sigma \cdot T(v_{\sigma(1)}, \ldots, v_{\sigma(k)})$$

  This map is a projection: $\mathrm{Alt}|_{\wedge^k(V^*)} = \mathrm{id}$, $\mathrm{Alt} \circ \mathrm{Alt} = \mathrm{Alt}$.

- The tensor product composed with the alternation map define the wedge product:

  $$\omega \wedge \eta = \frac{(k+l)!}{k!l!} \mathrm{Alt}(\omega \otimes \eta) \in \Lambda^{k+l}(V^*) \quad \text{for } \omega \in \Lambda^k(V^*), \eta \in \Lambda^l(V^*)$$

  The wedge product is bilinear, associative, and anti-commutative:

  $$\omega \wedge \eta = (-1)^{kl} \eta \wedge \omega$$

- In particular, we can form the wedge products of the basis covectors:

$$\theta^{i_1} \wedge \cdots \wedge \theta^{i_k} = \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma)\, \theta^{i_{\sigma(1)}} \otimes \cdots \otimes \theta^{i_{\sigma(k)}}$$

  We have

$$(\theta^{i_1} \wedge \cdots \wedge \theta^{i_k})(e_{i_1}, \ldots, e_{i_k}) = 1,$$

  where $(e_1, \ldots, e_n)$ is the basis of V dual to $(\theta^1, \ldots, \theta^n)$.

- The anticommutativity of the wedge product implies

$$\theta^{i_1} \wedge \cdots \wedge \theta^{i_k} = (\operatorname{sgn} \sigma) \theta^{i_{\sigma(1)}} \wedge \cdots \wedge \theta^{i_{\sigma(k)}},$$

  so that the forms

$$(\theta^{i_1} \wedge \cdots \wedge \theta^{i_k})_{1 \le i_1 < \cdots < i_k \le n}$$

  constitute a basis of $\Lambda^k(V^*)$. In particular, $\dim \Lambda^k(V^*) = \binom{n}{k}$.

Exercise. Is it true that $\omega \wedge \omega = 0$ for every $\omega \in \Lambda^k(V^*)$ with $k > 0$?

Definition 1.22 (a)**.** A *differential k-form* on a smooth manifold M is a choice of an element of $\Lambda^k(T_p^* M)$ at every $p \in M$, smoothly depending on the point p.

Definition 1.22 (b)**.** A *differential k-form* on a smooth manifold M is an alternating C^∞-multilinear map

$$\underbrace{\mathcal{X}(M) \times \cdots \times \mathcal{X}(M)}_{k} \to C^\infty(M)$$

The equivalence of the two definitions is proved similarly to the equivalence of definitions 1.16(a) and 1.16(b), see Lemma 1.17.

The space of differential k-forms on M is denoted by $\Omega^k(M)$. For $\omega \in \Omega^k(M)$ we write

$$\omega|_p(X_1, \ldots, X_k) \quad \text{or} \quad \omega(X_1, \ldots, X_k)(p)$$

for the value of ω on the vector fields $X_1, \ldots, X_k$ at the point p. Note that this value depends only on the vectors $X_1(p), \ldots X_k(p) \in T_p M$, and is independent of the values of the vector fields X_i at the other points.

The wedge product of two differential forms is defined as their pointwise wedge product:

$$(\omega \wedge \eta)|_p = \omega|_p \wedge \eta|_p$$

Note that we can put $\Omega^0(M) = C^\infty(M)$ and extend the definition of the wedge product by

$$f \wedge \omega = f\omega$$

For any chart around a point p, the forms on T_pM

$$(dx^{i_1} \wedge \cdots \wedge dx^{i_k})|_p, \quad 1 \leq i_1 < \cdots < i_k \leq n$$

constitute a basis of $\Lambda^k(T_p^*M)$. If we decompose a k-form $\omega \in \Omega^k(M)$ at every p with respect to the above basis, we obtain

$$\omega = \sum_{i_1 < \cdots < i_k} a_{i_1 \ldots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k} \tag{1.5}$$

with $a_{i_1 \ldots i_k} \in C^\infty(M)$.

Definition 1.23. Let $F \colon M \to N$ be a smooth map between smooth manifolds, and let $\omega \in \Omega^k(N)$ be a differential k-form on N. The *pullback* $F^*\omega$ of ω by F is a k-form on M acting on vectors by

$$F^*\omega(X_1, \ldots, X_k) = \omega(dF(X_1), \ldots, dF(X_k))$$

In particular, for $k = 0$ we have $F^*f = f \circ F$.

Note that F need not to be a diffeomorphism, and the dimension of M may be different from that of N.

One can check that $F^*\omega$ is a differential form by using either of our two definitions: pointwise $\mathbb{R}$-multilinearity or global $C^\infty(M)$-multilinearity. The alternating property is immediate.

It is easy to prove the compatibility of the pullback with the wedge product:

$$F^*(\omega \wedge \eta) = F^*\omega \wedge F^*\eta$$

Up to now everything was just the "pointwise linear algebra". Here comes an operation that involves not just forms and vector field values at a point, but also their behavior around the point.

Definition 1.24 (a)**.** The *exterior derivative* of a k-form (1.5) is the $(k+1)$-form

$$d\omega = \sum_{i_1 < \cdots < i_k} da_{i_1 \ldots i_k} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \tag{1.6}$$

It is not immediately clear that the definition of $d\omega$ is independent of the choice of coordinates. The coordinate independence can be proved in an axiomatic way.

Lemma-Definition 1.24 (b)**.** *The exterior derivative* $d \colon \Omega^k(U) \to \Omega^{k+1}(U)$ *is the unique map that has the following properties.*

1. $d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2$
2. $d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^{\deg \omega_1} \omega_1 \wedge d\omega_2$
3. $d \circ d = 0$

4. $d\colon \Omega^0(U) \to \Omega^1(U)$ *is the function differential.*

For a proof, see [Spi79, Propositions 7.10-11].

In particular, the differentials of ω computed in different coordinates on U coincide. This also implies that the map $d\colon \Omega^k(M) \to \Omega^{k+1}(M)$ is well-defined and is a unique map that satisfies the four properties stated in the lemma.

Exercise. Give an example of a 1-form on $\mathbb{R}^2$ that is not the differential of a function.

Your example provides a negative answer to Problem 1.19.

Exercise. Compute the differential of the form

$$\omega = \sum_{i=1}^{n} f_i \, dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n \in \Omega^{n-1}(\mathbb{R}^n)$$

Another way to prove the coordinate independence of Definition 1.24(a) is to show that it is equivalent to a coordinate-free formula.

Lemma-Definition 1.24 (c)**.** *Formula* (1.6) *is equivalent to*

$$\begin{aligned} d\omega(X_0, \ldots, X_k) &= \sum_{i=0}^{k} (-1)^i X_i(\omega(X_0, \ldots, \widehat{X_i}, \ldots, X_k)) \\ &+ \sum_{0 \le i < j \le k} (-1)^{i+j} \omega([X_i, X_j], X_0, \ldots, \widehat{X_i}, \ldots, \widehat{X_j}, \ldots, X_k) \end{aligned} \tag{1.7}$$

which can be used as a coordinate-free definition of $d\colon \Omega^k(M) \to \Omega^{k+1}(M)$.

Proof. We give a proof for $k = 1$ only, the general case is similar. For $k = 1$ the formula says

$$d\omega(X, Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([X, Y])$$

A priori it is not clear that this determines a differential form: the right hand side seems to depend on the local behaviour of X and Y, not just on their values at a point. Let us show that the map $d\omega$ is $C^\infty(M)$-bilinear.

$$\begin{aligned} d\omega(fX, Y) &= fX(\omega(Y)) - Y(\omega(fX)) - \omega([fX, Y]) \\ &= fX(\omega(Y)) - Y(f)\omega(X) - fY(\omega(X)) - \omega(-Y(f)X + f[X, Y]) \\ &= f \cdot (X(\omega(Y)) - Y(\omega(X)) - \omega([X, Y])) = f d\omega(X, Y) \end{aligned}$$

The additivity is clear, and the scaling property wrt Y is proved similarly. Thus formula (1.7) defines indeed a differential 2-form. In order to show that this form can be also computed by formula (1.6), compare the values at the pairs (∂_i, ∂_j) of coordinate vector fields. Both formulas yield $\frac{\partial a_j}{\partial x^i} - \frac{\partial a_i}{\partial x^j}$. $\square$

Lemma 1.25. *The exterior derivative commutes with the pullback operation:*

$$d(F^*\omega) = F^*(d\omega)$$

Proof. Induction on the degree k of ω. For $k = 0$ we have to prove

$$d(f \circ F) = F^*(df),$$

which holds by the chain rule. For $k = 1$ it suffices to consider $\omega = a_i dx^i$. Since the pullback is compatible with the (wedge) product, we have

$$F^*\omega = F^*a_i F^*(dx^i) = F^*a_i d(F^*x^i)$$

where we also used the already proved case $k = 0$. Hence

$$\begin{aligned} d(F^*\omega) &= d(F^*a_i) \wedge d(F^*x^i) + (F^*a_i)d^2(F^*x^i) \\ &= F^*(da_i) \wedge F^*(dx^i) = F^*(da_i \wedge dx^i) = F^*(d\omega) \end{aligned}$$

The induction step from k to $k + 1$ is done similarly.

□

By Lemma-Definition 1.24(b), if $\omega = d\eta$, then $d\omega = 0$. A partial converse is true.

Lemma 1.26 (Poincaré lemma)**.** *If $d\omega = 0$ for some $\omega \in \Omega^k(M)$, then for every subset $U \subset M$ there is a differential form $\eta \in \Omega^{k-1}(U)$ such that $\omega = d\eta$ on U.*

Now we are ready to solve Problem 1.11, at least locally.

Theorem 1.27. *Let $(X_1, \dots, X_n)$ be a frame field with $[X_i, X_j] = 0$ for all i, j. Then locally $(X_1, \dots, X_n)$ are the coordinate vector fields with respect to some chart.*

Proof. Let $\omega^1, \dots, \omega^n$ be the dual coframe. We have

$$\begin{aligned} d\omega^i(X_j, X_k) &= X_j(\omega^i(X_k)) - X_k(\omega^i(X_j)) - \omega^i([X_j, X_k]) \\ &= X_j(\delta^i_k) - X_k(\delta^i_j) - \omega^i(0) = 0 \end{aligned}$$

for all i, j, k. It follows that $d\omega^i = 0$ for all i. Hence for every simply-connected $U \subset M$ there are functions $\varphi^1, \dots, \varphi^n : U \to \mathbb{R}$ such that $\omega^i = d\varphi^i$ for all i. Since the linear functionals $d\varphi^1(p), \dots, d\varphi^n(p) \in T^*_p(M)$ are linearly independent, the map $\varphi : U \to \mathbb{R}^n, q \mapsto (\varphi^1(q), \dots, \varphi^n(q))$ has a non-degenerate Jacobi matrix at p. Hence for some neighborhood U' of p, $\varphi : U' \to \mathbb{R}^n$ is a homeomorphism onto the image. Since the coordinate vector fields of φ form the basis dual to $(d\varphi^1, \dots, d\varphi^n)$, they coincide with the vector fields $X_1, \dots, X_n$. □

1.6 Volume forms and orientability

For every n-dimensional vector space V we have

$$\dim \Lambda^n(V^*) = \binom{n}{n} = 1 \tag{1.8}$$

To obtain a non-zero element of $\Lambda^n(V^*)$, it suffices to take a basis $(\theta_1, \ldots, \theta_n)$ of V^* and form the wedge product $\theta_1 \wedge \cdots \wedge \theta_n$. We have $\theta_1 \wedge \cdots \wedge \theta_n \neq 0$ because $(\theta_1 \wedge \cdots \wedge \theta_n)(e_1, \ldots, e_n) = 1$ for the dual basis. Due to (1.8), every element of $\Lambda^n(V^*)$ has the form $\lambda\theta_1 \wedge \cdots \wedge \theta_n$ for some $\lambda \in \mathbb{R}$.

