Diffential Geometry: Lecture Notes

Dmitri Zaitsev

D. Zaitsev: School of Mathematics, Trinity College Dublin, Dublin 2, Ireland
E-mail address: zaitsev@maths.tcd.ie

Contents

CHAPTER 1

Introduction to Smooth Manifolds

Even things that are true can be proved.
Oscar Wilde, *The Picture of Dorian Gray*

1. Plain curves

Definition 1.1. A *regular arc* or *regular parametrized curve* in the plain $\mathbb{R}^2$ is any continuously differentiable map $f\colon I \to \mathbb{R}^2$, where $I = (a, b) \subset \mathbb{R}$ is an open interval (bounded or unbounded: $-\infty \leq a < b \leq \infty$) such that the $\mathbb{R}^2$-valued derivative $f'(t)$ is different from $0 = (0, 0)$ for all $t \in I$. That is for every $t \in I$, $f(t) = (f_1(t), f_2(t)) \in \mathbb{R}^2$ and either $f_1'(t) \neq 0$ or $f_2'(t) \neq 0$.

The variable $t \in I$ is called the *parameter* of the arc. One may also consider closed intervals, in that case their endpoints require special treatment, we'll see them as "boundary points".

Remark 1.2. There is a meaningful theory of nondifferentiable merely continuous arcs (including exotic examples such as Peano curves covering a whole square in $\mathbb{R}^2$) and of more restrictive injective continuous arcs (called Jordan curves) that is beyond the scope of this course.

The assumption $f'(t) \neq 0$ roughly implies that the image of f "looks smooth" and can be "locally approximated" by a line at each point. A map f with $f'(t) \neq 0$ for all t is also called *immersion*.

Example 1.3. Without the assumption $f'(t) \neq 0$ the image of f may look quite "unpleasant". For instance, investigate the images of the following C^∞ maps:

$$f(t) = (t^2, t^3), \quad f(t) := \begin{cases} (0, e^{1/t}) & t < 0 \\ (0, 0) & t = 0 \\ (e^{-1/t}, 0) & t > 0. \end{cases}$$

The first curve is called *Neil parabola* or *semicubical parabola*. Both maps are not regular at $t = 0$. Such a point is called a *critical point* or a *singularity* of the map f.

Definition 1.4. A *regular curve* is an equivalence class of regular arcs, where two arcs $f\colon I \to \mathbb{R}^2$ and $g\colon J \to \mathbb{R}^2$ are said to be equivalent if there exists a bijective continuously differential map $\varphi\colon I \to J$ with $\varphi'(t) > 0$ for all $t \in I$ (the inverse φ^{-1} is then exists and is automatically continuously differentiable) such that $f = g \circ \varphi$, i.e. $f(t) = g(\varphi(t))$ for all t. Sometimes a finite (or even countable) union of curves is also called a curve. A regular curve of class C^k for $1 \leq k \leq \infty$ is an equivalence class of regular arcs of class C^k (i.e. k times continuously differentiable), where the equivalence is defined via maps φ that are also of class C^k. "Smooth" usually stays for C^∞.

Recall that a map $f = (f_1, \dots, f_n)$ from an open set Ω in $\mathbb{R}^m$ into $\mathbb{R}^n$ is of class C^k (k times continuously differentiable) if all partial derivatives up to order k of every component f_i exist and are continuous everywhere on Ω.

A map φ as above will be called orientation preserving diffeomorphism between I and J; this will be defined later in a more general context.

Exercise 1.5. Give an example of a regular curve which is of class C^1 but not C^2.

The map $\varphi(t) := \tan t$ defines a C^∞ diffeomorphism between the intervals $(-\pi/2, \pi/2)$ and $(-\infty, \infty)$. Hence any regular curve can be parametrized by an arc defined over a bounded interval (why?).

Exercise 1.6. Give a parametrization of the line passing through 0 and a point $(a, b) \neq 0$ over the interval $(0, 1)$.

Example 1.7. One of the most important curves is the *unit circe.* Its standard parametrization is given by $f(t) = (\cos t, \sin t)$, $t \in \mathbb{R}$. Clearly f is not injective.

Exercise* 1.8. Show that the circle cannot be parametrized by an injective regular arc.

Definition 1.9. If a regular curve C is parametrized by an arc f, a *tangent vector* to C at a point $p = f(t_0)$ is any multiple (positive, negative or zero) of the derivative $f'(t_0)$. The *tangent line* to C at p is parametrized by $t \mapsto p + f'(t_0)t$. (Sometimes $t \mapsto f'(t_0)t$ is also called the tangent line).

Exercise 1.10. Show that the tangent line define above is independent of the choice of the parametrizing arc.

Exercise 1.11. In the above notation show that, if $t_n \in I$ is any sequence converging to t_0, then $f(t_n) \neq f(t_0)$ for n sufficiently large and that the line passing through $f(t_0)$ and $f(t_n)$ converges to the tangent line to C at p. Here "convergence" of lines can be defined as convergence of their unital directional vectors.

2. Surfaces in $\mathbb{R}^3$

The main difference between curves and surfaces is that the latter in general cannot be parametrized by a single regular map defined in an open set in $\mathbb{R}^2$. The simplest example is the unit sphere. Hence one may need different parametrizations for different points.

DEFINITION 1.12. A *parametrized (regular) surface element* (or *surface patch*) is a C^1 map $f\colon U \to \mathbb{R}^3$, where $U \subset \mathbb{R}^2$ is an open set, which is an immersion. The latter condition means that, for every $a \in U$, the differential $d_af\colon \mathbb{R}^2 \to \mathbb{R}^3$ is injective. Recall that

$$d_af(u_1, u_2) = u_1\frac{\partial f}{\partial t_1}(a) + u_2\frac{\partial f}{\partial t_2}(a).$$

The image of d_af is called the *tangent plane* at $f(a)$.

The condition that the differential d_af is injective is equivalent to linear independence of the partial derivative vectors $\frac{\partial f}{\partial t_1}(a)$ and $\frac{\partial f}{\partial t_2}(a)$. These span the *tangent plane* at $p = f(a)$.

EXAMPLE 1.13. Every open set $U \subset \mathbb{R}^2$ and every C^1 function f on U gives a surface in $\mathbb{R}^3$ via its *graph*

$$\{(t_1, t_2, f(t_1, t_2)) : (t_1, t_2) \in U\}.$$

In particular, a *hemisphere* is obtained for $f(t_1, t_2) = \sqrt{1 - t_1^2 - t_2^2}$ with U the open disc given by $t_1^2 + t_2^2 < 1$. One needs at least 2 surface elements to cover the sphere.

EXERCISE 1.14. Give two parametrized surface elements covering the unit sphere in $\mathbb{R}^3$. (Hint: Use stereographic projection).

EXAMPLE 1.15. A *torus (of revolution)* is an important surface that admits a global (but not injective) parametrization:

$$f(u, v) = \big((a + b\cos u)\cos v, (a + b\cos u)\sin v, b\sin u\big), \quad 0 < b < a, \quad (u, v) \in \mathbb{R}^2. \tag{2.1}$$

More generally, a *surface of revolution* is obtained by rotating a regular plane curve C parametrized by $t \mapsto (x(t), z(t))$ (called the *meridian* or *profile curve*) in the (x, z)-plane around the z-axis in $\mathbb{R}^3$, where C is assumed not to intersect the z-axis (i.e. $x(t) \neq 0$ for all t). It admits a parametrization of the form

$$(t, \varphi) \mapsto \big(x(t)\cos\varphi, x(t)\sin\varphi, z(t)\big). \tag{2.2}$$

Another important class of surfaces consists of *ruled surfaces.* A ruled surface is obtained by moving a line in $\mathbb{R}^3$ and admits a parametrization of the form

$$(t, u) \mapsto p(u) + tv(u) \in \mathbb{R}^3, \quad t \in \mathbb{R}, \quad u \in I, \tag{2.3}$$

where I is an interval in $\mathbb{R}$, $p, v\colon I \to \mathbb{R}^3$ are C^1 maps with v nowhere vanishing.

EXERCISE 1.16. Show that the map (2.2) defines a regular surface element. Find a condition on p and v in order that (2.3) define a regular surface element.

EXERCISE 1.17. Show that the hyperboloid given by $x^2 + y^2 - z^2 = 1$ is a ruled surface and the hyperboloid given by $x^2 + y^2 - z^2 = -1$ is not a ruled surface.

A general surface “in the large” is roughly defined by “patching together” surface elements. A precise definition (as an immersed 2-dimensional submanifold) will be given later. Here we give a definition of an “embedded regular surface”.

DEFINITION 1.18. A subset $S \subset \mathbb{R}^3$ is a *embedded regular surface* if, for each $p \in S$, there exist an open neighborhood V_p of p in $\mathbb{R}^3$ and a parametrized regular surface element $f_p \colon U_p \subset \mathbb{R}^2 \to V_p \cap S$ which is a homeomorphism between U_p and $V_p \cap S$. The map $f_p \colon U_p \to S$ is called a *parametrization* of S around p.

The most important consequence of the above definition is the fact that the change of parameters is a diffeomorphism:

THEOREM 1.19. *If $f_p \colon U_p \to S$ and $f_q \colon U_q \to S$ are two parametrizations as in Definition 1.18 such that $f_p(U_p) \cap f_q(U_q) = W \neq \emptyset$, then the maps*

$$(f_q^{-1} \circ f_p) \colon f_p^{-1}(W) \to f_q^{-1}(W), \quad (f_p^{-1} \circ f_q) \colon f_q^{-1}(W) \to f_p^{-1}(W) \tag{2.4}$$

are continuously differentiable.

The proof is based on the Implicit Function Theorem (or the Inverse Function Theorem), quoted here without proof:

THEOREM 1.20 (Implicit Function Theorem). *Consider an implicit equation*

$$F(x, y) = 0 \tag{2.5}$$

for a function $y = f(x)$, where $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$, $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$, and $F = (F_1, \ldots, F_n)$ is a C^1 map from an open neighborhood $U \times V$ of a point (a, b) in $\mathbb{R}^m \times \mathbb{R}^n$ into $\mathbb{R}^n$. Suppose that $F(a, b) = 0$ and the square matrix

$$\Big(\frac{\partial F_i}{\partial y_j}(a, b)\Big)_{1 \le i,j \le n} \tag{2.6}$$

is invertible. Then (2.5) is uniquely solvable near (a, b), i.e. there exists a possibly smaller open neighborhood $U' \times V' \subset U \times V$ of (a, b) in $\mathbb{R}^m \times \mathbb{R}^n$ and a C^1 map $f \colon U' \to V'$ such that, for $(x, y) \in U' \times V'$, (2.5) is equivalent to $y = f(x)$. If, moreover, F is of class C^k, $k > 1$, then f is also of class C^k.

An immediate consequence is the Inverse Function Theorem (sometimes the Implicit Function Theorem is deduced from the Inverse Function Theorem).

COROLLARY 1.21 (Inverse Function Theorem). *Let G be a C^1 map from an open neighborhood V of a point b in $\mathbb{R}^n$ into $\mathbb{R}^n$ with $a := G(b)$. Assume that the differential of G at b is invertible. Then G is also invertible near b, i.e. there exists an open neighborhood $V' \subset V$ of b in $\mathbb{R}^n$ such that $G(V')$ is open in $\mathbb{R}^n$, $G \colon V' \to G(V')$ is bijective onto and the inverse G^{-1} is C^1. If, moreover, G is C^k for $k > 1$, then also G^{-1} is C^k.*

PROOF OF THEOREM 1.19. Fix $b \in f_q^{-1}(W)$ and set $a := f_p^{-1}(f_q(b))$. By Definition 1.12, the differential $d_a f_p \colon \mathbb{R}^2 \to \mathbb{R}^3$ is injective. After a possible permutation of coordinates in $\mathbb{R}^3$, we may assume that the differential $d_a(f_p^1, f_p^2)$ is invertible, where $f_p = (f_p^1, f_p^2, f_p^3)$. Then, by the Inverse

Function Theorem, the map $(t_1, t_2) \mapsto (f_p^1(t_1, t_2), f_p^2(t_1, t_2))$ has a C^1 local inverse defined in a neighborhood of $(f_p^1(a), f_p^2(a))$ that we denote by φ. Then $f_p^{-1} \circ f_q = \varphi \circ (f_q^1, f_q^2)$ near b proving the conclusion of the theorem for the second map in (2.4). The proof for the first map is completely analogous. $\square$

3. Abstract Manifolds

DEFINITION 1.22. A *n-manifold* (or an *n-dimensional differentiable manifold) of class* C^k is a set M together with a family $(U_\alpha)_{\alpha \in A}$ of subsets and injective maps $\varphi_\alpha \colon U_\alpha \to \mathbb{R}^n$ whith open images $\varphi_\alpha(U_\alpha) \subset \mathbb{R}^n$ such that the union $\cup_{\alpha \in A} U_\alpha$ covers M and for any $\alpha, \beta \in A$ with $U_\alpha \cap U_\beta \neq \emptyset$, the sets $\varphi_\alpha(U_\alpha \cap U_\beta)$ and $\varphi_\beta(U_\alpha \cap U_\beta)$ are open in $\mathbb{R}^n$ and the composition map (called *transition map*)

$$(\varphi_\beta \circ \varphi_\alpha^{-1}) \colon \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta) \tag{3.1}$$

is of class C^k.

A family $(U_\alpha, \varphi_\alpha)_{\alpha \in A}$ as in Definition 1.22 is called a C^k *atlas* on M. Thus a C^k manifold is a set with a C^k atlas. A C^∞ manifold is often called *smooth manifold* and a C^0 manifold a *topological manifold*.

A comparison between Definition 1.22 and the defintion of an embedded regular surface (Definition 1.18) shows that the essential point (except for the change of dimension from 2 to n) was to distinguish the fundamental peroperty of the transition maps (3.1) (which is Theorem 1.19 for surfaces) and incorporate it as an axiom. This is the condition that will allow us to carry over ideas of differential calculus in $\mathbb{R}^n$ to abstract differential manifolds.

The condition on the transition maps (3.1) is only nontrivial in case there are at least two maps in the atlas. Hence, if the atlas consists of a single map, the manifold is C^∞ as in some of the examples below.

EXAMPLE 1.23. The most basic example of an n-manifold (of class C^∞) is the space $\mathbb{R}^n$ itself with the atlas consisting of the identity map $\varphi = \mathsf{id}$ from $U = \mathbb{R}^n$ into $\mathbb{R}^n$. More generally, every open subset $U \subset \mathbb{R}^n$ is an (C^∞) n-manifold with the atlas consisting again of the identity map.

EXAMPLE 1.24. A regular curve in $\mathbb{R}^2$ as defined in Definition 1.4, parametrized by an *injective* regular arc $f \colon I \to \mathbb{R}^2$ defines a 1-manifold structure on the image $f(I)$ via the atlas consisting of the inverse map $\varphi := f^{-1}$ defined on $U := f(I)$.

EXAMPLE 1.25. The simplest example of a 1-manifold for which one needs at least two maps in an atlas is the unit circle $S^1 := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$. An atlas with four maps can be given as follows. Let $U_1, \ldots, U_4$ be subsets of S^1 given by $x < 0$, $x > 0$, $y < 0$ and $y > 0$ respectively. Then a (C^∞) atlas is given by the maps $\varphi_\alpha \colon U_\alpha \to \mathbb{R}$, $\alpha = 1, \ldots, 4$, by $\varphi_1(x, y) := y$, $\varphi_2(x, y) := y$, $\varphi_3(x, y) := x$ and $\varphi_4(x, y) := x$. Indeed, it is easy to see that each φ_α is injective with open image $\varphi_\alpha(U_\alpha) = (-1, 1) \subset \mathbb{R}$, that the sets in (3.1) are open and the transition maps are C^∞. E.g., $(\varphi_2 \circ \varphi_4^{-1})(t) = \sqrt{1 - t^2}$ is smooth on $\varphi_4(U_2 \cap U_4) = (0, 1) \in \mathbb{R}$.

EXERCISE 1.26. Give an atlas on S^1 consisting only of 2 different maps φ_α.

EXAMPLE 1.27. An embedded regular surface S in $\mathbb{R}^3$ as in defined in Definition 1.18 can be given a structure of a 2-manifold with the atlas consisting of the inverse maps $f_p^{-1}\colon V_p \cap S \to \mathbb{R}^2$. The required property for the transition maps follows from Theorem 1.19.

EXAMPLE 1.28. Let $M \subset \mathbb{R}^3$ be the torus obtained as the image of the map $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ defined by (2.1). Then M has a natural structure of a smooth 2-manifold given by the atlas $(U_\alpha, \varphi_\alpha)_{\alpha\in A}$, where $A := \mathbb{R}^2$, $V_\alpha := (\alpha_1, 2\pi + \alpha_1) \times (\alpha_2, 2\pi + \alpha_2) \subset \mathbb{R}^2$ for $\alpha = (\alpha_1, \alpha_2) \in \mathbb{R}^2$, $U_\alpha := f(V_\alpha) \subset M$ and $\varphi_\alpha := (f|_{V_\alpha})^{-1}$. Each transition map $\varphi_\beta \circ \varphi_\alpha^{-1}$ equals the identity.

The following is an "exotic" manifold known as the real *line with double point*.

EXAMPLE 1.29. Take $M := \mathbb{R} \cup \{a\}$, where a is any point not in $\mathbb{R}$, and define a C^∞ atlas on M as follows. Set $U_1 := \mathbb{R}$ and $U_2 := \mathbb{R} \setminus \{0\} \cup \{a\}$, then clearly $M = U_1 \cup U_2$. Further define $\varphi_i\colon U_i \to \mathbb{R}$, $i = 1, 2$, by $\varphi_1(x) = x$ and $\varphi_2(x) := x$ for $x \neq a$ and $\varphi_2(a) := 0$. Then all assumptions of Definition 1.22 are satisfied (why?).

An easy way to obtain new manifolds is to take *products*. Let M be an n-manifold with an atlas $(U_\alpha, \varphi_\alpha)_{\alpha\in A}$ and M' be an n'-manifold with an atlas $(U'_\beta, \varphi'_\beta)_{\beta\in B}$, both of class C^k. Recall that $M \times M'$ is the set of all pairs (x, x') with $x \in M$ and $x' \in M'$. Then the collection of maps

$$\varphi_\alpha \times \varphi'_\beta\colon U_\alpha \times U_\beta \to \mathbb{R}^n \times \mathbb{R}^{n'} = \mathbb{R}^{n+n'}, \quad (\varphi_\alpha \times \varphi'_\beta)(x, y) := (\varphi_\alpha(x), \varphi'_\beta(y)), \quad (\alpha, \beta) \in C := A \times B,$$

defines a C^k atlas on $M \times M'$ making it an $(n + n')$-manifold of the same differentiability class (why?).

Any atlas as in Defintion 1.22 can be always completed to a maximal one by involving maps more general than φ_α that play the role of coordinates. It is very convenient and important for applications to have those general charts at our disposal.

DEFINITION 1.30. An injective map φ from a subset $U \subset M$ into $\mathbb{R}^n$ with open image $\varphi(U) \subset \mathbb{R}^n$ is called a *(coordinate) chart compatible with the atlas on* M or simply a *chart* on M if, by adding it to the given atlas one obtains a new (C^k) atlas on M, i.e. if, for every $\alpha \in A$, the images $\varphi(U \cap U_\alpha)$ and $\varphi_\alpha(U \cap U_\alpha)$ are open in $\mathbb{R}^n$ and the maps $\varphi \circ \varphi_\alpha^{-1}$ and $\varphi^{-1} \circ \varphi_\alpha$ are C^k in their domains of definition.

In particular, any map $\varphi_\alpha\colon U_\alpha \to \mathbb{R}^n$ in Definition 1.22 is a chart on M. The inverse $\varphi^{-1}\colon \varphi(U) \to U$ of a chart is called a *(local) parametrizations* of M and $\varphi(U) \subset \mathbb{R}^n$ the *parameter domain.*

LEMMA 1.31. *The set of all charts on M is a maximal atlas, i.e. it is an atlas and it cannot be extended to a larger family of maps satisfying the requirements of Definition 1.22.*

PROOF. Since each φ_α is a chart, the union of all charts covers M. In order to show that the set of all charts is an atlas, consider two charts $\varphi\colon U \to \mathbb{R}^n$ and $\psi\colon V \to \mathbb{R}^n$ with $U \cap V \neq \emptyset$. We first show that $\varphi(U \cap V \cap U_\alpha)$ is open (in $\mathbb{R}^n$) for every $\alpha \in A$. For this observe that both $\varphi_\alpha(U \cap U_\alpha)$ and $\varphi_\alpha(V \cap U_\alpha)$ are open and hence so is their intersection $W := \varphi_\alpha(U \cap V \cap U_\alpha)$. Since the map $h := (\varphi_\alpha \circ \varphi^{-1})\colon \varphi(U \cap U_\alpha) \to \varphi_\alpha(U \cap U_\alpha)$ is continuous, the set $\varphi(U \cap V \cap U_\alpha) = h^{-1}(W)$ is open. Interchanging the roles of φ and ψ we conclude that $\psi(U \cap V \cap U_\alpha)$ is also open. Now, since $M = \cup_\alpha U_\alpha$, $\varphi(U \cap V) = \cup_\alpha \varphi(U \cap V \cap U_\alpha)$ is open as a union of open sets and the same holds for $\psi(U \cap V) = \cup_\alpha \psi(U \cap V \cap U_\alpha)$.

In order to show that $g := (\psi \circ \varphi^{-1})\colon \varphi(U \cap V) \to \psi(U \cap V)$ is of class C^k, it is enough to show that so is the restriction of g to $\varphi(U \cap V \cap U_\alpha)$ for every $\alpha \in A$. The latter is C^k as composition of the C^k maps $(\varphi_\alpha \circ \varphi^{-1})\colon \varphi(U \cap V \cap U_\alpha) \to \varphi_\alpha(U \cap V \cap U_\alpha)$ and $(\psi \circ \varphi_\alpha^{-1})\colon \varphi_\alpha(U \cap V \cap U_\alpha) \to \psi(U \cap V \cap U_\alpha)$. Thus the family of all charts satifies the requirements of Definition 1.22 and hence forms an atlas.

Maximality follows directly: if there is a larger family of maps as in Definition 1.22, each of them must be a chart and hence the family is that of all charts. □

Definition 1.32. The set of all charts on M is called a *differentiable structure* on M.