For differential n-forms on an n-dimensional manifold M this has the following consequence.

Lemma 1.28. *Every differential n-form ω in every local chart (V, φ) can be written as $\omega(p) = a(p)dx^1 \wedge \cdots \wedge dx^n$ for some $a \in C^\infty(M)$.*

Definition 1.29. A nowhere vanishing differential n-form on an n-dimensional smooth manifold is called a *volume form*.

Problem 1.30. *What manifolds M admit volume forms?*

Definition 1.31. Two bases of a vector space V are said to have the same orientation if the matrix of the basis change has a positive determinant. This defines an equivalence relation on the set of all bases of V. An *orientation* of a vector space V is a choice of one of the equivalence classes.

Now, every local chart (U, φ) on a smooth manifold orients all tangent spaces T_pM, $p \in U$ at once. Namely, take at every p the orientation of the coordinate basis $(\partial_1, \ldots, \partial_n)$.

Definition 1.32. Two charts (U, φ) and (V, ψ) are called compatible if they define the same orientation on T_pM for every $p \in U \cap V$. A manifold M is called *orientable* if it has an atlas of pairwise compatible charts.

Lemma 1.33. *Let $\omega \in \Lambda^n(V^*)$, and let $(e_1, \ldots, e_n)$ be a basis of V. Then for every collection of vectors*

$$v_i = \sum_{j=1}^{n} v_i^j e_j$$

we have

$$\omega(v_1, \ldots, v_n) = \det(v_i^j) \cdot \omega(e_1, \ldots, e_n)$$

Proof.

$$\omega(v_1, \ldots, v_n) = \omega\left(\sum_{j=1}^{n} v_1^j, \ldots, \sum_{j=1}^{n} v_n^j\right) = \sum_{\sigma \in S_n} v_1^{\sigma(1)} \cdots v_n^{\sigma(n)} \omega(e_{\sigma(1)}, \ldots e_{\sigma(n)})$$

$$= \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) v_1^{\sigma(1)} \cdots v_n^{\sigma(n)} \omega(e_1, \ldots, e_n) = \det(v_i^j) \cdot \omega(e_1, \ldots, e_n)$$

□

Theorem 1.34. *Let $F\colon M \to N$ be a smooth map between n-dimensional smooth manifolds. Then for any local coordinates $(x^1, \dots, x^n)$ on M and $(y^1, \dots, y^n)$ on N we have*

$$F^*(b\, dy^1 \wedge \cdots \wedge dy^n) = (b \circ F) \det \left(\frac{\partial (y^i \circ F)}{\partial x^j} \right) dx^1 \wedge \cdots \wedge dx^n$$

Proof. Observe that $F^*(b\eta) = F^*(b)F^*(\eta) = (b \circ F)F^*(\eta)$. Therefore it remains to deal with the case $b = 1$.

For every n-form $\omega = adx^1 \dots dx^n$ the coefficient a can be determined by evaluating ω on the standard basis:

$$a = \omega \left(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n} \right)$$

So let us compute:

$$\begin{aligned} F^*(dy^1 \wedge \cdots \wedge dy^n) & \left(\frac{\partial}{\partial x^1}, \dots \frac{\partial}{\partial x^n} \right) \\ &= (dy^1 \wedge \cdots \wedge dy^n) \left(DF \left(\frac{\partial}{\partial x^1} \right), \dots, DF \left(\frac{\partial}{\partial x^n} \right) \right) \\ &= (dy^1 \wedge \cdots \wedge dy^n) \left(\sum_{j=1}^n \frac{\partial (y^j \circ F)}{\partial x^1} \frac{\partial}{\partial y^j}, \dots, \sum_{j=1}^n \frac{\partial (y^j \circ F)}{\partial x^n} \frac{\partial}{\partial y^j} \right) \\ &= \det \left(\frac{\partial (y^i \circ F)}{\partial x^j} \right) \end{aligned}$$

□

Corollary 1.35. *Let $F\colon M \to M$ be a smooth map. Then for every differential n-form ω we have*

$$(F^*\omega)(p) = \det(DF)(p)\omega(F(p))$$

Proof. Apply Theorem 1.34 to $M = N$ and $y^i = x^i$. □

Corollary 1.36. *Write a differential n-form ω on a smooth n-dimensional manifold M in two different coordinate systems:*

$$adx^1 \wedge \cdots \wedge dx^n = \omega = bdy^1 \wedge \cdots \wedge dy^n$$

Then $a = b \det \left(\frac{\partial y^i}{\partial x^j} \right)$.

Proof. Apply Theorem 1.34 to the identity map $M \to M$. □

Theorem 1.37. *A smooth manifold M is orientable if and only if there is a volume form $\omega \in \Omega^n(M)$.*

Proof. Assume there is $\omega \in \Omega^n(M)$ with $\omega(p) \neq 0$ for all p. Take an atlas $(U_\alpha, \varphi_\alpha)$ on M such that all U_α are connected. Call a chart of this atlas *good* if $\omega(\partial_1, \ldots, \partial_n) > 0$ for the coordinate fields of the chart; call it *bad* if $\omega(\partial_1, \ldots, \omega_n) < 0$. Corollary 1.36 implies that any two good charts are compatible. Every bad chart can be turned into a good one by composing it with a linear map $\mathbb{R}^n \to \mathbb{R}^n$ with a negative determinant (for example, permuting two coordinates). Making all bad charts good we obtain an atlas of pairwise compatible charts.

In the inverse direction, let $(U_\alpha, \varphi_\alpha)$ be an atlas of pairwise compatible charts. Put $\omega_\alpha = dx_\alpha^1 \wedge \cdots \wedge dx_\alpha^n$. Each ω_α is a nowhere vanishing n-form on U_α; besides, on $U_\alpha \cap U_\beta$ we have $\omega_\alpha = c\omega_\beta$ with $c > 0$. We need to find a way to patch all forms ω_α together.

This can be done with the help of a *partition of unity* subordinate to $\{U_\alpha\}$. This is a collection of functions $\rho_\alpha \colon M \to \mathbb{R}$ such that

$$\operatorname{supp} \rho_\alpha \subset U_\alpha, \quad \rho_\alpha \geq 0, \quad \sum_\alpha \rho_\alpha = 1$$

Due to $\operatorname{supp} \rho_\alpha \subset U_\alpha$, the product $\rho_\alpha \omega_\alpha$ can be viewed as an n-form on M. Now it is easy to check that

$$\omega = \sum_\alpha \omega_\alpha$$

is a nowhere vanishing differential n-form. □

1.7 Integration on manifolds

Let $f \colon \mathbb{R}^n \to \mathbb{R}$ be a continuous function with compact support. Then there is the Riemann integral

$$\int_{\mathbb{R}^n} f(x)\, dx = \int_{\mathbb{R}^n} f(x^1, \ldots, x^n)\, dx^1 \ldots dx^n$$

Sometimes we use a non-standard coordinate system in order to compute the integral. For this we use the following formula:

$$\int_{\mathbb{R}^n} f(x)\, dx = \int_{\mathbb{R}^n} f(F(y)) |\det(DF)(y)|\, dy, \tag{1.9}$$

where $x = F(y)$ is a substitution. This looks very similar to the formula in Corollary 1.35. One difference is the absolute value. To avoid problems, we first define integration on oriented manifolds only.

To define the integral of $\omega \in \Omega^n(M)$ we will "push" this form into the Euclidean space and integrate there in a usual way. Since ω cannot be pushed as a whole, we need first to split it into parts.

For n-forms on $\mathbb{R}^n$ we define the integral as

$$\int_{\mathbb{R}^n} f(x) dx^1 \wedge \cdots \wedge dx^n = \int_{\mathbb{R}^n} f(x) dx$$

Definition 1.38. Let ω be a differential n-form with a compact support on a smooth oriented n-dimensional manifold M. Choose an atlas $(U_\alpha, \varphi_\alpha)$ and a partition of unity $\{\rho_\alpha\}$ subordinate to $\{U_\alpha\}$ and put

$$\int_M \omega = \sum_\alpha \int_{\mathbb{R}^n} (\varphi_\alpha^{-1})^*(\rho_\alpha \omega)$$

Lemma 1.39. *The value of $\int_M \omega$ does not depend on a choice of an atlas and of a partition of unity.*

Proof. Let (V_β, ψ_β) be a different atlas with a subordinate partition of unity $\{\sigma_\beta\}$. Then $\{\tau_{\alpha\beta}\}$ with $\tau_{\alpha\beta} = \rho_\alpha\sigma_\beta$ is also a partition of unity. Due to $\rho_\alpha = \sum_\beta \tau_{\alpha\beta}$ and $\sigma_\beta = \sum_\alpha \tau_{\alpha\beta}$ we have

$$\int_M \omega = \sum_\alpha \int_{\mathbb{R}^n} (\varphi_\alpha^{-1})^*(\rho_\alpha \omega) = \sum_{\alpha,\beta} \int_{\mathbb{R}^n} (\varphi_\alpha^{-1})^*(\tau_{\alpha\beta}\omega)$$

and a similar formula for integration by means of $(V_\beta, \psi_\beta, \sigma_\beta)$. It suffices to show that

$$\int_{\mathbb{R}^n} (\varphi_\alpha^{-1})^*(\tau_{\alpha\beta}\omega) = \int_{\mathbb{R}^n} (\psi_\alpha^{-1})^*(\tau_{\alpha\beta}\omega)$$

or, in a simplified notation: for any two diffeomorphisms (onto image) $\varphi, \psi\colon W \to \mathbb{R}^n$ and any $\omega \in \Omega^n(W)$ show that

$$\int_{\mathbb{R}^n} (\varphi^{-1})^*\omega = \int_{\mathbb{R}^n} (\psi^{-1})^*\omega \tag{1.10}$$

We have $(\psi^{-1})^*\omega = (\varphi\circ\psi^{-1})^*((\varphi^{-1})^*\omega)$. Denote $F = \varphi\circ\psi^{-1}$ and $(\varphi^{-1})^*\omega = f(x)dx^1 \wedge \cdots \wedge dx^n$. Then by Corollary 1.35 we have

$$((\psi^{-1})^*\omega)(y) = \det(DF)(y) f(F(y)) dy^1 \wedge \cdots \wedge dy^n,$$

so that (1.10) follows from (1.9). □

In a similar way one proves the following.

Lemma 1.40. *For any diffeomorphism $F\colon M \to N$ between n-dimensional smooth manifolds we have $\int_M F^*\omega = \int_N \omega$ for any $\omega \in \Omega^n(N)$.*

Once we fix a volume form ω, we can define the integral of a function f (with respect to the volume form ω) to be the integral of the n-form $f\omega$.

Definition 1.41. A topological manifold *with boundary* is defined in the same way as a manifold without boundary, only this time a chart φ_α maps U_α homeomorphically onto an open subset of either $\mathbb{R}^n$ or $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x^1 \geq 0\}$.

A smooth manifold with boundary is defined again in the same way, only we need to say what we mean by smooth maps between non-necessarily open subsets of $\mathbb{R}^n$. These can be defined as maps that are extendable to smooth maps between some open neighborhoods of the sets.

Boundary points of a manifold with boundary are those that are sent to the boundaries of "half-open" sets in $\mathbb{R}^n$. The set of all boundary points is denoted by ∂M.

Lemma 1.42. *Let M be an n-dimensional smooth manifold with boundary. Then ∂M is an $(n-1)$-dimensional smooth manifold without boundary.*

Let M be an oriented n-dimensional smooth manifold with boundary. The *induced orientation* on ∂M is defined as follows: for every $p \in \partial M$ choose local coordinates as in the definition, that is with $x^1 = 0$ corresponding to the boundary points and $x^1 > 0$ away from the boundary. If $(\partial_1, \ldots, \partial_n)$ agrees with the chosen orientation of T_pM, then we orient $T_p\partial M$ *opposite* to the basis $(\partial_2, \ldots, \partial_n)$.

This choice looks a bit arbitrary, but it saves us from possible minus signs in the following theorem.

Theorem 1.43. *Let ω be an $(n-1)$-form on an n-dimensional manifold M. Then*

$$\int_M d\omega = \int_{\partial M} \omega$$

1.8 Tensors and tensor fields

1.8.1 Vectors and covectors

In what follows, V is an n-dimensional real vector space.

We are using the upper and lower indices in the following way:

- The elements of a basis of V are numbered using lower indices, for example $e_1, \ldots, e_n$.
- The components of a vector with respect to a basis are numbered using upper indices, for example $v = (v^1, \ldots, v^n)$.
- The elements of a basis of V^* are numbered using upper indices, for example $\eta^1, \ldots, \eta^n$.
- The components of a covector (linear functional) are numbered using lower indices, for example $\ell = (\ell_1, \ldots, \ell_n)$.

Thus we can write a vector v and a covector ℓ as

$$v = \sum_{i=1}^n v^i e_i, \quad \ell = \sum_{i=1}^n \ell_i \eta^i.$$

The *Einstein summation convention* discards the sum sign. Always when an index appears twice in a term: once as an upper index and once as a lower index, the term is summed over all values of that index. So we will write

$$v = v^i e_i, \quad \ell = \ell_i \eta^i.$$

Here is an example of a calculation with the Einstein summation convention. Let $(\eta^1, \dots, \eta^n)$ be the basis of V^* dual to a basis $(e_1, \dots, e_n)$ of V. This means $\eta^i(e_j) = \delta^i_j$ (the Kronecker delta). For any $v \in V$, $\ell \in V^*$ we have

$$\ell(v) = \ell_i \eta^i(v^j e_j) = \ell_i v^j \eta^i(e_j) = \ell_i v^j \delta^i_j = \ell_i v^i.$$

Assume we have two different bases $(e_1, \dots, e_n)$ and $(f_1, \dots, f_n)$ of V. Then we have

$$f_i = a^j_i e_j \tag{1.11}$$

for certain numbers $a^j_i \in \mathbb{R}$.

Remark. In the matrix language, equation (1.11) can be written as

$$\begin{pmatrix} f_1 \cdots f_n \end{pmatrix} = \begin{pmatrix} e_1 \cdots e_n \end{pmatrix} \begin{pmatrix} a^1_1 & \cdots & a^1_n \\ \vdots & \ddots & \vdots \\ a^n_1 & \cdots & a^n_n \end{pmatrix}$$

It is convenient to write arrays with lower indices as row vectors, and arrays with upper indices as column vectors. The same applies to matrices: the lower index varies along a row, the upper index varies along a column.

Lemma 1.44. *A basis change* (1.11) *results in the following transformation rules for vector and covector components:*

$$w^i = b^i_j v^j, \quad m_i = a^j_i \ell_j$$

Here (v^i) *and* (w^i) *are the coordinates of a vector in the bases* (e_i) *and* (f_i) *respectively, and* (ℓ_i) *and* (m_i) *are the coordinates of a covector in the bases dual to* (e_i) *and* (f_i)*, respectively.*

Proof. The first formula follows from

$$v^i e_i = w^j f_j = w^j e_i a^i_j$$

For the second formula first prove the transformation rule for the dual bases: $\eta^i = a^i_j \varphi^j$. We have

$$\eta^i(f_j) = \eta^i(e_k a^k_j) = \delta^i_k a^k_j = a^i_j$$

On the other hand,

$$a^i_k \varphi^k(f_j) = a^i_k \delta^k_j = a^i_j$$

Two linear functionals that take the same values on all basis elements are equal. Therefore $\eta^i = a^i_k \varphi^k = a^i_j \varphi^j$.

The rest is similar to the argument at the beginning. □

1.8.2 Tensors

Definition 1.45. A *tensor* over V is a multilinear map

$$\alpha\colon \underbrace{V^* \times \cdots \times V^*}_{r} \times \underbrace{V \times \cdots \times V}_{s} \to \mathbb{R}$$

The numbers r and s are called the *contravariant degree* and the *covariant degree* of the tensor α.

Notation for the space of (r,s)-tensors: $T^r_s(V)$ or simply T^r_s if it is clear what vector space is meant.

The space $T^0_1(V)$ consists of the linear functionals $V \to \mathbb{R}$, therefore $T^0_1(V) = V^*$. The space $T^1_0(V)$ consists of the linear functionals on V^*, hence

$$T^1_0(V) = (V^*)^* = V.$$

Definition 1.46. If $\alpha \in T^r_s$ and $\beta \in T^t_u$, then we define their *tensor product* $\alpha \otimes \beta \in T^{r+t}_{s+u}$ as

$$\begin{aligned}(\alpha \otimes \beta)&(\ell^1, \dots, \ell^{r+t}, v_1, \dots, v_{s+u})\\ &= \alpha(\ell^1, \dots, \ell^r, v_1, \dots, v_s) \cdot \beta(\ell^{r+1}, \dots, \ell^{r+t}, v_{s+1}, \dots, v_{s+u})\end{aligned}$$

The tensor product is associative and distributive but not commutative.

From the elements of a basis (e_i) of V and a basis (η^i) of V^* we can form the tensor products

$$e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes \eta^{j_1} \otimes \cdots \otimes \eta^{j_s} \in T^r_s(V). \tag{1.12}$$

Exercise. If $\ell^i = \ell^i_j \eta^j$ and $v_i = v^j_i e_j$, then what value does the tensor (1.12) take on the arguments $(\ell^1, \dots, \ell^r, v_1, \dots, v_s)$?

Exercise. Show that the tensors (1.12) with all indices ranging independently from 1 to n form a basis of $T^r_s(V)$. In particular, $\dim T^r_s(V) = n^{r+s}$.

We define the tensor product of two tensor spaces as

$$T^r_s \otimes T^t_u = T^{r+t}_{s+u}.$$

This notation fits well with Definition 1.46: in view of the last exercise the right hand side is spanned by the elements of the form $\alpha \otimes \beta$ with $\alpha \in T^r_s$, $\beta \in T^t_u$. Because of $T^1_0(V) = V$ and $T^0_1(V) = V^*$ we can now write

$$T^r_s(V) = \underbrace{V \otimes \cdots \otimes V}_{r} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{s}$$

Denote by $\mathrm{Hom}(V, W)$ the space of linear homomorphisms from a vector space V to a vector space W.

Lemma 1.47. *There are natural isomorphisms*

$$T^1_1(V) \cong \mathrm{Hom}(V, V) \cong \mathrm{Hom}(V^*, V^*)$$

Proof. Let $\alpha\colon V^* \times V \to \mathbb{R}$ be a $(1,1)$-tensor. It defines a linear map $\beta\colon V^* \to V^*$ by the rule

$$\beta(\ell)(v) = \alpha(\ell, v).$$

Conversely, every linear map β determines a $(1,1)$-tensor by the above formula.

Similarly, α gives rise to a linear map $\gamma\colon V \to V$: the vector $\gamma(v)$ is uniquely determined by

$$\ell(\gamma(v)) = \alpha(\ell, v) \quad \text{for all } \ell \in V^*.$$

Again, every γ determines an α. □

The above isomorphisms are natural in the sense that we did not make any choices during their construction. By contrary, any two vector spaces of the same dimension are isomorphic, but a construction of an isomorphism usually involves a choice of bases.

Similarly to the above lemma one can construct two natural isomorphisms

$$T^0_2(V) \cong \mathrm{Hom}(V, V^*).$$

One isomorphism takes a tensor α to the map β_1 defined by $\beta_1(v)(w) = \alpha(v, w)$, the other isomorphism takes α to the map β_2 defined by $\beta_2(v)(w) = \alpha(w, v)$.

For tensor spaces of higher rank there are more and more natural isomorphisms. For example, $\alpha \in T^1_2(V)$ can be interpreted (in two different ways) as a bilinear map

$$V^* \times V \to V$$

or as a linear map

$$V^* \to V^* \otimes V^*$$

or (again in two different ways) as a linear map

$$V \to V \otimes V^*.$$

Exercise. Describe the natural isomorphism

$$T^1_2(V) \cong \mathrm{Hom}(V^*, V^* \otimes V^*).$$

1.8.3 Transformation rules

Let $\alpha \in T^1_1$. Pick a basis (e_i) of V; let (η^i) be the dual basis of V^*. Write α with respect to these bases:

$$\alpha = \alpha^i_j e_i \otimes \eta^j$$

What are the components of α if we change the bases to (f_i) and (φ^i)? By Lemma 1.44 we have

$$\alpha = \alpha^i_j (f_k b^k_i) \otimes (a^j_l \varphi^l) = \alpha^i_j b^k_i a^j_l f_k \otimes \varphi^l.$$

As a result, ${\alpha'}^k_l = \alpha^i_j b^k_i a^j_l$ are the components of α in the new bases. (Compare this with the basis change formula for a matrix of a linear transformation.)

A similar argument proves the following.

Lemma 1.48. *The components of an (r, s)-tensor α are transformed under a basis change according to*

$${\alpha'}^{k_1 \ldots k_r}_{l_1 \ldots l_s} = \alpha^{i_1 \ldots i_r}_{j_1 \ldots j_s} b^{k_1}_{i_1} \cdots b^{k_r}_{i_r} a^{j_1}_{l_1} \cdots a^{j_s}_{l_s}$$

1.8.4 Trace and contraction

Definition 1.49. The trace of a $(1, 1)$-tensor α is defined as

$$\operatorname{tr} \alpha = \alpha^i_i,$$

where (α^i_j) are the components of α with respect to any dual pair of bases.

Lemma 1.50. *The trace of a $(1, 1)$-tensor is well-defined.*

Proof. For any other basis we have

$${\alpha'}^i_i = \alpha^j_k b^i_j a^k_i = \alpha^j_k \delta^k_j = \alpha^j_j.$$

□

An (obviously) equivalent definition of the trace:

$$\operatorname{tr} \alpha = \sum_i \alpha(\eta^i, e_i).$$

It is possible to generalize the trace to tensors of higher degree by setting equal one of the upper and one of the lower indices.

Definition 1.51. The *contraction* of a (r, s)-tensor α with respect to the p-th upper and the q-th lower index is the $(r - 1, s - 1)$-tensor β with the components

$$\beta^{i_1 \ldots i_{r-1}}_{j_1 \ldots j_{s-1}} = \alpha^{i_1 \ldots i_{p-1} k i_p \ldots i_r}_{j_1 \ldots j_{q-1} k j_q \ldots j_s}$$

The resulting tensor β is independent of the choice of a basis. We use the notation $\beta = \mathrm{tr}^p_q(\alpha)$. An equivalent definition:

$$\mathrm{tr}^p_q(\alpha)(\ell^1, \ldots, \ell^{r-1}, v_1, \ldots, v_{s-1}) = \sum_{i=1}^n \alpha(\ell^1, \ldots, \eta^i, \ldots, \ell^{r-1}, v_1, \ldots, e_i, \ldots, v_{s-1}),$$

where (e_i) is any basis of V and (η^i) is the dual basis of V^*; the basis vectors and covectors are inserted as the q-th contravariant and the p-th covariant argument.