The same differentiable structure can be obtained through a different atlas. Indeed, it follows from the proof of Lemma 1.31 that any covering of M by charts is an atlas defining the same differentiable structure.

Exercise 1.33. Show that the atlas consisting of the map $\varphi(x) = x^3$ defines a differentiable structure on $\mathbb{R}$ which is different from the standard one (defined in Example 1.23). That is, construct a chart for one atlas which is not a chart for the other atlas.

We shall use the convention that for a chart $\varphi\colon U \to \mathbb{R}^n$ on M, we shall write $(x^1, \ldots, x^n)$ for the standard coordinates of $\mathbb{R}^n$ and identify them with corresponding functions on U as follows. For $p \in U$, $x^i(p)$ is the ith coordinate of $\varphi(p)$, i.e. $\varphi(p) = (x^1(p), \ldots, x^n(p)) \in \mathbb{R}^n$. We call $(x^1(p), \ldots, x^n(p))$ coordinates of p.

4. Topology of abstract manifolds

Let M be an n-manifold with an atlas $(U_\alpha, \varphi_\alpha)_{\alpha\in A}$ as defined in previous section. Then M carries canonical topology called *manifold topology* defined as follows. A set $U \subset M$ is said to be *open* if for every $\alpha \in A$, the image $\varphi_\alpha(U \cap U_\alpha)$ is open in $\mathbb{R}^n$. Recall that a *topological space* is a set X together with a collection of subsets of X called open sets such that (1) the empty set and the whole set X are open; (2) union of any family of open sets is open; (3) intersection of any two open sets is open.

LEMMA 1.34. *With open sets defined above M is a topological space.*

EXERCISE 1.35. Prove the lemma.

LEMMA 1.36. *If $\varphi\colon U \to \mathbb{R}^n$ is any chart on M (see Definition 1.30), then U is open in M and φ is a homeomorphism onto its image $\varphi(U) \subset \mathbb{R}^n$. In particular, any φ_α is a homeomorphism onto its image.*

PROOF. By Definition 1.30, $\varphi\colon U \to \varphi(U)$ is a bijection. To see that it is continuous we have to prove that $\varphi^{-1}(V)$ is open in M for any open set V in $\mathbb{R}^n$. According to the definition of open sets this means that $\varphi_\alpha(\varphi^{-1}(V) \cap U_\alpha)$ has to be open for any α. The latter set coincides with $h^{-1}(V)$ where the transition map $h := (\varphi \circ \varphi_\alpha^{-1})\colon \varphi_\alpha(U \cap U_\alpha) \to \varphi(U \cap U_\alpha)$ is continuous. Hence $\varphi_\alpha(\varphi^{-1}(V) \cap U_\alpha) = h^{-1}(V)$ is open as desired.

To see that φ^{-1} is continuous we have to prove that $(\varphi^{-1})^{-1}(W) = \varphi(W \cap U)$ is open in $\mathbb{R}^n$ for any open set W in M. According to the definition of open sets, $\varphi_\alpha(W \cap U_\alpha)$ is open for every α. Since (U, φ) is a chart, $\varphi_\alpha(U \cap U_\alpha)$ is open and hence so is the intersection $\varphi_\alpha(W \cap U \cap U_\alpha)$. Since the transition map $h := \varphi_\alpha \circ \varphi^{-1}$ is continuous, the preimage $h^{-1}(\varphi_\alpha(W \cap U \cap U_\alpha)) = \varphi(W \cap U \cap U_\alpha)$ is open. Then also the union $\cup_\alpha \varphi(W \cap U \cap U_\alpha) = \varphi(W \cap U)$ is open as desired. $\square$

As a consequence, the topology on M depends only on the differentiable structure but not on the atlas defining this structure (why?).

EXERCISE 1.37. Check that the topology defined above is the only one for which every map φ_α is a homeomorphism onto its image $\varphi_\alpha(U_\alpha)$.

The following observation can be used to obtain new examples of manifolds generalizing Example 1.23. Any open subset U in a manifold M is automatically a manifold with induced differentiable structure defined as follows. If $(U_\alpha, \varphi_\alpha)_{\alpha\in A}$ is an atlas on M, the restrictions $(U_\alpha \cap U, \varphi_\alpha|_{U_\alpha})_{\alpha\in A}$ define an atlas on U.

EXERCISE 1.38. Check that the latter is indeed an atlas on U.

An important property of the topology of a manifold is that locally (or in the small) it is the same as that of $\mathbb{R}^n$. This is illustrated by the following statements.

LEMMA 1.39. *A manifold M is always locally compact and locally connected. That is, every point $p \in M$ has a fundamental system of open neighborhoods in M which are relatively compact and connected.*

PROOF. For any $p \in M$, there exists a map $\varphi_\alpha \colon U_\alpha \to \mathbb{R}^n$ from the atlas with $p \in U_\alpha$. Then take $V_p := \varphi_\alpha^{-1}(B)$ with B being sufficiently small open ball in $\mathbb{R}^n$ with center $\varphi_\alpha(p)$. The desired properties of V_p follow from the fact that φ_α is a homeomorphism onto its image. $\square$

A basic question about any topological space: Is it Hausdorff? Recall that a topological space X is called *Hausdorff* (or satisfying T_2-axiom) if any two distinct points $p, q \in X$ have disjoint open neighborhoods, i.e. open sets $U_p, U_q \in X$ with $p \in U_p$, $q \in U_q$ and $U_p \cap U_q = \emptyset$. A manifold does not have to be Hausdorff in general (see e.g. Example 1.29 or Berger-Gostiaux, 2.2.10.4). In most books (including Berger-Gostiaux) a manifold is always assumed to be Hausdorff. We shall also make this assumption in the sequel.

5. Submanifolds

We begin the discussion of submanifolds by introducing submanifolds of $\mathbb{R}^n$ which form an important class of manifolds. This is a direct generalization of Definition 1.18 of an embedded surface in $\mathbb{R}^3$.

Definition 1.40. A subset $S \subset \mathbb{R}^n$ is an *m-dimensional submanifold of $\mathbb{R}^n$ of class C^k* ($m \leq n$), if, for each $p \in S$, there exist an open neighborhood V_p of p in $\mathbb{R}^n$, an open set $\Omega_p \subset \mathbb{R}^m$ and a homeomorphism $f_p \colon \Omega_p \to V_p \cap S$ which is C^k and regular in the sense that for every $a \in \Omega_p$, the differential $d_a f_p \colon \mathbb{R}^m \to \mathbb{R}^n$ is injective. The map $f_p \colon \Omega_p \to S$ is called a *(local) parametrization* of S around p.

An m-dimensional submanifold S of $\mathbb{R}^n$ of class C^k carries a structure of an m-manifold of class C^k with the atlas $(U_p, \varphi_p)_{p\in S}$ given by $U_p := V_p \cap S$, $\varphi_p := f_p^{-1}$, where V_p and f_p are as the above definition. A direct generalization of Theorem 1.19 (with the same proof based on the Inverse Function Theorem) implies that the transition maps (3.1) are of class C^k. Hence the collection $(U_p, \varphi_p)_{p\in S}$ is indeed a C^k atlas on M.

The following is an important characterization of submanifolds of $\mathbb{R}^n$.

Theorem 1.41. *Let S be a subset of $\mathbb{R}^n$. The following properties are equivalent:*

(i) *S is an m-dimensional submanifold of $\mathbb{R}^n$ of class C^k.*
(ii) *For every $p \in S$ there exists an open neighborhood $V \subset \mathbb{R}^n$ of p and a C^k diffeomorphism φ between V and an open set in $\mathbb{R}^n$ such that $S \cap V = \varphi^{-1}(\mathbb{R}^m \times \{0\})$, where $\mathbb{R}^m \subset \mathbb{R}^n$ is the set of vectors whose last $n-m$ coordinates are* 0.
(iii) *For every $p \in S$ there exists an open neighborhood $V \subset \mathbb{R}^n$ of p and C^k functions $g_i \colon V \to \mathbb{R}$, $i = 1, \ldots, n-m$, such that the gradients $\nabla g_i(x)$ are linearly independent and $S \cap V = \{x \in V : g_1(x) = \cdots = g_{n-m}(x) = 0\}$.*
(iv) *For every $p \in S$ there exists an open neighborhood $V \subset \mathbb{R}^n$ and a submersion $g \colon V \to \mathbb{R}^{n-m}$ (i.e. the differential $d_a g \colon \mathbb{R}^n \to \mathbb{R}^{n-m}$ is surjective for each $a \in V$) such that $S \cap V = g^{-1}(0)$.*
(v) *S is locally a graph of a C^k function from $\mathbb{R}^m$ to $\mathbb{R}^{n-m}$, i.e. for every $p \in S$ having, after possible reordering the coordinates, the form $p = (p^1, \ldots, p^n) \in \mathbb{R}^n$, there exist neighborhoods $V_1 \subset \mathbb{R}^m$ of $(p^1, \ldots, p^m)$ and $V_2 \subset \mathbb{R}^{n-m}$ of $(p^{m+1}, \ldots, p^n)$ and a C^k map $h \colon V_1 \to V_2$ such that $S \cap (V_1 \times V_2) = \{(x, h(x)) : x \in V_1\}$ (the graph of h).*

In (i) we think of a submanifold as locally regularly parametrized, in (ii) as locally diffeomorphic to an open set in $\mathbb{R}^m$ embedded into $\mathbb{R}^n$ in the standard way, in (iii) as given by a regular system of equations, in (iv) as the zero-set of a submersion, and in (v) as locally a graph of a smooth map.

Proof of Theorem 1.41. Given (i) and a local parametrization $f_p \colon \Omega_p \to S$, we embedd $\Omega_p \subset \mathbb{R}^m$ into $\mathbb{R}^n$ in the standard way and extended f_p to a neighborhood of $a := f_p^{-1}(p)$ in $\mathbb{R}^n$ as follows. After a permutation of coordinates, we first assume that $d_a(f_p^1, \ldots, f_p^m) \colon \mathbb{R}^m \to \mathbb{R}^m$ is invertible. Then set

$$\psi(u^1, \ldots, u^n) := f_p(u^1, \ldots, u^m) + (0, \ldots, 0, u^{m+1}, \ldots, u^n).$$

Then it is easy to check that $d_a\psi\colon \mathbb{R}^n \to \mathbb{R}^n$ is invertible and hence, by the Inverse Function Theorem, $\varphi := \psi^{-1}$ exists and satisfies (ii) for a possibly smaller neighborhood V of p.

Given (ii), we can take $g_i := x^{m+i} \circ \varphi^{-1}$, $i = 1, \ldots, n-m$, to satisfy (iii). Given (iii), we obtain (iv) with g being a restriction of $(g_1, \ldots, g_{n-m})$ to a suitable neighborhood of x.

Given (iv), observe that the $n \times (n-m)$ matrix associated with $d_p g$ has the maximal rank $n-m$. Therefore, after a permutation of coordinates of $\mathbb{R}^n$, we can apply Theorem 1.20 (Implicit Function Theorem) whose conclusion implies (v).

Finally, given (v), a local parametrization of S as in Definition 1.40 can be easily obtained by setting $\Omega_p := V_1$, $f_p(x) := (x, h(x))$. Hence (v) implies (i). □

An important special case of a submanifold of $\mathbb{R}^n$ is a *hypersurface* which is by definition, an $(n-1)$-dimensional submanifold. According to property (iii) in Theorem 1.41 a subset $S \subset \mathbb{R}^n$ is a C^k hypersurface if and only if it can be locally defined as the zero-set of a single function with nontrivial gradient.