Lemma 1.52. *For every $v \in V$ and $\ell \in V^*$, the evaluation of ℓ on V is the trace of the tensor product $v \otimes \ell$:*

$$\ell(v) = \mathrm{tr}(v \otimes \ell).$$

The composition of two endomorphisms $\alpha, \beta \in \mathrm{Hom}(V) \cong T^1_1(V)$ is a contraction of their tensor product:

$$\alpha \circ \beta = \mathrm{tr}^2_1(\alpha \otimes \beta).$$

Proof. Both statements follow from the coordinate descriptions of a linear functional evaluation and composition of endomorphisms:

$$\ell(v) = \ell^i v_i, \quad (\alpha \circ \beta)^i_k = \alpha^i_j \beta^j_k.$$

□

Lemma 1.53. *The space of $(1,1)$-tensors is canonically self-dual. The non-degenerate pairing is given by*

$$\langle \alpha, \beta \rangle = \mathrm{tr}^{12}_{21}(\alpha \otimes \beta) = \mathrm{tr}(\alpha \circ \beta) = \mathrm{tr}(\beta \circ \alpha).$$

Proof. The pairing is bilinear and independent of the choice of a basis. It is non-degenerate: if $\alpha^i_j \neq 0$ for some basis and for some i, j, then put $\beta^i_j = \alpha^i_j$ and $\beta^k_l = 0$ for $(k, l) \neq (i, j)$. This results in $\langle \alpha, \beta \rangle = (\alpha^i_j)^2 \neq 0$. □

Exercise. For $\alpha, \beta, \gamma \in \mathrm{Hom}(V, V)$ write $\mathrm{tr}(\alpha \circ \beta \circ \gamma)$ in terms of $\otimes$ and tr. Show that $\mathrm{tr}(\alpha \circ \beta \circ \gamma) = \mathrm{tr}(\beta \circ \gamma \circ \alpha)$ but in general $\mathrm{tr}(\alpha \circ \beta \circ \gamma) \neq \mathrm{tr}(\alpha \circ \gamma \circ \beta)$.

Lemma 1.54. *For every r, s there is a canonical isomorphism $(T^r_s)^* \cong T^s_r$.*

Exercise. Prove this lemma. What is the image of the tensor (1.12) (with the basis (η^i) dual to the basis (e_i)) under your isomorphism?

1.8.5 Symmetric and antisymmetric tensors

We say that an (r, s)-tensor is *symmetric in the p-th and q-th vector arguments*, if its value remains unchanged when we interchange these variables. Similarly, an *antisymmetric in the p-th and q-th vector arguments* tensor changes its sign, when these variables are interchanged.

Lemma 1.55. *The following three statements are equivalent.*

1. *A tensor α is symmetric (resp. antisymmetric) in the p-th and q-th contravariant arguments.*
2. *The components of α with respect to some basis are unchanged (resp. change the sign) when the p-th and the q-th lower indices are interchanged.*
3. *The same as above but with respect to* all *bases.*

Proof. Third statement implies the second, and the first implies the third (the components with respect to a basis are values on the basis elements). Let us show that the second statement implies the first. We provide an argument for a $(0, 2)$-tensor, the general case does not differ much.

So assume that in some basis $\alpha_{ij} = \alpha_{ji}$ for all i, j. Then we have

$$\alpha(v, w) = \alpha_{ij} v^i w^j = \alpha_{ji} v^i w^j = \alpha(w, v),$$

and similarly in the case of antisymmetry. □

Exercise. Show that every $(0, 3)$-tensor that is symmetric in one pair of arguments and antisymmetric in another pair is identically zero.

1.8.6 Tensor fields

Just as in the case of vector fields and differential forms, the *bundle of (r, s)-tensors over M* is the union of corresponding tensor spaces over the tangent spaces to M:

$$T^r_s(TM) = \bigcup_{p \in M} T^r_s(T_pM).$$

An *(r, s)-tensor field* is a choice of an (r, s)-tensor over every tangent space, smoothly depending on the point.

Equivalently, this is a $C^\infty(M)$-multilinear map that has as arguments r differential 1-forms and s vector fields and takes values in $C^\infty(M)$:

$$\alpha(\omega^1, \dots, \omega^r, X_1, \dots, X_s) \in C^\infty(M).$$

Equivalently, in local coordinates an (r, s)-tensor field is given by its components. Every component is a smooth function:

$$\alpha^{i_1 \dots i_r}_{j_1 \dots j_s}(x), \quad x \in \varphi(U) \subset \mathbb{R}^n.$$

The transformation rule with respect to a coordinate change $y = y(x)$ is

$$\alpha'^{k_1 \ldots k_r}_{l_1 \ldots l_s}(y) = \alpha^{i_1 \ldots i_r}_{j_1 \ldots j_s}(x) \frac{\partial y^{k_1}}{\partial x^{i_1}} \cdots \frac{\partial y^{k_r}}{\partial x^{i_r}} \frac{\partial x^{j_1}}{\partial y^{l_1}} \cdots \frac{\partial x^{j_s}}{\partial y^{l_s}}.$$

(In a denominator of a fraction, an upper index becomes a lower index.) This follows from the fact that the basis change is done with the help of the Jacobi matrix $\frac{\partial x}{\partial y}$:

$$\frac{\partial}{\partial y^i} = \frac{\partial x^j}{\partial y^i} \frac{\partial}{\partial x^j}.$$

1.8.7 The interior product and the divergence

We can perform over the tensor fields the same operations as over the tensors: addition, multiplication with a function (which corresponds to multiplication with a scalar in every tangent space), tensor product, contraction.

Definition 1.56. The *interior product* of a vector and a differential k-form is a differential $(k-1)$-form defined as

$$i_X \omega = \mathrm{tr}^1_1(X \otimes \omega).$$

That is,

$$(i_X \omega)(X_1, \ldots, X_{k-1}) = \omega(X, X_1, \ldots, X_{k-1}).$$

Exercise. Show that the interior product satisfies the Leibniz rule

$$i_X(\omega \wedge \eta) = (i_X \omega) \wedge \eta + (-1)^{\deg \omega} \omega \wedge (i_X \eta).$$

Note that the interior product is a purely algebraic operation. By this we mean that it is fiberwise: to compute $(i_X \omega)(p)$, we only need to know $X(p)$ and $\omega(p)$.

Definition 1.57. Let a volume form $\omega \in \Omega^n(M)$ be fixed, and let X be a vector field on M. The *divergence of X with respect to ω* is a function $\operatorname{div} X \in C^\infty(M)$ defined by

$$d(i_X \omega) = (\operatorname{div} X) \omega.$$

Exercise. Check that for $\omega = dx^1 \wedge \cdots \wedge dx^n \in \Omega^n(\mathbb{R}^n)$ this definition yields $\operatorname{div} X = \sum_{i=1}^n \frac{\partial X^i}{\partial x^i}$.

1.9 Smooth manifolds: where to find them

1.9.1 Product and quotient

Lemma-Definition 1.58. *The product of two smooth manifolds without boundary is a smooth manifold without boundary.*

Let $(U_\alpha, \varphi_\alpha)$ and (V_β, ψ_β) be smooth atlases on M and N. The atlas on $M \times N$ is given by $(U_\alpha \times V_\beta, \varphi_\alpha \times \psi_\beta)$, where

$$\varphi_\alpha \times \psi_\beta \colon U_\alpha \times V_\beta \to \mathbb{R}^{m+n}, \quad (\varphi_\alpha \times \psi_\beta)(p, q) = (\varphi_\alpha(p), \psi_\beta(q)).$$

The product of a manifold without boundary and a manifold with non-empty boundary is a manifold with boundary. The product of two smooth manifolds with boundary is a "smooth manifold with corners": the images of charts around boundary points may be open subsets of "quarter-spaces" rather than of half-spaces. However, there is a canonical way to straighten the corners.

Example 1.59. The 2-torus is the cartesian product of two circles. A circle has a smooth structure (can take two charts). This defines a smooth structure on the torus (get four charts).

The tangent space of a product is the direct sum of the tangent spaces of the factors:

$$T_{(p,q)}(M \times N) = T_pM \oplus T_qN.$$

Exercise. Show that the product of two orientable manifolds is orientable.

Definition 1.60. The standard n-sphere is defined as

$$\mathbb{S}^n = \{x \in \mathbb{R}^{n+1} \mid (x^0)^2 + (x^1)^2 + \cdots + (x^n)^2 = 1\}.$$

A smooth structure on the sphere can be defined by means of two stereographic projections: one from the "north pole", the other from the "south pole". Let us introduce a different atlas that consists of $2n + 2$ charts

$$(U_i^\pm, \varphi_i^\pm), i = 0, \dots, n,$$

where

$$\begin{aligned} U_i^+ &= \{x \in \mathbb{R}^{n+1} \mid x^i > 0\} \\ U_i^- &= \{x \in \mathbb{R}^{n+1} \mid x^i < 0\} \\ \varphi_i^\pm(x) &= (x^0, \dots \hat{\imath} \dots x^n) \end{aligned}$$

(Here $\hat{\imath}$ means that the entry with index i is omitted.) The second atlas defines the same smooth structure: its charts are compatible with the stereographic projections.

Definition 1.61. The n-dimensional projective space $\mathbb{R}P^n$ is the quotient $\mathbb{S}^n/\sim$, where $x \sim y \Leftrightarrow x = -y$.

The projective space is a smooth manifold. Take $n + 1$ charts (U_i, φ_i), $i = 0, \dots, n$, where

$$U_i = \{\{x, -x\} \in \mathbb{R}P^n \mid x^i \neq 0\}, \quad \varphi_i(\{x, -x\}) = \operatorname{sign}(x^i)(x^0, \dots \hat{\imath} \dots, x^n).$$

(Here $\operatorname{sign}(x) = \pm 1$, according to whether $x > 0$ or $x < 0$.)

Exercise. Show that $\mathbb{R}P^n$ is orientable if and only if n is odd.

Exercise. Convince yourself that $\mathbb{R}P^2$ without a point is homeomorphic to the (open) Möbius band.

1.9.2 Immersions and embeddings

Definition 1.62. A smooth map $F\colon M \to N$ is called an *immersion*, if for every $p \in M$ the linear map

$$dF\colon T_pM \to T_{F(p)}N$$

is injective.

If an immersion $F\colon M \to N$ maps M homeomorphically onto its image $F(M)$, then F is called a *smooth embedding*.

- If there is an immersion $M \to N$, then $\dim M \le \dim N$.
- A smooth curve $\gamma\colon I \to \mathbb{R}^n$ is an immersion if and only if $\left\| \frac{d\gamma}{dt} \right\| \neq 0$ for all t.
- If an immersion is injective, it does not yet mean that it is an embedding. Counterexample: $\gamma\colon [0, 2\pi) \to \mathbb{R}^2$, $\gamma(t) = (\cos t, \sin t)$.

Example 1.63. Identify $\mathbb{R}^2$ with $\mathbb{C}$. Then the map $\mathbb{R}^2 \setminus \{0\} \to \mathbb{R}^2 \setminus \{0\}$ that sends z to z^2 is an immersion (but not an embedding).

Example 1.64. The map $\Delta\colon M \to M \times M$, $\Delta(p) = (p, p)$ is an embedding.

Example 1.65. An embedding of the 2-torus in $\mathbb{R}^4 = \mathbb{R}^2 \times \mathbb{R}^2$:

$$F(\varphi, \psi) = (\cos\varphi, \sin\varphi, \cos\psi, \sin\psi).$$

Here $\varphi, \psi \in \mathbb{R}/2\pi$ are points on a circle.

Theorem 1.66. *Every compact smooth manifold can be embedded in a Euclidean space.*

Idea of proof. A coordinate chart φ_α embeds U_α into $\mathbb{R}^n$. Use these coordinates and a partition of unity to construct an embedding of M into $\mathbb{R}^{An+n}$, where A is the number of charts in some finite atlas on M. $\square$

Whitney showed that any smooth n-dimensional manifold (compact or not) can be embedded in $\mathbb{R}^{2n}$ and immersed in $\mathbb{R}^{2n-1}$. For example, the projective plane can be immersed in $\mathbb{R}^3$ although it is not so easy to describe such an immersion.

1.9.3 Local immersion theorem

The most elementary example of an immersion is

$$\mathbb{R}^m \to \mathbb{R}^n, \quad (x^1, \dots, x^m) \mapsto (x^1, \dots, x^m, 0, \dots, 0). \tag{1.13}$$

It turns out that every immersion has locally this form after an appropriate coordinate change.

Theorem 1.67. *Let $F\colon M \to N$ be an immersion,* $\dim M = m$*,* $\dim N = n$*. Then for every point $p \in M$ there are charts (U, φ) around p and (V, ψ) around $F(p)$ such that the map $\psi \circ F \circ \varphi^{-1}$ is given by the formula* (1.13)*.*

Definition 1.68. A subset $M \subset N$ of a smooth manifold is called a *smooth submanifold* of N if there is a smooth atlas on M such that the inclusion $M \to N$ is a smooth embedding.

An equivalent definition: $M \subset N$ is a submanifold if for every point $p \in M$ there is a local chart (V, ψ) on N around p such that $\psi^{-1}(\mathbb{R}^m) = V \cap M$ and ψ sends $V \cap M$ homeomorphically onto its image in $\mathbb{R}^m \subset \mathbb{R}^n$.

Example 1.69. The sphere $\mathbb{S}^n$ is a submanifold of $\mathbb{R}^{n+1}$. Construct appropriate charts $(V_i^\pm, \psi_i^\pm)$ based on the charts $(U_i^\pm, \varphi_i^\pm)$ from Section 1.9.1.

1.9.4 Submersions

Definition 1.70. A smooth map $F\colon M \to N$ is called *regular* at $p \in M$ if the linear map $dF\colon T_pM \to T_{F(p)}N$ is surjective.
If F is regular at all points of M, then it is called a *submersion.*

Sometimes "regular at p" is also called "submersion at p".

- If there is a map $F\colon M \to N$ that is regular at some point $p \in M$, then $\dim M \geq \dim N$.
- A smooth function $f\colon M \to \mathbb{R}$ is regular at p if and only if $df_p \neq 0$.

Example 1.71.
- The projection $M \times N \to N$ is a submersion.
- A smooth function $f\colon M \to \mathbb{R}$ is regular at p if and only if $df\big|_p \neq 0$.
- The map $f\colon \mathbb{R}^n \to \mathbb{R}$, $f(x) = \|x\|^2$ is regular at every point except 0.

Definition 1.72. Let $F\colon M \to N$ be a smooth map. A point $q \in N$ is called a *regular value* of F if every point in $F^{-1}(q)$ is regular.

Theorem 1.73. *The preimage of a regular value is a submanifold.*

1.10 Lie derivative

1.10.1 Integral curves of a vector field

Definition 1.74. Let X be a vector field on a smooth manifold M. A curve $\gamma\colon (a,b) \to M$ is called an *integral curve* of X if

$$\dot\gamma(t) = X(\gamma(t)) \quad \text{for all } t \in (a,b).$$

Here $\dot\gamma(t) = d\gamma\big|_t(\partial_t)$ is the tangent vector to γ at time t. The interval endpoints may be $\pm\infty$.

Example 1.75. If $X(p) = 0$ at some point p (such a point is called a *stationary point*), then $\gamma(t) = p$ is an integral curve of X.

In local coordinates $\varphi(p) = (x^1(p), \ldots, x^n(p))$, finding an integral curve is equivalent to solving a system of first-order ODE on functions $x^1(t), \ldots, x^n(t)$:

$$\frac{dx^i}{dt} = v^i(x^1(t), \ldots, x^n(t)).$$

Here $v^i(x)$ are the components of the vector field X with respect to the coordinate system φ.

Example 1.76. Let us find the integral curves of the radial vector field $X = r\partial_r$ on the plane. In the polar coordinates we have the system

$$\frac{dr}{dt} = r, \quad \frac{d\varphi}{dt} = 0.$$

The general solution is $r(t) = r(0)e^t$, $\varphi(t) = \varphi(0)$. Every integral curve is defined on $(-\infty, \infty)$.

If we take a vector field $X = r^2\partial_r$, then the system becomes

$$\frac{dr}{dt} = r^2, \quad \frac{d\varphi}{dt} = 0,$$

and the general solution is

$$r(t) = \frac{r(0)}{1 - r(0)t}, \quad \varphi(t) = \varphi(0).$$

Integral curves are defined only on the intervals $(-\infty, \frac{1}{r(0)})$. One says: the solution escapes to infinity in a finite time.

It is possible to construct examples of vector fields where the integral curves are defined on arbitrarily small time intervals.

Theorem 1.77. *For every $p \in M$ there is $\epsilon > 0$ such that an integral curve $\gamma\colon (-\epsilon, \epsilon) \to M$ with $\gamma(0) = p$ exists. Besides, this integral curve is unique.*

Proof. Via a coordinate representation of the vector field (see above) this is the existence and uniqueness theorem for solutions of ODE. □

Here are some strengthenings of this theorem.

Theorem 1.78. *If the manifold M is compact, then every integral curve can be extended to all of $\mathbb{R}$.*

Roughly speaking, on a compact manifold there is no place to escape.

Theorem 1.79. *For every $p \in M$ there is an open neighborhood $U \ni p$ and an $\epsilon > 0$ such that for all $q \in U$ there is an integral curve $\gamma\colon (-\epsilon, \epsilon) \to M$ with $\gamma(0) = q$.*

Roughly speaking, near p the "speed" is bounded, so starting from any point you will need a long time to escape.

1.10.2 The flow of a vector field

Let X be a vector field on M. For simplicity, assume that for all $p \in M$ the integral curve $\gamma_p(t)$ with the initial condition $\gamma_p(0) = p$ is defined for all $t \in \mathbb{R}$. (This holds for example if M is compact, see Theorem 1.78.) Then for every $t \in \mathbb{R}$ there is a map

$$\varphi_t\colon M \to M, \quad \varphi_t(p) = \gamma_p(t). \tag{1.14}$$

Definition 1.80. The family of maps (1.14) is called the *flow* of the vector field X.

Lemma 1.81. *The flow of a vector field has the following properties.*

1. $\varphi_0 = \mathrm{id}$
2. $\varphi_t \circ \varphi_s = \varphi_{t+s}$
3. $(\varphi_t)^{-1} = \varphi_{-t}$
4. *Every map φ_t is a diffeomorphism $M \to M$.*

Proof. The first three properties follow easily from the definition. The last one follows from the C^∞-dependence of a solution of an ODE on the initial conditions. □

In other words, the flow of a vector field is a 1-parameter group of diffeomorphisms of M.

Example 1.82.

- For the radial vector field $X = r\partial_r$ on $\mathbb{R}^2$ (where we put $X(0) = 0$ which results in a smooth field on $\mathbb{R}^2$) we have
$$\varphi_t(p) = e^t \cdot p.$$
Thus, the flow consists of homotheties with the center at 0. The homothety coefficient is exponential in t, which is forced by the group property of the flow.

- Let $X = \partial_\varphi$ and $X(0) = 0$. (Again, this is a smooth vector field, because in the cartesian coordinates it is written as $X(x, y) = (-y, x)$.) Then $\varphi_t \colon \mathbb{R}^2 \to \mathbb{R}^2$ is the rotation around 0 by the angle t.

What can we do if the integral curves of X are not defined for all t? For every $t \in \mathbb{R}$ denote by $M_t \subset M$ the set of all points p such that the curve γ_p is defined at time t (by which we mean that it is defined on some open interval containing 0 and t). This yields smooth maps $\varphi_t \colon M_t \to M$, called the *local flow* of X. These maps satisfy the properties formulated in Lemma 1.81, with obvious modifications, which allows to call them a local one-parameter group.

An important observation is that for every $p \in M$ there is an open neighborhood $U \ni p$ and an $\epsilon > 0$ such that $U \subset M_t$ for all $t \in (-\epsilon, \epsilon)$. That is the maps $\varphi_t \colon U \to M$ are defined for all t sufficiently close to 0. It follows that for every p there is an ϵ such that the differentials

$$d\varphi_t \colon T_pM \to T_{\varphi_t(p)}M \tag{1.15}$$

are defined for all $t \in (-\epsilon, \epsilon)$.

1.10.3 Definition of the Lie derivative

The (local) flow of a vector field X defines the linear maps (1.15) between the tangent spaces at two points on the same integral curve. This allows to compare tangent vectors (and more generally, tensors) living at different points of the manifold M.

Definition 1.83. Let X and Y be vector fields on M. The *Lie derivative* of Y with respect to X is the time derivative of X pulled back along the flow of Y:

$$L_XY = \left.\frac{d}{dt}\right|_{t=0} d\varphi_{-t}(Y).$$

To work with the above formula, we will need to write it in an expanded form. To compute the vector $(L_XY)(p)$, we are sending the vector $Y(\varphi_t(p)) \in T_{\varphi_t(p)}M$ to T_pM and comparing it there with $Y(p)$. This means

$$(L_XY)(p) = \lim_{t\to 0} \frac{d\varphi_{-t}(Y(\varphi_t(p))) - Y(p)}{t}$$

Example 1.84. Let $X = \partial_\varphi$, $Y = \partial_x$. The map φ_{-t} is rotation by $-t$ around 0. It sends the tangent plane at $\varphi_t(p)$ to the tangent plane at p by rotating it by $-t$ as well. Therefore $d\varphi_{-t}(Y)$ is Y rotated by $-t$. It follows that $L_X Y = -\partial_y$.

From the definition it is clear that $L_X(Y + Z) = L_X Y + L_X Z$.

Exercise. Show that $L_X(fY) = X(f) \cdot Y + f \cdot L_X Y$.

1.10.4 The Lie derivative and the commutator

It is not easy to figure out what is $L_{X+Y} Z$ and what is $L_{fX} Y$. The former uses the flow of $X+Y$ that does not stand in a simple relation with the flows of X and Y. The latter uses the flow of fX. Although the integral curves of fX are just reparametrized integral curves of X, the reparametrization depends on the starting point, which makes it difficult to work with it.

Theorem 1.85. *The Lie derivative of a vector field coincides with the commutator:* $L_X Y = [X, Y]$.

Proof. A vector field is completely determined by its action on the functions. Therefore it suffices to show that $(L_X Y)(f) = X(Y(f)) - Y(X(f))$ for every function f. Note that

$$d\varphi_{-t}(Y)(f) = Y(f \circ \varphi_{-t})$$

because of the definition of the directional derivative and the chain rule. Thus we have

$$(L_X Y)(f)(p) = \lim_{t \to 0} \frac{Y(f \circ \varphi_{-t})(p_t) - Y(f)(p)}{t}.$$

Here we put $p_t = \varphi_t(p)$. One must be careful and not forget at what points the individual derivatives in the above formula are evaluated.