We now extend the notion of submanifolds in $\mathbb{R}^n$ by defining submanifolds of abstract manifolds following the line of the property (ii) in Theorem 1.41 and using general charts as defined in Definition 1.30.

Definition 1.42. Let M be an n-manifold of class C^k. A subset $S \subset M$ is called an *m-dimensional submanifold* of M (of the same class C^k) if, for every $p \in S$, there exists a chart (U, φ) on M with $p \in U$ and $S \cap U = \varphi^{-1}(\mathbb{R}^m)$.

Remark 1.43. More generally, a C^r submanifold of M can be defined for $r \le k$ by viewing M as a C^r manifold.

It is straightforward to see that, for (U, φ) ranging over all charts satisfying the conditions in Definition 1.42, the pairs $(S \cap U, \varphi|_{S\cap U})$ form a C^k atlas on S definining on S a structure of a C^k manifold.

Exercise 1.44. Show that the manifold topology on a submanifold $S \subset M$ coincides with that induced by the manifold topology on M.

Exercise 1.45. Show that being submanifold is a transitive relation: If M_1 is a submanifold of M_2 and M_2 of M_3, then M_1 is a submanifold of M_3.

6. Differentiable maps, immersions, submersions and embeddings

Let $\Omega \subset \mathbb{R}^n$ and $\Omega' \subset \mathbb{R}^{n'}$ be open sets. Recall that a map $f\colon \Omega \to \Omega'$ is of class C^k if all partial derivatives of f up to order k exist and are continuous on Ω.

Definition 1.46. Let M be an n-manifold with an atlas $(U_\alpha, \varphi_\alpha)_{\alpha\in A}$ of class C^r and M' be an n'-manifold with an atlas $(U'_\beta, \varphi'_\beta)_{\beta\in B}$ of class $C^{r'}$. A *continuous* map $f\colon M \to M'$ is called *differentiable of class* C^k with $k \le \min(r, r')$ if, for every $x \in M$ and every $\alpha \in A$ and $\beta \in B$ with $x \in U_\alpha$ and $f(x) \in U'_\beta$, the composition $\varphi'_\beta \circ f \circ \varphi_\alpha^{-1}$ is of class C^k in an open neighborhood of $\varphi_\alpha(x)$.

Remark 1.47. The continuity of f is assumed with respect to the canonical topologies on M and M' as defined in previous section. The continuity of f guarantees that $f(y) \in U'_\beta$ for y in a neighborhood of x and hence the composition $\varphi'_\beta \circ f \circ \varphi_\alpha^{-1}$ is defined in a neighborhood of $\varphi_\alpha(x)$ (by Lemma 1.36, both φ_α^{-1} and φ'_β are continuous on their domains of definition).

Remark 1.48. Given any $x \in M$, there always exist α and β as in Definition 1.46 because $M = \cup_{\alpha\in A} U_\alpha$ and $M' = \cup_{\beta\in B} U'_\beta$.

Remark 1.49. The reason to choose $k \le \min(r, r')$ is to guarantee that the property of f to be C^k does not change when composing with C^r maps on the right and $C^{r'}$ maps on the left. This choice is also illustrated by the following statements.

Example 1.50. If $S \subset M$ is a submanifold of a manifold M, the inclusion $\imath\colon S \to M$ is automatically differentiable (why?)

Definition 1.51. A C^k *diffeomorphism* between two manifolds M and M' of the same dimension is a bijection $f\colon M \to M'$ such that both f and f^{-1} are of class C^k. Two manifolds are said to be C^k *diffeomorphic* if there exists a C^k diffeomorphism between them.

Being C^k diffeomorphic is the basic equivalence relation in Differential Topology.

We mention without proof the following fundamental result (see M. Hirsch, *Differential Topology*, Chapter 2, Theorem 2.10 for a proof):

Theorem 1.52. *For $r \ge 1$, every C^r manifold is C^r diffeomorphic to a C^∞ manifold. If for $1 \le r < s \le \infty$, two C^s manifolds are C^r diffeomorphic, they are C^s diffeomorphic.*

Definition 1.53. In the notation of Definition 1.46, a C^1 map f is said to be an *immersion* at a point $x \in M$ if $\varphi'_\beta \circ f \circ \varphi_\alpha^{-1}$ is an immersion at $\varphi_\alpha(x)$, i.e. if the differential

$$d_{\varphi_\alpha(x)}(\varphi'_\beta \circ f \circ \varphi_\alpha^{-1})\colon \mathbb{R}^n \to \mathbb{R}^{n'} \tag{6.1}$$

is injective. Similarly f is said to be a *submersion* at $x \in M$ if the differential in (6.1) is surjective. Without point specified, f is an immersion or submersion if it so at every point. Finally, f is called an *embedding* if it is an injective immersion and is a homeomorphism onto its image $f(M)$.

An important property of being immersion or submersion at a point is that the latter properties automatically hold in a neighborhood of that point.

Proposition 1.54. *If $S \subset M$ is a submanifold, the inclusion $\imath\colon S \to M$ is an embedding. Vice versa, if S and M are two manifolds and $\imath\colon S \to M$ is an embedding, then the image $\imath(S)$ is a submanifold of M.*

Exercise 1.55. Prove this proposition.

An important relation between immersions and embedding is the following statement which is a consequence of the Inverse Function Theorem:

Proposition 1.56. *If $f\colon M \to M'$ is an immersion at $x \in M$, there exists an open neighborhood U of x in M such that the restriction $f|_U$ is an embedding of U into M'.*

Example 1.57. Both regular arc and regular surface elements are examples of immersions.

Exercise 1.58. Prove that an injective immersion of a compact Hausdorff manifold is an embedding. Can "compact" be dropped?

Exercise 1.59. Given two C^k manifolds M_1 and M_2, prove that the canonical projections $\pi_i\colon M_1 \times M_2 \to M_i$, $i = 1, 2$, are C^k submersions.

Exercise 1.60. For each $k = 0, 1, \ldots$, define $\varphi_k\colon \mathbb{C} \to \mathbb{C}$ and $\psi_k\colon \mathbb{C}^* \to \mathbb{C}$ ($\mathbb{C}^* := \mathbb{C} \setminus \{0\}$) by $\varphi_k(z) = \psi_k(z) := z^k$. For which k are φ_k and ψ_k immersions, submersions or embeddings?

CHAPTER 2

Basic results from Differential Topology

A topologist is one who doesn't know the difference between a doughnut and a coffee cup.

John Kelley, In N. Rose, *Mathematical Maxims and Minims*

1. Manifolds with countable bases

From now one a manifold M will be assumed to be Hausdorff and to have a countable basis:

DEFINITION 2.1. A topological space T is said to have *countable basis (or base)* or to be *separable* if there exists a countable family (U_i) of open sets which form a *basis*, that is, one (and hence the other) of the following equivalent conditions holds:

(1) For any $x \in M$ and any neighborhood V of x, there is α with $x \in U_i \subset V$;
(2) Any open set in T is a union of some of the U_i.

EXERCISE 2.2. Show that (1) and (2) are indeed equivalent.

EXAMPLE 2.3. $\mathbb{R}^n$ has a countable basis. The basis can be taken among open balls with rational radius and center.

EXERCISE 2.4. Show that any submanifold of $\mathbb{R}^n$ has also countable basis.

On a manifold a basis can be chosen with additional properties:

LEMMA 2.5. *A manifold M has a countable basis (U_i) whose elements are relatively compact (i.e. their closures $\overline{U}_i$ in M are compact).*

PROOF. Given a countable basis (V_α), the subfamily consisting of relatively compact V_α's is again a basis satisfying the conclusion. □

The following simple but important statement asserts the existence of countable compact exhaustions:

LEMMA 2.6. *A manifold M can be exhausted by countably many compact subsets in the sense that there exists a sequence (K_m) of compact sets in M with $K_m \subset \mathrm{Int}K_{m+1}$ for every n and $\cup_m K_m = M$.*

PROOF. Given a countable relatively compact basis $(V_m)_{m\in\mathbb{N}}$ as in Lemma 2.5, we construct the sequence (K_m) inductively as follows. Set $K_1 := \overline{V}_1$. Assuming K_m is constructed, since it is compact, it is covered by finitely many V_j's. We can choose $j > m$ such that

$$K_m \subset V_1 \cup \cdots \cup V_j$$

and set $K_{m+1} := \overline{V}_1 \cup \cdots \cup \overline{V}_j$. The sequence K_m, inductively constructed, satisfies the desired properties. □

2. Partition of unity

Partition of unity is an important tool for glueing together functions (and other objects) defined locally.

Definition 2.7. A C^k partition of unity on a C^k manifold M is a family $(h_i)_{i\in I}$ of C^k functions $h_i\colon M \to \mathbb{R}$ satisfying the following:

(1) $0 \le h_i \le 1$ for all $i \in I$;
(2) every point $p \in M$ has a neighborhood intersecting only finitely many sets

$$\operatorname{supp} h_i := \overline{x \in M : h(x) \neq 0}; \tag{2.1}$$

(3) $\sum_{i\in I} h_i \equiv 1$ (locally this is always a finite sum in view of (2)).

In practice it is useful to have partitions of unity subordinate to a given covering of M in the sense that each $\operatorname{supp} h_i$ is contained in (at least one) single element of the covering:

Definition 2.8. A partition of unity $(h_i)_{i\in I}$ is said to be *subordinate* to an open covering (W_α) of M if for every $i \in I$ there is an α with $\operatorname{supp} h_i \subset W_\alpha$.

Construction of a partition of unity is based on the existence of the following "cut off functions":

Lemma 2.9. *For every n, there exists a C^∞ function $h\colon \mathbb{R}^n \to \mathbb{R}$ with $h \ge 0$ such that $h(x) = 1$ for $\|x\| \le 1$ and $h(x) = 0$ for $\|x\| \ge 2$.*

Proof. Let

$$f(r) := \begin{cases} e^{-1/r^2} & r > 0 \\ 0 & r \le 0. \end{cases}$$

Then $f\colon \mathbb{R} \to \mathbb{R}$ is C^∞ and the function

$$g(r) := \frac{f(2-r)}{f(2-r) + f(r-1)}$$

is C^∞ with values in $[0,1]$ such that $g(r) = 1$ for $r \le 1$ and $g(r) = 0$ for $r \ge 2$. It remains to set $h(x) := g(\|x\|)$. □

We can now prove the main existence theorem for a partition of unity.

Theorem 2.10. *Let (W_α) be an open covering of a C^k manifold M. Then there exists a C^k partition of unity subordinate to (W_α).*

Proof. Let $(K_m)_{m\in\mathbb{N}}$ be a compact exhaustion of M as in Lemma 2.6. For each fixed m, consider the open sets

$$(\operatorname{Int} K_{m+2} \setminus K_{m-1}) \cap W_\alpha.$$

These sets (for fixed m and varying α) cover the compact set $C_m := K_{m+1} \setminus \mathrm{Int}K_m$. For every $x \in C_m$, choose a chart (U_x, φ_x) with

$$x \in U_x \subset (\mathrm{Int}K_{m+2} \setminus K_{m-1}) \cap W_\alpha$$

for some α and such that $\varphi_x(x) = 0$ and $\varphi_x(U_x)$ contains the ball $B_3(0) \subset \mathbb{R}^n$ with center 0 and radius 3. These properties are easy to obtain starting from any φ_x and composing it with a suitable affine transformation of $\mathbb{R}^n$. Since C_m is compact, it is covered by finitely many open sets of the form $\varphi_x^{-1}(B_1(0))$. By collecting these sets for all m, we obtain a family of charts (U_j, φ_j) satisfying the following properties:

(1) for every U_j there exists α with $U_j \subset W_\alpha$;
(2) (U_j) is locally finite, i.e. for every $x \in M$, there exists a neighborhood of x that meets only finitely many U_j's;
(3) $\varphi_j(V_j) \supset B_3(0)$ for all j;
(4) $\cup_j \varphi_j^{-1}(B_1(0)) = M$.