Consider the function

$$g \colon (-\epsilon, \epsilon) \times U \to \mathbb{R}, \quad g_t = \begin{cases} \frac{f \circ \varphi_t - f}{t} & \text{for } t \neq 0 \\ X(f) & \text{for } t = 0 \end{cases}$$

We claim (without a proof) that g is smooth. (The continuity is easy: to see that $\lim_{t \to 0} g_t(p) = g_0(p)$ remember how the directional derivative $X(f)$ is defined if X is interpreted as an equivalence class of paths. The smoothness is not hard either, see [Spi79, Lemma 5.9].) Thus we have

$$f \circ \varphi_{-t} = f - t g_{-t}.$$

Hence

$$(L_XY)(f)(p) = \lim_{t\to 0} \frac{Y(f - tg_{-t})(p_t) - Y(f)(p)}{t}$$
$$= \lim_{t\to 0} \frac{Y(f)(p_t) - Y(f)(p)}{t} - \lim_{t\to 0} Y(g_{-t})(p_t).$$

By the path definition of the directional derivative we have

$$\lim_{t\to 0} \frac{Y(f)(p_t) - Y(f)(p)}{t} = X(Y(f))(p)$$

($Y(f)$ is a function, and we look how it changes along the path p_t whose tangent vector at $t = 0$ is X). And due to the smoothness of g_t we have

$$\lim_{t\to 0} Y(g_{-t})(p_t) = Y(g_0)(p) = Y(X(f))(p).$$

The theorem is proved. □

Corollary 1.86. *We have*

$$L_{X+Y}Z = L_XZ + L_YZ$$
$$L_{fX}Y = f \cdot L_XY - Y(f) \cdot X$$
$$L_XX = 0$$
$$L_{[X,Y]}Z = L_X(L_YZ) - L_Y(L_XZ)$$

Remark. There is an equivalent definition of the Lie derivative:

$$L_XY = -\left.\frac{d}{dt}\right|_{t=0} d\varphi_t(Y)$$

Explicitely,

$$(L_XY)(p) = \lim_{t\to 0} \frac{Y(p) - d\varphi_t(Y(\varphi_{-t}(p)))}{t}$$

(This can be interpreted as watching the sticks floating on the river.)

Exercise.

1. Give an example of two vector fields X and Y such that $X(p) = 0$ but $(L_XY)(p) \neq 0$.
2. Show that if $X(p) = 0$ and $Y(p) = 0$, then $(L_XY)(p) = 0$.

Exercise. Show that $[X, Y] = 0$ everywhere on M if and only if the flows of X and Y commute: $\varphi_t \circ \psi_s = \psi_s \circ \varphi_t$ for all t, s.

1.11 Covariant derivative

1.11.1 Motivation

On a smooth manifold, we have the notion of a directional derivative of a function. Recall that, given $f \in C^\infty(M)$ and $X \in T_pM$, the derivative $X(f)(p)$ of f in the direction X can be defined in two equivalent ways:

- as the directional derivative in $\mathbb{R}^n$, by using any coordinate system;
- as
$$X(f)(p) = \left.\frac{d}{dt}\right|_{t=0} f(\gamma(t)),$$
where γ is any path such that $\gamma(0) = p$ and $\left.\frac{d}{dt}\right|_{t=0} \gamma(t) = X$.

If X is not just one tangent vector, but a vector field, then $X(f)$ becomes a function on M. It is important to remember that the value $X(f)(p)$ depends on the vector $X(p)$ only, and not on the values of the vector field X at the nearby points. By contrast, $X(f)(p)$ depends on the values of f around p, and not only on $f(p)$.

We would like to be able to differentiate vector fields in a similar way: given a vector field Y and a vector $X \in T_pM$, compute the "directional derivative" of Y at the point p in the direction X. If we try to imitate the coordinate definition of $X(f)$, then we will obtain an expression that is not coordinate-independent. If we try to imitate the path definition, then we write
$$\nabla_X Y(p) \stackrel{?}{=} \left.\frac{d}{dt}\right|_{t=0} Y(\gamma(t)) = \lim_{t\to 0} \frac{Y(\gamma(t)) - Y(\gamma(0))}{t},$$
which makes no sense because we cannot compare two vectors $Y(\gamma(t))$ and $Y(\gamma(0))$ that belong to the tangent spaces at different points.

Note the difference with the Lie derivative: there we used the vector field X to construct a flow, and the flow transports tangent vectors along the integral curves. Now X is only a tangent vector, and not a vector field. We have no integral curves and no flow.

1.11.2 Definition, existence, and non-uniqueness of a covariant derivative

As we cannot give a constructive definition at the moment, let's formulate a set of axioms for this operation that we want to define.

Definition 1.87 (a)**.** A *covariant derivative* of vector fields is an operation that associates to two vector fields X and Y a new vector field $\nabla_X Y$. The covariant derivative is

- $C^\infty(M)$-linear in X:

$$\nabla_{X_1+X_2} Y = \nabla_{X_1} Y + \nabla_{X_2} Y, \quad \nabla_{fX} Y = f \nabla_X Y;$$

  additive and satisfies the Leibniz rule in Y:

$$\nabla_X(Y_1 + Y_2) = \nabla_X Y_1 + \nabla_X Y_2, \quad \nabla_X(fY) = X(f) \cdot Y + f \cdot \nabla_X Y.$$

Definition 1.87 (b)**.** A *covariant derivative* of vector fields is an operation that associates to every vector field Y a $(1,1)$-tensor field ∇Y. The map $Y \mapsto \nabla Y$ must satisfy the Leibniz rule with respect to multiplication by functions:

$$\nabla(fY) = df \cdot Y + f \cdot \nabla Y.$$

Instead of $(\nabla Y)(X)$ we write $\nabla_X Y$.

A covariant derivative is also called an *affine connection.*

The equivalence of these definitions is shown similarly to the equivalence of definitions 1.16(a) and 1.16(b) of differential 1-forms. The main point is to show that the C^∞-linearity with respect to X implies that $\nabla_X Y(p)$ depends on the value of $X(p)$ only, disregarding the values of X near p.

Example 1.88. The following defines a covariant derivative on $\mathbb{R}^n$. For $X = (X^1, \ldots, X^n)$ and $Y = (Y^1, \ldots, Y^n)$ put

$$D_X Y = \left(\sum_{j=1}^n X^j \frac{\partial Y^i}{\partial x^j} \right)_{i=1}^n .$$

That is, this covariant derivative consists of directional derivatives of components of the vector field Y.

In the sense of the second definition, DY is a linear endomorphism with the matrix $\left(\frac{\partial Y^i}{\partial x^j}\right)$.

An axiomatic definition does not say anything about the existence and uniqueness of an object. The following theorem describes the extent of the non-uniqueness of the covariant derivative.

Theorem 1.89. *For any two covariant derivatives ∇, ∇' on a manifold M (or on an open subset of M) their difference*

$$\nabla' - \nabla \colon (X, Y) \mapsto \nabla'_X Y - \nabla_X Y$$

is a $(1,2)$-tensor. Conversely, for every covariant derivative ∇ and $(1,2)$-tensor A the sum $\nabla + A$ is a covariant derivative.

Proof. The $C^\infty(M)$-linearity of $(X,Y) \mapsto \nabla'_X Y - \nabla_X Y$ with respect to X and the additivity with respect to Y follow from the corresponding properties of ∇' and ∇. Let us prove the $C^\infty(M)$-linearity with respect to Y.

$$\nabla'_X(fY) - \nabla_X(fY) =$$
$$= f \cdot (\nabla'_X Y - \nabla_X Y)$$

It follows that $\nabla' - \nabla$ is a bilinear vector-valued form, that is a $(1,2)$-tensor field.

In the same way one shows that adding to a covariant derivative a $(1,2)$-tensor field results in another covariant derivative. □

Corollary 1.90. *Every connection on an open subset $U \subset \mathbb{R}^n$ has the form*

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

where $\overline{\nabla}$ is the standard connection on $\mathbb{R}^n$, and Γ is an arbitrary $(1,2)$-tensor field.

In the standard coordinates on $\mathbb{R}^n$ the above formula writes as

$$\nabla_X Y = \left(X^j \frac{\partial Y^i}{\partial x^j} + \Gamma^i_{jk} X^j Y^k \right) e_i. \tag{1.16}$$

As a result, every connection on a chart domain $U \subset M$ has in the local coordinates the form (1.16). The numbers Γ^i_{jk} are called the *Christoffel symbols* of the connection ∇ with respect to a given chart.

Remark. Although the difference of two connections is a tensor, the Christoffel symbols don't form a tensor. This is because the "standard" connection D changes when we change the coordinates. To give a concrete example, consider the standard connection on $\mathbb{R}^2$. With respect to the cartesian coordinates it has zero Christoffel symbols. With respect to the polar coordinates it has non-zero Christoffel symbols (because the standard connection with respect to (r,φ) differs from the standard connection with respect to (x,y)). But there is no tensor that has zero components in one coordinate system and non-zero components in another.

Theorem 1.91. *On every manifold there is a covariant derivative.*

Idea of proof. Take an atlas and consider the standart covariant derivatives on each of the chart domains. Then use a partition of unity to glue these covariant derivatives together. This works because a convex combination of covariant derivatives is a covariant derivative. □

Exercise. Let $\nabla^1, \ldots, \nabla^n$ be covariant derivatives. Show that their linear combination $\sum_{i=1}^n \lambda_i \nabla^i$ is a covariant derivative if $\sum_{i=1}^n \lambda_i = 1$.

1.11.3 Torsion tensor

Lemma-Definition 1.92. *The* torsion tensor *of a connection* ∇ *is a* $(1,2)$*-tensor field defined by*

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

A connection is called torsion-free *if its torsion tensor is identically zero.*

We need to prove that the above formula defines a tensor field.

Proof. Let us show that $T(X,Y)$ is $C^\infty(M)$-linear in both arguments. It is obviously additive in each of the arguments. It remains to check the homogenuity.

$$T(fX,Y) =$$

$$= fT(X,Y)$$

For the homogenuity in Y either proceed similarly, or note that $T(X,Y) = -T(Y,X)$, so that

$$T(X,fY) = -T(fY,X) = -fT(Y,X) = fT(X,Y).$$

□

The result is quite surprising: although the individual summands in the definition of $T(X,Y)(p)$ depend on a local behavior of either X or Y or both, the combination depends only on the values $X(p)$ and $Y(p)$.

Since $T(X,Y) = -T(Y,X)$, the tensor T may be called a vector-valued differential 2-form.

1.12 Riemannian metrics

1.12.1 Definition and main properties

Definition 1.93. A *Riemannian metric* g on a smooth manifold M is a field of positive definite symmetric bilinear forms.

That is, at every point p we have a scalar product $g(X,Y)$ for $X,Y \in T_pM$, depending smoothly on p.

We use the notation

$$\langle X,Y\rangle_g = g(X,Y), \quad \|X\|_g^2 = g(X,X).$$

The Riemannian metric allows to define the angle between two tangent vectors at the same point:

$$\cos\angle(X,Y) = \frac{g(X,Y)}{\sqrt{g(X,X)g(Y,Y)}},$$

and allows to speak of orthonormal bases in T_pM.

Theorem 1.94. *On every smooth manifold M there exist a Riemannian metric.*

Proof. Partition of unity. □

1.12.2 Riemannian metric in local coordinates

As every $(0,2)$-tensor field, a Riemannian metric has in every coordinate system the form

$$g(X,Y) = g_{ij}X^iY^j.$$

The symmetry of g is equivalent to $g_{ij} = g_{ji}$, and the positive definiteness to the positive definiteness of the matrix

$$G = (g_{ij})_{i,j=1}^n.$$

Note that $g_{ij} = g\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right)$.

When we perform a coordinate change $y = y(x)$, then in the new coordinates y the components of the metric tensor are

$$g'_{kl} = g_{ij}\frac{\partial x^i}{\partial y^k}\frac{\partial x^j}{\partial y^l}. \tag{1.17}$$

This follows from

$$g = g_{ij}dx^i \otimes dx^j \tag{1.18}$$

and

$$dx^i = \frac{\partial x^i}{\partial y^k}dy^k.$$

The rule (1.17) corresponds to the rule of change of a matrix of a symmetric bilinear form on a vector space:

$$G' = A^\top GA.$$

In our case, the matrix A is the Jacobi matrix of the coordinate change; usually it depends on the point $p \in M$.

Following a long tradition, in (1.18) one usually omits the sign of the tensor product. For example, the standard Euclidean metric on $\mathbb{R}^n$ is written as

$$g = (dx^1)^2 + \cdots + (dx^n)^2.$$

Example 1.95. Let's write the Euclidean metric in the polar coordinates. We have

$$G = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}.$$

This implies

$$G' = A^\top A = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}.$$

The same result can be obtained by computing the scalar products of the coordinate basis vectors:

$$\|\partial_r\|^2 = 1, \quad \langle \partial_r, \partial_\varphi \rangle = 0, \quad \|\partial_\varphi\|^2 = r^2.$$

1.12.3 Looking for good coordinates

Since the components of the metric tensor change when we are changing the coordinates, it is natural to try to find a coordinate system, in which the components take the simplest possible form. The simplest of all is of course the unit matrix.

Question. Given a Riemannian manifold (M, g) and a point $p \in M$. Do there exist coordinates in some neighborhood U of p such that in these coordinates $g = \sum_{i=1}^n (dx^i)^2$ everywhere in U?

Answer. In general not. Every Riemannian metric has *curvature* (that is independent of the coordinate system). The curvature of the Euclidean metric is 0. Therefore if the curvature of g is not 0, then g cannot be transformed to a Euclidean metric by a coordinate change.

Lemma 1.96. *For every Riemannian manifold (M, g) and every point $p \in M$ there are coordinates around p for which $g = \sum_{i=1}^n (dx^i)^2$ at p (but not at points different from p).*

Proof. Take any coordinate chart around p. Let $G(p)$ be the matrix of g at p. Since $G(p)$ is positive definite, there is $A \in GL(n)$ such that $G(p) = A^\top A$. Perform the coordinate change $y = Ax$. This is a linear substitution, it has the Jacobi matrix $\left(\frac{\partial y}{\partial x}\right) = A$ everywhere. The Jacobi matrix of $x = x(y)$ is

$$\left(\frac{\partial x}{\partial y}\right) = A^{-1}.$$

Therefore by (1.17) the matrix of g at the point p in the coordinates y has the form

$$G'(p) = (A^{-1})^\top G A^{-1} = (A^\top)^{-1} A^\top A A^{-1} = \mathrm{Id}\,.$$

□

If we try to bring G to the standard form at all points, we might choose a smooth family of matrices $A(q)$ such that $G(q) = A(q)^\top A(q)$ for all points q in a neighborhood of p. But then these matrices must be the Jacobi matrices of some coordinate change:

$$A(q) = \left(\frac{\partial y}{\partial x}\right).$$

The commutation of the second derivatives $\frac{(\partial y^i)^2}{\partial x^j \partial x^k}$ imposes restrictions on the components of the matrices $A(q)$, therefore not every family of matrices corresponds to a coordinate change. The matter is complicated by the non-uniqueness of A satisfying $A^\top A = G$. Therefore it is not easy to decide whether G can be transformed to the identity matrix in a neighborhood of p.

Here is a weaker positive result that we give without a proof.

Theorem 1.97. *Let (M, g) be a Riemannian manifold of dimension 2. Then in the neghborhood of every point p there is an* isotropic coordinate system*, that is one with*

$$g(q) = f(q)(dx^2 + dy^2) \text{ for all } q \in U$$

where f is a smooth function.

1.12.4 Coordinate systems vs. frames

The negative answer to the question in the previous subsection means that for a general Riemannian metric we cannot find a coordinate system with orthonormal coordinate vector fields (even if we are looking for it in an arbitrarily small neighborhood). However, it is possible to find an orthonormal frame.

Definition 1.98. An *orthonormal frame* on an open subset $U \subset M$ of a Riemannian manifold (M, g) is a collection of vector fields $(e_1, \ldots, e_n)$ on M such that for every $p \in M$ the vectors $(e_1(p), \ldots, e_n(p))$ form an orthonormal basis of T_pM.

Lemma 1.99. *Orthonormal frames exist locally.*

Proof. On a domain U of a coordinate chart take the coordinate vector fields $\partial_1, \ldots, \partial_n$. Apply the Gram-Schmidt process to these fields pointwise. This gives a new collection of vector fields, which form an orthonormal frame. □

There is a price for everything. The vector fields $e_1, \ldots, e_n$ are orthonormal, but they are in general not coordinate vector fields. That is, their pairwise commutators might not vanish.

Orthonormal frame is an alternative to a coordinate system: it allows to represent a vector field as a linear combination of basic vector fields. By taking the dual frame for covectors (coframe), we also get a coordinate presentation of a tensor of any type.

Remember: coordinate vector fields commute but are usually not orthonormal; frames can be orthonormal, but usually don't commute.

1.12.5 Lengths of curves

Definition 1.100. Let $\gamma\colon [a,b] \to M$ be a smooth path on a Riemannian manifold (M,g). The length of γ is defined as

$$L_g(\gamma) = \int_a^b \left|\frac{d\gamma}{dt}\right|_g dt = \int_a^b \sqrt{g\left(\frac{d\gamma}{dt}, \frac{d\gamma}{dt}\right)}\, dt.$$

Lemma 1.101. *The length of a path is independent of the parametrization. That is, if $\varphi\colon [c,d] \to [a,b]$ is a diffeomorphism, then $L_g(\gamma \circ \varphi) = L_g(\gamma)$.*

Proof. An easy application of the substitution rule in the integral. □

Exercise. Compute the length of the logarithmic spiral given in the polar coordinates by $r = e^{\varphi}$, $r \le 1$.

1.12.6 The volume form on a Riemannian manifold

Lemma-Definition 1.102. *Let (M,g) be an oriented Riemannian manifold. Then*

$$\sqrt{\det G} \cdot dx^1 \wedge \cdots \wedge dx^n,$$

where $(x^1, \ldots, x^n)$ is any coordinate system compatible with the orientation chosen on M, defines a volume form on M.

Proof. We need to show that the form is preserved by an orientation-preserving coordinate change $y = y(x)$. This is true because

$$G' = A^\top G A \Rightarrow \det G' = (\det A)^2 \det G \Rightarrow \sqrt{\det G'} = \det A \cdot \sqrt{\det G},$$

where $A = \left(\frac{\partial y}{\partial x}\right)$ (its determinant is positive, because the coordinate systems x and y belong to the same orientation), and

$$dy^1 \wedge \cdots \wedge dy^n = \det A \cdot dx^1 \wedge \cdots \wedge dx^n.$$

□

1.12.7 Type change of tensors

Recall that to a bilinear form $\alpha \in T_2^0(V)$ one can associate two homomorphisms $\beta_1, \beta_2 \in \mathrm{Hom}(V, V^*)$:

$$\beta_1(v)(w) := \alpha(v,w), \quad \beta_2(v)(w) := \alpha(w,v).$$

The form α is symmetric if and only if $\beta_1 = \beta_2$. In that case we denote $\beta = \beta_1 = \beta_2$. A bilinear form α is called *non-degenerate* if for every $v \in V$ there is $w \in V$ such that $\alpha(v,w) \neq 0$. We have

$$\alpha \text{ is positive definite} \Rightarrow \alpha \text{ is non-degenerate} \Leftrightarrow \beta \text{ is an isomorphism.}$$

The same holds on for fields of symmetric bilinear forms on a smooth manifold, in particular for a Riemannian metric g. The pointwise isomorphism $TM \to T^*M$ that corresponds to g allows to transform vectors to differential 1-forms and back. These are the so-called musical isomorphisms.

Definition 1.103. Let X be a vector field on M. Denote by $X^\flat$ the 1-form given by

$$X^\flat(Y) = g(X, Y).$$

Let ω be a 1-form on M. Denote by $\omega^\sharp$ the vector field determined by

$$g(\omega^\sharp, X) = \omega(X).$$

In particular, we can define the gradient vector field ∇f of a function f by

$$g(\nabla f, X) = X(f).$$

Let (g_{ij}) be the components of g in a local coordinate system. Then we have

$$X_i^\flat = g_{ij} X^j.$$

The components of $\omega^\sharp$ are expressed through those of ω by means of the components of the inverse isomorphism $T^*M \to TM$:

$$(\omega^\sharp)^i = g^{ij} \omega_j.$$

We then have

$$g_{ij} g^{jk} = \delta_i^k,$$

that is the matrix of g^{-1} with respect to a basis of TM and the dual basis of T^*M is the inverse of the matrix of g in the same bases.

Remark. The inverse $g^{-1} \colon T^*M \to TM$ gives rise to a symmetric bilinear form on T^*M.

The isomorphisms g and g^{-1} can be used to raise or lower an index in any tensor field. In particular, a homomorphism $\alpha \in T_1^1(TM)$ can be turned into a bilinear form $\beta \in T_2^0(TM)$ by putting

$$\beta(v, w) = g(\alpha(v), w).$$

1.13 The Levi-Civita connection

1.13.1 Covariant derivative of tensors

A covariant derivative of vector fields allows to define a covariant derivative of arbitrary tensors. The idea is to impose the Leibniz rule with respect to the contraction and with respect to the tensor product.

Definition 1.104. Let ∇ be a covariant derivative on M. We define the covariant derivative of a 1-form ω by

$$(\nabla_X \omega)(Y) = X(\omega(Y)) - \omega(\nabla_X Y).$$

Equivalently,

$$\nabla_X(\omega(Y)) = (\nabla_X \omega)(Y) + \omega(\nabla_X Y),$$

where the covariant derivative of a function (on the left hand side of the equation) is just the directional derivative.

This operation on 1-forms has the derivation property.

Exercise. Check that $\nabla_X(f\omega) = X(f)\omega + f \cdot \nabla_X \omega$.

For tensors of higher order we define the covariant derivative inductively by

$$\nabla(Y \otimes \alpha) = (\nabla Y) \otimes \alpha + Y \otimes (\nabla \alpha)$$
$$\nabla(\omega \otimes \alpha) = (\nabla \omega) \otimes \alpha + \omega \otimes (\nabla \alpha)$$

and using the additivity. Again, one can show that the operation thus defined has the derivation property.

In particular, if α is a bilinear form, then we have

$$X(\alpha(Y, Z)) = (\nabla_X \alpha)(Y, Z) + \alpha(\nabla_X Y, Z) + \alpha(Y, \nabla_X Z).$$

1.13.2 Metric connections

Definition 1.105. A *metric connection* is a connection ∇ with respect to which the metric tensor is parallel, that is $\nabla g = 0$.