With h as in Lemma 2.9, we now define functions $f_j \colon M \to \mathbb{R}$ by

$$f_j(y) = \begin{cases} h(\varphi_j(y)) & y \in U_j \\ 0 & y \notin U_j. \end{cases}$$

It follows from property (3) above and the properties of h that f_j is a C^k function on M. Furthermore, $f_j = 1$ on $\varphi_j^{-1}(B_1(0))$. Then it remains to set

$$g := \sum_i g_i, \quad h_i := g_i/g$$

to obtain a partition of unity (h_i) as desired. Indeed, (2) implies that the sum is locally finite and yields a well-defined C^k function. Furthermore, (4) implies that $g > 0$ everywhere on M and hence h_i is well-defined. Since $\operatorname{supp} h_j \subset U_j$, (h_j) is subordinate to (W_α) in view of (1). □

A partition of unity is used to glue together locally defined functions. For instance, we obtain the following simple application of Theorem 2.10:

LEMMA 2.11. *Let M be a C^k manifolds with $p, q \in M$, $p \neq q$. Then there exists a C^k function $f \colon M \to \mathbb{R}$ with $0 \leq f \leq 1$, $\{p\} = f^{-1}(0)$, $\{q\} = f^{-1}(1)$.*

PROOF. Let (U_p, φ_p) and (U_q, φ_q) be disjoint charts at p and q respectively with $\varphi_p(p) = 0$, $\varphi_p(q) = 0$ and $\|\varphi_p(x)\| < 1$, $\|\varphi_q(y)\| < 1$ for $x \in U_p$, $y \in U_q$. Consider the covering of M by U_p, U_q and $U_0 := M \setminus \{p, q\}$. For each set in this covering we can solve the problem by taking $f_p := \|\varphi_p\|^2$ on U_p, $f_q := 1 - \|\varphi_q\|^2$ on U_q and $f_0 := 1/2$ on U_0 respectively. By Theorem 2.10, there exists a partition of unity (h_i) subordinate to the covering $\{U_p, U_q, U_0\}$. By summing h_i's together we may assume that $i = p, q, 0$ and $\operatorname{supp} h_i \subset U_i$. Then

$$f := h_p f_p + h_q f_q + h_0 f_0$$

is a C^k function on M satisfying the conclusion of the lemma. □

3. Regular and critical points and Sard's theorem

We shall assume all manifolds to be Hausdorff and to have a countable topology basis.

In view of Property (iv) in Theorem 1.41, points where maps are submersive are of particular interest.

DEFINITION 2.12. Let M and M' be C^k manifolds ($k \geq 1$) and $f\colon M \to M'$ a C^k map. A point $x \in M$ is said to be *regular* for f if f is a submersion at x, and *critical* otherwise. A point $y \in M'$ is a *regular value* for f is every point $x \in f^{-1}(y)$ is regular, and a *critical value* otherwise (even if $y \notin f(M)$).

In case $\dim M' > \dim M$ it is easy to see that all points of M are critical and all points of $f(M)$ are critical values.

The following *regular value theorem* is often used to construct manifolds:

THEOREM 2.13. *Let $f\colon M \to M'$ be a C^k map between two C^k manifolds. If $y \in M'$ is a regular value, then $f^{-1}(y)$ is a C^k submanifold of M.*

The proof follows from Theorem 1.41.

EXERCISE 2.14. Give an example of a C^∞ map $f\colon \mathbb{R} \to \mathbb{R}$ with infinitely many critical values.

A profound result known as *Sard's theorem* says that the set of critical values cannot be "too large". To measure the "largeness" we have to introduce sets of measure zero.

DEFINITION 2.15. An *n-cube* $C \subset \mathbb{R}^n$ of *edge* $\lambda > 0$ is any set C obtained from $[0,\lambda]^n$ by euclidean motion. The measure (volume) of C is $\mu(C) = \lambda^n$. A subset $S \subset \mathbb{R}^n$ is *of measure zero* if for every $\varepsilon > 0$, it can be covered by a countable union of n-cubes, the sum of whose measures is less than ε. A subset S in a manifold is said to be *of measure zero* if for every chart (U,φ) on M the set $\varphi(S \cap U) \subset \mathbb{R}^n$ is of measure zero.

EXERCISE 2.16. Show that a countable set is of measure zero.

Note that we have not defined the "measure" of a subset of M but only a certain kind of subsets that are of measure zero. A countable union of sets of measure zero is again a set of measure zero (why?). On the other hand, it follows from elementary Measure Theory that a cube is not of measure zero. The latter fact implies the following important property used in many applications:

LEMMA 2.17. *The complement of a set of measure zero is dense.*

The property to be of measure zero is a rather flexible notion invariant under diffeomorphisms and even under general C^1 maps:

LEMMA 2.18. *Let $U \subset \mathbb{R}^n$ be an open set and $f\colon U \to \mathbb{R}^n$ a C^1 map. If $X \subset U$ is of measure zero, so is $f(X)$.*

PROOF. The set U can be covered by countably many balls $B \subset U$ such that the norm $\|d_x f\|$ is uniformly bounded on B by a constant $C = C(B)$. Then

$$|f(x) - f(y)| \leq C|x - y|$$

for all $x, y \in B$. It follows that if $C \subset B$ is an n-cube of edge λ, then $f(C)$ is contained in an n-cube C' of edge less than $\sqrt{n}C\lambda$. Hence $\mu(C') \leq (\sqrt{n}C)^n \mu(C)$. The latter bound implies that $f(X \cap B)$ is of measure zero. Then $f(X)$ is also of measure zero as a countable union of sets of measure zero. □

We now state the following important theorem:

THEOREM 2.19 (Sard's Theorem). *The set of critical values of a C^k map $f\colon M \to M'$ between C^k manifolds ($k \geq 1$) is of measure zero provided*

$$k > \max(0, \dim M - \dim M'). \tag{3.1}$$

In particular, the set of critical values of a C^∞ map between C^∞ manifolds is always of measure zero.

With Lemma 2.17 we obtain the following important consequence:

COROLLARY 2.20. *Under the assumptions of Theorem 2.19 the set of regular values of f is dense in M'.*

EXAMPLE 2.21. A constant map has all points as critical but only one critical value.

In case of a map $f\colon M \to M'$ with $\dim M' > \dim M$ the set of critical points is exactly $f(M)$ and hence Sard's Theorem asserts that $f(M)$ is of measure zero. This assertion is relatively easy to prove.

EXERCISE 2.22. Give a proof of this assertion. Hint. Use Lemma 2.18.

We prove Sard's theorem here only in the equidimensional case $\dim M = \dim M'$. The proof for $\dim M > \dim M'$ is more difficult and requires consideration of higher order partial derivatives.

PROOF OF SARD'S THEOREM FOR $\dim M = \dim M'$. Since M and M' have countable bases consisting of charts, the general case is reduced to that where M and M' are relatively compact open sets in $\mathbb{R}^n$ and f is C^1 map from a neighborhood of $K := \overline{M} \subset \mathbb{R}^n$ into $\mathbb{R}^n$. We can further assume that K is contained in the cube $[0,1]^n$.

If a point $a \in M$ is critical, the image $d_a f(\mathbb{R}^n)$ is a lower-dimensional plane which is of measure zero in $\mathbb{R}^n$. The idea of the proof is to compare the actual map f with its first order Taylor expansion $f(a) + d_a f(x-a)$ by estimating the difference $f(x) - f(a) - d_a f(x-a)$.

Since the first order partial derivatives of f are continuous on the compact set K, we have $\|d_a f\| \leq C$ for some $C > 0$ and all $a \in K$. Fix $\varepsilon > 0$. Since the first order partial derivatives of f are also uniformly continuous on K, there exists $\delta > 0$ such that $\|d_a f - d_b f\| \leq \varepsilon$ whenever $\|a - b\| \leq \delta$ for $a, b \in K$. We now estimate the difference mentioned above for $\|x - a\| \leq \delta$:

$$\begin{aligned}\|f(x) - f(a) - d_a f(x-a)\| &= \left\| \int_0^1 \frac{d}{dt} f(a + t(x-a))dt - d_a f(x-a) \right\| \\ &= \left\| \int_0^1 (d_{a+t(x-a)} f - d_a f)(x-a)dt \right\| \leq \sup_{0 \leq t \leq 1} \|d_{a+t(x-a)} f - d_a f\| \cdot \|x - a\| \leq \varepsilon\delta. \end{aligned} \tag{3.2}$$

Let now C_δ be an n-cube of edge $\delta/\sqrt{n}$ containing a critical point a. Then the affine map $h_a(x) := f(a) + d_a f(x-a)$ sends C_δ into an $(n-1)$-cube of edge $2C\delta$ inside a hyperplane $H \subset \mathbb{R}^n$ (with respect to some euclidean coordinates in H), where $C > 0$ is the bound for $\|df\|$ chosen above. Now (3.2) implies that $f(C_\delta)$ is contained in a parallelepiped of measure $2C\varepsilon\delta^n$ which can be covered by finitely many cubes whose sum of measures does not exceed $4C\varepsilon\delta^n$.

Without loss of generality, $\delta = 1/m$ for some integer m. Then $K \subset [0,1]^n$ can be covered by m^n disjoint n-cubes of edge δ. Some of these cubes contain critical points. Above we have seen that the image of each of those cubes is covered by cubes of total measure not greater than $4C\varepsilon\delta^n$. Summing over all cubes we see that the total set of critical values can be covered by cubes whose total measure does not exceed $4C\varepsilon$. It remains to observe that ε can be arbitrarily small. $\square$

As an application of Sard's theorem we show that there is no retraction from the closed unit ball in $\mathbb{R}^n$ onto the unit sphere. Recall that a *retraction* from a topological space X onto a subspace Y is any continuous map $r\colon X \to Y$ such that the restriction of r to X is the identity.

THEOREM 2.23. *There is no retraction from the ball*

$$\overline{B^n} := \{\|x\| \le 1\} \subset \mathbb{R}^n$$

onto the sphere $S^{n-1} := \{\|x\| = 1\}$.

PROOF. Suppose there is a retraction $r\colon \overline{B^n} \to S^{n-1}$. We first reduce the general case to the case when r extends to a C^∞ map in a neighborhood of $\overline{B^n}$. For this, we find a new retraction $g\colon \overline{B^n} \to S^{n-1}$ which is C^∞ in a neighborhood of S^{n-1} by extending r continuously to a closed ball of radius $1+\varepsilon$ as constant in radial directions and scaling the ball with coefficient $(1+\varepsilon)^{-1}$. Then we approximate g by a C^∞ map into S^{n-1} which agrees with g on a neighborhood of S^{n-1}.