In other words, a metric connection is one that satisfies

$$X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z).$$

Lemma 1.106. *Let ∇ and $\nabla' = \nabla + A$ be two metric connections. Then the $(1,2)$-tensor A satisfies*

$$g(A_X(Y), Z) + g(Y, A_X(Z)) = 0. \tag{1.19}$$

Here we denote $A_X(Y) = \nabla'_X Y - \nabla_X Y$.

Proof. Follows directly from

$$g(\nabla'_X Y, Z) + g(Y, \nabla'_X Z) = X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z).$$

□

1.13.3 Existence and uniqueness of the Levi-Civita connection

Definition 1.107. A *Levi-Civita connection* on a Riemannian manifold (M, g) is a torsion-free metric connection.

(A connection is called torsion-free if its torsion tensor vanishes.)

Theorem 1.108. *For every Riemannian metric there is a unique Levi-Civita connection.*

Proof of the uniqueness. Let ∇', ∇ be two Levi-Civita connections, and let $A = \nabla' - \nabla$ be the corresponding $(1, 2)$-tensor field. Compute the difference of the torsion tensors of ∇' and ∇:

$$T'(X,Y) - T(X,Y) = \\ = A_X(Y) - A_Y(X).$$

Since both ∇' and ∇ are torsion-free, we have

$$A_X(Y) = A_Y(X). \tag{1.20}$$

We want to show that (1.19) together with (1.20) imply $A = 0$. The argument will look nicer if we lower the upper index of A so that it becomes a $(0, 3)$-tensor:

$$A^\flat(X,Y,Z) = g(A_X(Y), Z).$$

Then (1.19) and (1.20) become

$$A^\flat(X,Y,Z) = -A^\flat(X,Z,Y) \quad \text{and} \quad A^\flat(X,Y,Z) = A^\flat(Y,X,Z).$$

It follows that the cyclic permutation of arguments changes the sign of A. By performing the cyclic permutation thrice, we obtain

$$A^\flat(X,Y,Z) = -A^\flat(X,Y,Z).$$

Hence $A = 0$ and the uniqueness of a Levi-Civita connection is proved. □

Proof of the existence. Let ∇ be any connection on M. We want to find a $(1, 2)$-tensor A such that the connection $\nabla' = \nabla + A$ is metric and torsion-free. This is equivalent to the following conditions on $A^\flat$:

$$A^\flat(X,Y,Z) + A^\flat(X,Z,Y) = (\nabla_X g)(Y,Z) \tag{1.21}$$

$$A^\flat(X,Y,Z) - A^\flat(Y,X,Z) = -g(T(X,Y),Z). \tag{1.22}$$

It follows that

$$A^\flat(X,Y,Z) = -A^\flat(Y,Z,X) + (\nabla_Y g)(X,Z) - g(T(X,Y),Z).$$

By repeating the cyclic permutation of arguments two more times we obtain

$$2A^\flat(X,Y,Z) = (\nabla_X g)(Y,Z) + (\nabla_Y g)(Z,X) - (\nabla_Z g)(X,Y) \\ - g(T(Z,X),Y) - g(T(X,Y),Z) + g(T(Y,Z),X). \quad (1.23)$$

A direct check shows that this expression for A satisfies conditions (1.21) and (1.22). Thus $\nabla' = \nabla + A$ is a Levi-Civita connection. □

Lemma 1.109 (Koszul formula)**.** *The Levi-Civita connection is determined by*

$$g(\nabla_X Y, Z) = \frac{1}{2}(X(g(Y,Z)) + Y(g(Z,X)) - Z(g(X,Y)) \\ + g([Z,X],Y) + g([X,Y],Z) - g([Y,Z],X)).$$

Proof. Due to the uniqueness of a Levi-Civita connection, an application of the "symmetrization-antisymmetrization" procedure from the previous proof should yield $A^\flat = 0$. After expanding the right hand side of (1.23) we obtain the Koszul formula. □

1.13.4 Isometric immersions and induced metrics

Definition 1.110. A smooth map $F\colon M \to N$ between Riemannian manifolds (M,g) and (N,h) is called an *isometric immersion* if for every $p \in M$ and every $X, Y \in T_pM$ we have

$$h(dF(X), dF(Y)) = g(X,Y).$$

Every isometric immersion is an immersion: for $X \neq 0$ we have

$$h(dF(X), dF(X)) = g(X,X) > 0,$$

therefore $dF(X) \neq 0$.

If an immersion between smooth manifolds is given and the target manifolds is equipped with a Riemannian metric, this metric can be pulled back to the source manifold.

Definition 1.111. Let $F\colon M \to N$ be an immersion of smooth manifolds, and let h be a Riemannian metric on N. The *induced Riemannian metric* g on M is defined by

$$g(X,Y) = h(dF(X), dF(Y)).$$

In other words, the induced metric is the pullback F^*h. The assumption that f is an immersion and not just any smooth map is needed for g to be positive definite:

$$g(X,X) = h(dF(X), dF(X)) > 0 \text{ for } X \neq 0.$$

Example 1.112. The Riemannian manifold (N, h) may be $\mathbb{R}^n$ with the Euclidean metric. The manifold M may be not just immersed but also embedded; in that case we sometimes consider it as a subset of N.

In particular, we now have a Riemannian metric on $\mathbb{S}^n \subset \mathbb{R}^{n+1}$.

1.13.5 The Levi-Civita connection on an immersed manifold

Let (M, g) and (N, h) be Riemannian manifolds, and $F\colon M \to N$ be an isometric immersion. What is the relation between the Levi-Civita connections on (M, g) and (N, h)? More concretely, let $p \in M$, $X \in T_pM$ and Y be a vector field in a neighborhood of p. Take the covariant derivative $\nabla_X Y$. Then consider the vector $dF(X) \in T_{F(p)}N$ and the vector field $dF(Y)$, and take the covariant derivative $\widetilde{\nabla}_{dF(X)}(dF(Y))$, where $\widetilde{\nabla}$ is the Levi-Civita connection of (N, h). A natural conjecture would be $\widetilde{\nabla}_{dF(X)}(dF(Y)) = dF(\nabla_X Y)$, but we will see that it is not quite true.

Every immersion is locally an embedding. Thus by taking, if needed, a subset of M we can assume that F is an embedding and identify M with its image. (Thus we can forget about F and dF, the question is now about the relation between $\nabla_X Y$ and $\widetilde{\nabla}_X Y$.)

For every $p \in M$ we have two vector spaces $T_pM \subset T_pN$. On T_pN we have a scalar product h, and its restriction to T_pM is g. Denote by $\perp_p M$ the orthogonal complement to T_pM in T_pN:

$$\perp_p M = \{V \in T_pN \mid h(V, X) = 0 \text{ for all } X \in T_pM\}.$$

A vector $V \in \perp_p M$ is called a *normal* to M.

Definition 1.113. The union of the vector spaces $\perp_p M$ is called the *normal bundle* of M in N:

$$\perp M = \bigcup_{p \in M} \perp_p M.$$

Since $T_pN = T_pM \oplus \perp_p M$, every vector $V \in T_pN$ has a unique decomposition into a tangential and a normal component:

$$V = \top V + \perp V, \quad \top V \in T_pM, \perp V \in \perp_p M.$$

Theorem 1.114. *Let M be an immersed submanifold of a Riemannian manifold N, and let X and Y be two vector fields on M. Then*

$$\nabla_X Y = \top(\widetilde{\nabla}_X Y). \tag{1.24}$$

That is, the covariant derivative on a submanifold is the tangential component of the ambient covariant derivative.

Note that Y is a vector field on M. The covariant derivative $\widetilde{\nabla}$ derivates fields defined on N. Strictly speaking, in order to give sense to $\widetilde{\nabla}_X Y$, we need to take an extension $\widetilde{Y}$ of Y.

Exercise. Show that $\widetilde{\nabla}_X Y$ is independent of an extension $\tilde{Y}$ of Y.

Proof. Due to the uniqueness of the Levi-Civita connection it suffices to prove that

$$\nabla'_X Y := \top(\widetilde{\nabla}_X Y) \tag{1.25}$$

is a torsion-free metric connection on M.

Since $\top$ is a linear map, the map (1.25) is linear in X and additive in Y. Let us check the derivation property.

$$\nabla'_X(fY) =$$

$$= X(f)Y + f \cdot \nabla'_X Y$$

Thus ∇' is a connection.

Let X, Y be vector field on M. Since $\widetilde{\nabla}$ is torsion-free, we have

$$\widetilde{\nabla}_X Y - \widetilde{\nabla}_Y X = [X, Y].$$

On the right hand side we actually have the commutator of arbitrary vector fields $\tilde{X}, \tilde{Y}$ on N whose restrictions to M are X and Y. But this commutator is independent of the choice of extensions. One reason for this is that it is equal to the left hand side whose independence was established earlier. Another, more basic reason is that the commutator is "functorial":

$$[dF(X), dF(Y)] = dF([X, Y]).$$

In our case, $F\colon M \to N$ is the inclusion of M into N. By taking the tangential components, we obtain

$$\nabla'_X Y - \nabla'_Y X = [X, Y],$$

where the commutator on the right hand side can now be understood as the commutator in M. Thus the connection ∇' is torsion-free.

Finally, let us check that ∇' is metric.

$$X(g(Y, Z)) =$$

$$= g(\nabla'_X Y, Z) + g(Y, \nabla'_X Z).$$

Hence ∇' is the Levi-Civita connection on M, and therefore $\nabla' = \nabla$. □

1.13.6 The second fundamental form

Lemma 1.115. *Let X and Y be vector fields on $M \subset N$. Then the map*

$$B(X, Y) = \perp(\widetilde{\nabla}_X Y) \in \perp M$$

is symmetric and $C^\infty(M)$-bilinear. In particular, its value at $p \in M$ depends only on the vectors $X(p)$ and $Y(p)$.

Proof. By Theorem 1.114, $B(X,Y) = \widetilde{\nabla}_X Y - \nabla_X Y$ is the difference of two Levi-Civita connections ($\widetilde{\nabla}$ of the metric h on N and ∇ of the metric g on M). The argument from the proof of Theorem 1.89 shows that B is linear in Y. The argument from the proof of Theorem 1.108 shows that B is symmetric (as the difference of two torsion-free connections). □

Let us assume that $\dim N = \dim M + 1$. Then $\dim \perp_p M = 1$ and in $\perp_p M$ there are two unit vectors. Choose one of them. There is a unique way to extend this to a smooth field ν of unit normals in a neighborhood of p. Since $B(X,Y) \in \perp M$, we have

$$B(X,Y) = h(B(X,Y),\nu) \cdot \nu.$$

Lemma-Definition 1.116. *The map*

$$II(X,Y) := h(\widetilde{\nabla}_X \nu, Y)$$

is a symmetric bilinear form on M*, and*

$$B(X,Y) = -II(X,Y) \cdot \nu.$$

Proof. It suffices to prove the last formula. Since $B(X,Y)$ is symmetric bilinear, this will imply that $II(X,Y)$ is.

$$0 = X(h(Y,\nu)) = h(\widetilde{\nabla}_X Y, \nu) + h(Y, \widetilde{\nabla}_X \nu) = h(B(X,Y),\nu) + II(X,Y).$$

□

Bibliography

[BG80] Richard L. Bishop and Samuel I. Goldberg. *Tensor analysis on manifolds*. Dover Publications, Inc., New York, 1980. Corrected reprint of the 1968 original.

[GHL04] Sylvestre Gallot, Dominique Hulin, and Jacques Lafontaine. *Riemannian geometry*. Universitext. Springer-Verlag, Berlin, third edition, 2004.

[Spi79] Michael Spivak. *A comprehensive introduction to differential geometry. Vol. I*. Publish or Perish, Inc., Wilmington, Del., second edition, 1979.