Now, by Sard's theorem, g defined in a neighborhood of $\overline{B^n}$, has a regular value $y \in S^{n-1}$ and therefore the preimage $V := g^{-1}(y)$ is an one-dimensional closed submanifold in a neighborhood of $\overline{B^n}$. We have $y \in V$ and the connected component of y in V (which is homeomorphic either to a circle or to an interval) intersects the interior of $\overline{B^n}$ and hence must contain another point of S^{n-1}. However the retraction condition implies $S^{n-1} \cap V = \{y\}$ which is a contradiction. $\square$

Theorem 2.23 implies Brouwer's fixed-point theorem:

THEOREM 2.24 (Brouwer's fixed-point theorem). *Any continuous map* $f\colon \overline{B^n} \to \overline{B^n}$ *has a fixed point.*

Indeed, if there is a continuous map as in Theorem 2.24 without fixed points, then one can obtain a retraction r as in Theorem 2.23 by letting $r(x)$ to be the intersection of S^{n-1} with the ray passing starting from $f(x)$ and passing through x.

4. Whitney embedding theorem

THEOREM 2.25 (Whitney embedding theorem). *Any n-manifold M of class C^k $(k \geq 1)$ can be C^k embedded into $\mathbb{R}^{2n+1}$.*

We prove Theorem 2.25 in the easy case when M is compact and of class at least C^2. The proof is split into two steps. First we construct an embedding into $\mathbb{R}^m$ for some possibly large m. Then we project the embedded copy of M to a smaller linear subspace of dimension $(2n+1)$ in such a way that the projection is still an embedding.

PROPOSITION 2.26. *Let M be a compact manifold. Then there exists an integer m and an embedding of M into $\mathbb{R}^m$.*

PROOF. We use the proof of Theorem 2.10 on the existence of a partition of unity. There it was constructed a covering of M by coordinate charts (U_i, φ_i) and smooth functions $f_i \colon M \to \mathbb{R}$, $0 \leq f_i \leq 1$, with $\mathrm{supp} f_i \subset U_i$ such that the interiors of the preimages $f_i^{-1}(1) \subset U_i$ for all i still cover M. Since M is compact we may assume that the covering is finite, i.e. $i = 1, \ldots, l$. The the map

$$g(x) := (f_1(x)\varphi_1(x), \ldots, f_l(x)\varphi_l(x), f_1, \ldots, f_l) \in \mathbb{R}^n \times \cdots \times \mathbb{R}^n \times \mathbb{R} \times \cdots \times \mathbb{R} = \mathbb{R}^{nl+l},$$

where each component $f_i(x)\varphi_i(x)$ is extended by 0 for $x \notin U_i$, defines an embedding of M into $\mathbb{R}^{nl+l}$. $\square$

In view of Proposition 2.26 we may assume that M is a submanifold of $\mathbb{R}^m$.

PROPOSITION 2.27. *Let $M \subset \mathbb{R}^m$ be a compact submanifold of dimension n of class at least C^2 such that $m \geq 2n+1$. Then, after a suitable linear change of coordinates in $\mathbb{R}^m$, the standard projection from M to the subspace $\mathbb{R}^{2n+1} \subset \mathbb{R}^m$ is an embedding.*

PROOF. We may assume that $m > 2n+1$. For every unit vector $v \in S^{m-1} \subset \mathbb{R}^m$, denote by f_v the orthogonal projection along v from $\mathbb{R}^m$ onto the orthogonal complement of v (isomorphic to $\mathbb{R}^{m-1}$). We want to choose v such that $f_v|_M$ is an embedding.

The injectivity of $f_v|_M$ means that v is not parallel to any secant of M, i.e.

$$v \neq \frac{x-y}{\|x-y\|}, \quad \forall x \neq y \in M, \tag{4.1}$$

or, equivalently, v is not in the image of the map $c \colon (M \times M) \setminus \Delta \to S^{m-1}$, $(x, y) \mapsto (x-y)/\|x-y\|$, where $\Delta := \{(x, x) : x \in M\}$ is the diagonal. Since $\dim(M \times M) = 2n < m-1$, there is a dense set of points v not in the image of c (e.g. by the easy case of Sard's theorem).

The property of $f_v|_M$ to be an immersion means that v is not tangent to M at any point. Again, since v is a unit vector, the latter property means that v is not in the in the image of the map

$$\tau \colon S \to S^{m-1}, \quad S := \{(x, v) \in M \times \mathbb{R}^m : v \in T_x M, \|v\| = 1\}, \tag{4.2}$$

where $T_x M$ denotes the space of all vectors tangent to M at x. Then S is a $(2n-1)$-dimensional submanifold of $M \times \mathbb{R}^m$. Since $2n - 1 < m - 1$, it follows from the easy case of Sard that the complement of the image $\tau(S)$ in S^{m-1} is dense. Moreover, since S is compact, this complement is open. Hence we can choose v in both complements of τ and c guaranteeing that $f_v|M$ is an injective immersion. Since M is compact, $f_v|M$ is in fact an embedding as desired. $\square$

Theorem 2.25 in case M is compact and $k \geq 2$ follows now from Propositions 2.26 and 2.27.

CHAPTER 3

Tangent spaces and tensor calculus

1. Tangent spaces

We start by defining tangent vectors to submanifolds in $\mathbb{R}^n$.

DEFINITION 3.1. A vector $X \in \mathbb{R}^n$ is said to be *tangent* to a submanifold $S \subset \mathbb{R}^n$ at a point $p \in S$ if there exists $\varepsilon > 0$ and a C^1 arc $c\colon (-\varepsilon, \varepsilon) \to S$ with $c(0) = p$ and $c'(0) = X$.

PROPOSITION 3.2. *The set T_pS of all tangent vectors to an m-dimensional submanifold $S \subset \mathbb{R}^n$ at a fixed point $p \in S$ is a vector subspace of $\mathbb{R}^n$. Given a local parametrization $f_p\colon \Omega_p \subset \mathbb{R}^m \to S$ of S, one has $T_pS = d_a f(\mathbb{R}^m)$ with $a := f_p^{-1}(p)$.*

PROOF. Given $f_p\colon \Omega_p \to \mathbb{R}^n$ as above, C^1 arcs $c\colon (-\varepsilon, \varepsilon) \to S$ with $c(0) = p$ are in one-to-one correspondence with C^1 maps $\widetilde{c}\colon (-\varepsilon, \varepsilon) \to \Omega_p$ with $\widetilde{c}(0) = a$ (as a consequence of the Inverse Function Theorem). The correspondence is defined by composing with f_p: $c = f_p \circ \widetilde{c}$. Then $c'(0) = d_a f_p(\widetilde{c}'(0))$ completes the proof since $d_a f_p$ is a linear map. □

The various characterizations of submanifolds of $\mathbb{R}^n$ (Theorem 1.41) lead to equivalent characterizations of tangent spaces:

THEOREM 3.3. *Let $S \subset \mathbb{R}^n$ be an m-dimensional C^k submanifold and let $g_i\colon V \to \mathbb{R}$, $i = 1, \dots, n-m$, be local defining functions of S in a neighborhood V of p with linear independent gradients as in Theorem 1.41 (iii). Then X is tangent to S at p if and only if $d_p g_i(X) = 0$ for all i.*

We now give three equivalent definitions of a tangent vector to an abstract C^k manifold M ($1 \le k \le \infty$) at a point $p \in M$.

DEFINITION 3.4 (**Geometric Definition of tangent vectors as classes of equivalent curves**). A *tangent vector* at p is an equivalence class of C^k arcs $c\colon (-\varepsilon, \varepsilon) \subset \mathbb{R} \to M$, $\varepsilon > 0$, with $c(0) = p$ where two arcs

$$c_1\colon (-\varepsilon_1, \varepsilon_1) \to M \text{ and } c_2\colon (-\varepsilon_2, \varepsilon_2) \to M$$

are called *equivalent* if $(\varphi \circ c_1)'(0) = (\varphi \circ c_2)'(0)$ for some (and hence for any) chart (U, φ) on M with $U \ni p$. The tangent space T_pM of M at p is the set of all tangent vectors at p.

Briefly: Tangent vectors are tangents to arcs lying on the manifold. "Unfortunately" there is no privileged representative arc for a given tangent vector.

EXAMPLE 3.5. If $U \subset \mathbb{R}^n$ is an open set, a curve $c\colon (-\varepsilon, \varepsilon) \subset \mathbb{R} \to U$ with $c(0) = p$ is equivalent to the affine curve $\lambda \mapsto p + c'(0)\lambda$. Then $c \mapsto c'(0)$ defines a one-to-one correspondence between the equivalence classes of curves and the space $\mathbb{R}^n$. Hence the tangent space T_pU can be identified with $\mathbb{R}^n$.

Given a tangent vector $X \in T_pM$ and a C^k function $f\colon U \to \mathbb{R}$ defined in an open neighborhood U of p in M, we can define the *directional derivative of f in the direction of X* by

$$D_X f = Xf := (f \circ c)'(0), \tag{1.1}$$

where c is any curve representing X. It is easy to see that the right-hand side of (1.1) does not depend on the particular representative c of X. Furthermore, $D_X f$ does not change if f is replaced by its restriction to a smaller neighborhood of p. In other words, $D_X f = D_X g$ if f and g coincide in a neighborhood of p. This motivates the following

Definition 3.6. A *germ* at p of a C^k function on M is an equivalence class of C^k functions defined in open neighborhoods of p such that two such functions $f_1\colon U_1 \to \mathbb{R}$ and $f_2\colon U_2 \to \mathbb{R}$ with $p \in U_1 \cap U_2$ are equivalent if their restrictions $f_1|_U$ and $f_2|_U$ coincide for some smaller neighborhood U of p, $U \subset U_1 \cap U_2$.

Denote by $\mathcal{F}_p(M)$ the set of all germs of C^k functions at p. Unlike the problem to add or multiply functions with different domains of definitions, there is no problem to add or multiply germs of functions. In particular, $\mathcal{F}_p(M)$ has a natural structure of an $\mathbb{R}$-algebra. We now have the following equivalent definition of tangent vectors in the C^∞ case.

Definition 3.7 (**Algebraic Definition of tangent vectors as derivations**). A tangent vector X on a C^∞ manifold M at a point $p \in M$ is a *derivation* on the set $\mathcal{F}_p(M)$, i.e a map $X\colon \mathcal{F}_p(M) \to \mathbb{R}$ with following properties:

(1) linearity: $X(af + bg) = aXf + bXg$ for $a, b \in \mathbb{R}$, $f, g \in \mathcal{F}_p$;
(2) Leibnitz rule: $X(fg) = (Xf)g + f(Xg)$ for $f, g \in \mathcal{F}_p$.

Briefly: tangent vectors are derivations acting on germs of scalar functions.

In view of formula (1.1), every tangent vector X in the sense of Definition 3.4 defines a tangent vector $X = D_X$ in the sense of Definition 3.7. Vice versa, we show that any derivation as in Definition 3.7 arises in this way in the C^∞ case. The proof is based on the following lemma:

Lemma 3.8. *Let $B^n \subset \mathbb{R}^n$ be the open unit ball and $f\colon B^n \to \mathbb{R}$ a C^∞ function. Then there exist C^∞ functions $f_1, \ldots, f_n\colon B^n \to \mathbb{R}$ such that*

$$f(x) = f(0) + x_1 f_1(x) + \cdots + x_n f_n(x). \tag{1.2}$$

Proof. Since

$$f(x) - f(0) = \int_0^1 \frac{d}{dt} f(tx_1, \ldots, tx_n)\, dt = \sum_{i=1}^n x_i \int_0^1 \frac{df}{dx_i}(tx_1, \ldots, tx_n)\, dt,$$

it remains to set $f_i(x) := \int_0^1 \frac{df}{dx_i}(tx_1, \ldots, tx_n)\, dt$, $i = 1, \ldots, n$. □

Note that if f is C^k, f_i is only C^{k-1}, hence we may not find functions in the same class unless $k = \infty$.

Given a derivation X in the sense of Definition 3.7 and a local system of coordinates (or chart) $(x^1, \dots, x^n)$ on M at p, vanishing at p, we obtain a tangent vector in the sense of Defintion 3.4 as follows. Denote by $\xi^i \in \mathbb{R}$ the value of X on the germ defined by the coordinate function x^i. Then $Xf = (f \circ c)'(0)$ with $c(t) := (\xi^1, \dots, \xi^n)t$. Indeed, X vanished on constant functions (as follows from the Leibnitz rule) and hence, in view of formula (1.2),

$$Xf = \sum_i \xi^i f_i(0) = \sum_i \xi^i \frac{df}{dx^i}(0) = (f \circ c)'(0).$$

A basis of the tangent space to M at p is given in local coordinates $(x^1, \dots, x^n)$ by

$$\frac{\partial}{\partial x^1}(p), \dots, \frac{\partial}{\partial x^n}(p),$$

where $\partial/\partial x^i$ is the derivation given by $f \mapsto \partial f/\partial x^i$. Other notation is also used frequently:

$$\left(\frac{\partial}{\partial x^i}\right)_p \quad \text{or} \quad \frac{\partial}{\partial x^i}\Big|_p$$

Finally we have the following analytic (also called "physical") definition:

Definition 3.9 (**Analytic Definition of tangent vectors via coordinates and their transformation rules**). A tangent vector ξ at the point p is an association to every coordinate chart $(x^1, \dots, x^n)$ at p an n-tuple $(\xi^1, \dots, \xi^n)$ of real numbers is such a way that, if $(\widetilde{\xi}^1, \dots, \widetilde{\xi}^n)$ is associated with another coordinate system $(\widetilde{x}^1, \dots, \widetilde{x}^n)$, then it satisfies the transformation rule

$$\widetilde{\xi}^i = \frac{\partial \widetilde{x}^i}{\partial x^j}(p)\xi^j, \tag{1.3}$$

where there is a summation over j (Einstein's convention).

Briefly: tangent vectors are represented by elements of $\mathbb{R}^n$ for each coordinate chart transforming via differentials of a coordinate change at the reference point.

Given a tangent vector represented by a curve c as in Definition 3.4, it is easy to define an association in the sense of Definition 3.9 by taking $(\xi^1, \dots, \xi^n)$ to be the derivative of c at $t = 0$ computed with respect to the given coordinate chart. Vice versa, given $(\xi^1, \dots, \xi^n)$, a representing curve c can be chosen $t \mapsto (\xi^1, \dots, \xi^n)t$. A direct equivalence with Definition 3.7 is also easy to obtain by letting $\xi^1, \dots, \xi^n$ to be the values of a given derivation on the (germs of) coordinate functions $x^1, \dots, x^n$ respectively. In other words, we have

$$\xi = \xi^i \frac{\partial}{\partial x^i}(p). \tag{1.4}$$

Definition 3.10. Let $f\colon M \to M'$ be a C^k map. The *differential of f* or *tangent map to f* at a point $p \in M$ is the map $d_p f\colon T_x M \to T_{f(x)}M'$ (very often denoted by df_p or $(df)_p$) that sends an equivalence class defined by a curve $c\colon (-\varepsilon, \varepsilon) \to M$ to the equivalence class defined by $(f \circ c)\colon (-\varepsilon, \varepsilon) \to M'$.

If $(x^1, \ldots, x^n)$ and $(y^1, \ldots, y^m)$ are local coordinates on M and M' at p and $p' := f(p)$ respectively, we write $f = (f^1, \ldots, f^m)$ and have the formula

$$d_p f \left(\frac{\partial}{\partial x^j}(p) \right) = \frac{\partial f^i}{\partial x^j}(p) \frac{\partial}{\partial y^i} \tag{1.5}$$

2. Vector fields and Lie brackets

DEFINITION 3.11. The *tangent bundle* TM of an n-manifold M is the disjoint union of all tangent spaces T_xM for $x \in M$ with the natural structure of a $2n$-manifold. A *vector field* X on a manifold M is a correspondence that associates to each point $p \in M$ a vector $X(p) \in T_pM$, i.e. a mapping $X\colon M \to TM$. The field is differentiable if this mapping is differentiable.

Denote by $C^\infty(M)$ the algebra of C^∞ functions on M. Then, in view of the algebraic definition of a tangent space in the previous section, a vector field X can be identified with a *derivation* of $C^\infty(M)$ (or an *operator acting there*), i.e. a map

$$f \in C^\infty(M) \mapsto Xf \in C^\infty(M)$$

satisfying both linearity and the Leibnitz rule as in Defintion 3.7.

It makes sense to iterate the above operators. If X and Y are two vector fields, $X \circ Y$ and $Y \circ X$ are (in general different) operators involving higher order derivatives. However the difference of them turns out to be a new vector field called their Lie bracket $[X, Y]$:

$$X, Y]f = X(Yf) - Y(Xf). \tag{2.1}$$

LEMMA 3.12. *There exists a unique vector field Z such that, for all $f \in C^\infty(M)$, $Zf = X(Yf) - Y(Xf)$.*

PROOF. Write $X = \sum_i a^i \frac{\partial}{\partial x^i}$ and $Y = \sum_j b^j \frac{\partial}{\partial x^j}$, where a^i and b^i are smooth functions on M. Then for all f,

$$X(Yf) - Y(Xf) = X(\sum_j b^j \frac{\partial f}{\partial x^j}) - Y(\sum_i a^i \frac{\partial f}{\partial x^i}) = \sum_{ij} (a^i \frac{\partial b^j}{\partial x^i} - b^i \frac{\partial a^j}{\partial x^i}) \frac{\partial f}{\partial x^j},$$

where we have cancellation for terms involving second order derivatives of f. Then the formula $Z = \sum_{ij}(a^i \frac{\partial b^j}{\partial x^i} - b^i \frac{\partial a^j}{\partial x^i})$ shows both existence and uniqueness. □

The basic properties of the Lie brackets are bilinearity, anticommutativity and the Jacobi identity:

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

DEFINITION 3.13. A *trajectory* or *integral curve* of a vector field X on a manifold M is a curve (arc) γ from an open interval $(a, b) \subset \mathbb{R}$ into M safisfying the equation

$$\frac{d\gamma}{dt}(t_0) = X(\gamma(t_0))$$

for every $t_0 \in (a, b)$.

It follows from the local theory of Ordinary Differential Equations that a vector field (at least of class C^1) locally has a unique integral curve γ_p defined over an interval (a, b) containing 0 and passing through any given point $p \in M$ at $t = 0$. The uniqueness means that any two such curves coincide on the intersection of their intervals of definition. There is a unique choice of the maximal possible intervale. Recall that the manifold M is assumed Hausdorff which is important for the mentioned uniqueness.

It is a further consequence of the local theory of Ordinary Differential Equations that the integral curve γ_p also smoothly depends on p, i.e. the map $\gamma(t,p) := \gamma_p(t)$ is a smooth function in $(t,p) \in \mathbb{R} \times M$. The smoothness of γ is the same as that of the vector field. If for each p, the interval for γ_p is chosen to be the maximal one, the domain of definition of γ becomes an open subset of $\mathbb{R} \times M$. The map γ is called the *local flow* of the vector field X.

Definition 3.14. A vector field on M is called *complete* if its flow is defined on the whole $\mathbb{R} \times M$ or, equivalently, for every $p \in M$, the integral curve γ_p is defined on $\mathbb{R}$.

Proposition 3.15. *If M is compact, every vector field is complete.*

If X is a complete vector field, by fixing $t \in \mathbb{R}$ we obtain a self-diffeomorphism $\gamma(t,\cdot)\colon M \to M$ because its inverse is $\gamma(-t,\cdot)$. In general we have

$$\gamma(t_1, \gamma(t_2, p)) = \gamma(t_1 + t_2, p))$$

justifying the name *one-parameter group of diffeomorphisms.*

In general, the local flow $\gamma(t,x)$ of a vector field X can be used to calculate Xf for a smooth function f (exercise):

$$Xf(\cdot) = \frac{d}{dt} f(\gamma(t,\cdot)). \tag{2.2}$$

Let $\varphi(s,y)$ be the (local) flow of another vector field Y. Iterating (2.2) we obtain

$$[X,Y]f(\cdot) = X(Yf)(\cdot) - Y(Xf)(\cdot) = \frac{d^2}{dt\,ds}\Big|_{t=s=0} \Big(f(\gamma(t,\varphi(s,\cdot))) - f(\varphi(s,\gamma(t,\cdot))) \Big), \tag{2.3}$$

from where we see that, if the flows of X and Y commute (i.e. $\gamma(t,\varphi(s,\cdot)) \equiv \varphi(s,\gamma(t,\cdot))$), then $[X,Y] \equiv 0$. The converse is also true (here without proof):

Theorem 3.16. *Given vector fields X and Y on a manifold M, their flows commute if and only if $[X,Y] \equiv 0$.*

3. Frobenius Theorem

The problem of finding integral curves for single vector fields naturally extends to the more general problem of finding integral surfaces for systems of vector fields.

Definition 3.17. Given a system of vector fields $X_1, \ldots, X_d$, on M, linearly independent at every point, an *integral submanifold* S is a submanifold of M such that for every $p \in S$, the tangent subspace $T_pS \subset T_pM$ is spanned by $X_1(p), \ldots, X_d(p)$.

In contrast to the integral curves of single vector fields, an integral submanifold may not exist in general. A necessary condition comes from the following property:

Lemma 3.18. *If $S \subset M$ is a submanifold and two vector fields X and Y on M are tangent to S, then their Lie bracket is also tangent to S.*

Here X is called tangent to S if $X(p) \in T_pS$ for all $p \in S$.

Proof. By definition, $Xf(p) = \frac{d}{dt}|_{t=0} f(c(t))$, where $c\colon (-\varepsilon, \varepsilon) \to M$ is a curve representing $X(p) \in T_pM$. If $p \in S$ and X is tangent to S, the curve c can be chosen in S. Then

$$(Xf)|_S = (X|_S)(f|_S), \tag{3.1}$$

i.e. differentiating f along X and restricting to S yields the same result as differentiating the restriction of f along the restriction of X. Then the Lie bracket (2.1) can be computed separately for X, Y on M and for their restrictions to S. It follows from (3.1) that both brackets are equal when restricted to S and hence are tangent to S. □

Hence, if S is an integral submanifold of the system $X_1, \ldots, X_d$, and $p \in S$, all Lie brackets $[X_i, X_j](p)$ must be contained in the span of $X_1(p), \ldots, X_d(p)$. This condition turns out to be also sufficient if assumed at every point:

Theorem 3.19 (Frobenius). *Let $X_1, \ldots, X_d$, be vector fields on a manifold M, linearly independent at every point. A necessary and sufficient condition for the existence, for every $p \in M$, of an integral submanifold $S \subset M$ passing through p is that $[X_i, X_j](p)$ belongs to the span of $X_1(p), \ldots, X_d(p)$ for every $p \in M$.*

4. Lie groups and Lie algebras

Lie groups are important example of differentiable manifolds.

DEFINITION 3.20. A *Lie group* is a set G which is both a C^k manifold and a group such that the maps

$$G \times G \to G,\ (x,y) \mapsto xy, \quad G \to G,\ x \mapsto x^{-1}$$

are of class C^k.

EXERCISE 3.21. Show that for every $g \in G$, the left and right translations

$$L_g, R_g \colon G \to G, \quad L_g(h) := gh, \quad R_g(h) := hg$$

are C^k diffeomorphisms.

REMARK 3.22. We mention without proof that a Lie group of class C^1 has an unique C^∞ (even real-analytic) differentiable structure compatible with the group operation in the sense of Definition 3.20. In particular, we may always assume a Lie group to be C^∞.

EXAMPLES 3.23. (1) Any finite-dimensional real vector space is a Lie group with respect to addition. (2) For $\mathbb{K} = \mathbb{R}$ (real numbers), $\mathbb{C}$ (complex numbers) or $\mathbb{H}$ (quaternions), set $\mathbb{K}^* := \mathbb{K} \setminus \{0\}$. Then $\mathbb{K}^*$ is a Lie group with respect to multiplication.

The most important Lie groups, called *classical Lie groups*, are $GL(n,\mathbb{R})$, $GL(n,\mathbb{C})$, $O(n)$, $U(n)$ and their corresponding subgroups $SL(n,\mathbb{R})$, $SL(n,\mathbb{C})$, $SO(n)$, $SU(n)$, defined as follows. We write $\mathbb{R}^{n\times n}$ (resp. $\mathbb{C}^{n\times n}$) for the sets of all $n \times n$ matrices with real (resp. complex) entries. For a matrix $A \in \mathbb{C}^{n\times n}$, denote by A^t the transpose of A and by $\bar{A}$ its conjugate. Then

$$\begin{aligned}
GL(n,\mathbb{R}) &:= \{A \in \mathbb{R}^{n\times n} : \det A \neq 0\},\\
GL(n,\mathbb{C}) &:= \{A \in \mathbb{C}^{n\times n} : \det A \neq 0\},\\
O(n) &:= \{A \in \mathbb{R}^{n\times n} : AA^t = \mathsf{id}\},\\
U(n) &:= \{A \in \mathbb{C}^{n\times n} : A\bar{A}^t = \mathsf{id}\},\\
SL(n,\mathbb{R}) &:= \{A \in \mathbb{R}^{n\times n} : \det A = 1\},\\
SL(n,\mathbb{C}) &:= \{A \in \mathbb{C}^{n\times n} : \det A = 1\},\\
SO(n) &:= O(n) \cap SL(n,\mathbb{R}),\\
SU(n) &:= U(n) \cap SL(n,\mathbb{C}).
\end{aligned}$$

The groups $GL(n,\mathbb{R})$ and $GL(n,\mathbb{C})$ are open (and dense) subsets of the matrix spaces $\mathbb{R}^{n\times n}$ and $\mathbb{C}^{n\times n}$ respectively. The other groups are submanifolds. To see this it is sufficient to check the submanifold property near the identity matrix id and use left (or right) multiplications to check it near other points.

For every Lie group G there exists unique associated finite-dimensional *Lie algebra* $\mathfrak{g}$ that as a vector space can be identified with the tangent space T_eG, where $e \in G$ is the unit of the group. In order to define an algebra operation (called the *commutator*) on $\mathfrak{g}$, identify a vector $\xi \in T_eG$

with the unique left-invariant vector field X on G with $X(e) = \xi$. A vector field X on G is called *left-invariant* if

$$d_a L_g(X(a)) = X(ga).$$

for all $a, g \in G$, where $L_g(x) := gx$ is the left translation on G. The the usual Lie bracket induces a Lie algebra structure on the (finite-dimensional) space of all left-invariant vector fields. In the light of the above identification of $\mathfrak{g}$ with left-invariant vector fields, this construction yields a Lie algebra structure on $\mathfrak{g}$ which is the Lie algebra structure induced by the group operation of G. (Right-invariant vector fields would yield an isomorphic Lie algebra).

The Lie algebras corresponding to the above classical Lie groups are the classical matrix Lie algebras $gl(n,\mathbb{R})$, $gl(n,\mathbb{C})$, $o(n)$, $u(n)$ and their corresponding subalgebras $sl(n,\mathbb{R})$, $sl(n,\mathbb{C})$, $so(n)$, $su(n)$:

$$\begin{aligned}
&gl(n,\mathbb{R}) := \mathbb{R}^{n\times n},\\
&gl(n,\mathbb{C}) := \mathbb{C}^{n\times n},\\
&o(n) := \{a \in \mathbb{R}^{n\times n} : a + a^t = 0\},\\
&u(n) := \{a \in \mathbb{C}^{n\times n} : a + \bar{a}^t = 0\},\\
&sl(n,\mathbb{R}) := \{a \in \mathbb{R}^{n\times n} : \operatorname{tr} a = 0\},\\
&sl(n,\mathbb{C}) := \{a \in \mathbb{C}^{n\times n} : \operatorname{tr} a = 0\},\\
&so(n) := o(n) \cap sl(n,\mathbb{R}),\\
&su(n) := u(n) \cap sl(n,\mathbb{C}),
\end{aligned}$$

where “tr” stands for the trace of the matrix, i.e. the sum of all diagonal entries.

For every matrix group, its Lie algebra consists of matrices tangent to the group at the identity matrix and their commutator is given by the formula $[a, b] = ab - ba$.

5. Tensors and differential forms

A *tensor* Θ of type (m, n) on a manifold M is a correspondence that associates to each point $p \in M$ a multilinear map

$$\Theta_p \colon T_pM \times \cdots \times T_pM \times T_p^*M \times \cdots \times T_p^*M \to \mathbb{R},$$

where the tangent space T_pM appears m times and the cotangent space T_p^*M (the dual of the tangent) appears n times. For instance, a vector field X can be seen as a collection of linear maps $X_p \colon T_pM \to \mathbb{R}$.

An (exterior) *m-form* is a tensor ω of type $(m, 0)$ which is skew symmetric in its arguments, i.e.

$$\omega_p(\xi_{\sigma(1)}, \ldots, \xi_{\sigma(m)}) = (-1)^\sigma \omega_p(\xi_1, \ldots, \xi_m)$$

for any tangent vectors $\xi_1, \ldots, \xi_m \in T_pM$ and any permutation $\sigma \in S_m$. In any local coordinates $(x^1, \ldots, x^m)$ it can be expressed as

$$\omega = \omega_{i_1,\ldots,i_m} dx^{i_1} \wedge \cdots \wedge dx^{i_m} = \omega_I dx^I,$$

where the summation is taken over all sets of indices $1 \le i_1 < \cdots < i_m \le n$ and $I = (i_1, \ldots, i_m)$, $\omega_I = \omega_{i_1,\ldots,i_m}$, $dx^I = dx^{i_1} \wedge \cdots \wedge dx^{i_m}$. Here ω_I is a function of the reference point and ω is said to be smooth (or differentiable) if each ω_I is smooth. The latter notion is independent of the choice of coordinates.

The *wedge product* of an m-form ω and an m'-form ω' is an $(m + m')$-form which is defined in coordinates by

$$(\omega_I dx^I) \wedge (\omega'_{I'} dx^{I'}) = \omega_I \omega'_{I'} dx^I \wedge dx^{I'},$$

i.e. the coefficients are multiplied as functions and the wedge products of differentials are written together with a wedge inbetween. A coordinate-free definition can be also given:

$$(\omega \wedge \omega')_p(\xi_1, \ldots, \xi_{m+m'}) = \sum_{\sigma \in S_{m+m'}} \frac{(-1)^\sigma}{m!\, m'!} \omega_p(\xi_{\sigma(1)}, \ldots, \xi_{\sigma(m)})\, \omega'_p(\xi_{\sigma(m+1)}, \ldots, \xi_{\sigma(m+m')}),\ p \in M.$$

Another important operation with forms is their pullback. Let ω be an m-form on M and $\varphi \colon N \to M$ be any differentiable map from another manifold N to M. Then the *pullback* $\varphi^*\omega$ is an m-form on N defined by

$$(\varphi^*\omega)_p(\xi_1, \ldots, \xi_m) = \omega_{\varphi(p)}(\varphi_*\xi_1, \ldots, \varphi_*\xi_m),\ p \in N,$$

for $\xi_1, \ldots, \xi_m \in T_pN$, where $\varphi_*\xi = d_p\varphi(\xi) \in T_{\varphi(p)}M$ is the pushforward of $\xi \in T_pN$.

6. Orientation and integration of differential forms

DEFINITION 3.24. A manifold M is *orientable* if it has an atlas $(U_\alpha, \varphi_\alpha)$ such that, if $(x^1, \ldots, x^n)$ and $(x^{1'}, \ldots, x^{n'})$ are any coordinate systems defined by charts in this atlas, then the Jacobian determinant $\det(\frac{\partial x^i}{\partial x^{i'}})$ is positive. An *orientation* of M is a choice of such an atlas. Any other chart that can be added to this atlas without changing the above property is called positively oriented with respect to the given orientation. If a manifold is orientable, it can be oriented in precisely two different ways.

REMARK 3.25. It can be shown that a compact surface (2-dimensional submanifold) in $\mathbb{R}^3$ is always orientable.

The integral of an n-form ω on an orientable n-manifold M can be defined in three steps.

In the first step assume that $M = U \subset \mathbb{R}^n$ is an open set and the support of ω is compact. Writing $\omega = f(x)dx^1 \wedge \cdots \wedge dx^n$, we set

$$\int_U \omega = \int_U f(x)\, dx^1 \cdots dx^n,$$

where the right-hand side is the usual multiple integral in $\mathbb{R}^n$.

In the second step assume that ω has compact support in an open set $U \subset M$ where a chart $\varphi\colon U \to \Omega \subset \mathbb{R}^n$ is defined and positively oriented. Then we set

$$\int_M \omega = \int_U \omega = \int_\Omega (\varphi^{-1})^* \omega,$$

where the right-hand side is an integral on an open set in $\mathbb{R}^n$ and is hence defined in the first step. An important part of the justification of this definition involves checking that the integral is in fact independent of the choice of the coordinate system which is based on the transformation formula

$$\int_{\psi(U)} f(x)\, dx^1 \cdots dx^n = \int_U f(x(u))\, \det\left(\frac{\partial x^i}{\partial u^j}\right) du^1 \cdots du^n$$

which holds for any orientation preserving diffeomorphism $\psi\colon U \to \psi(U)$ between open sets in $\mathbb{R}^n$.

Finally, in the third step, consider a covering of M by positively oriented coordinate charts $(U_\alpha, \varphi_\alpha)$ and choose a partition of unity $(f_i)_{i\in I}$ subordinate to the covering (U_α) (see Section 2 of Chapter 2). Then, for any i, the product $\omega_i = f_i\omega$ is a new n-form on M with compact support in some U_α and hence its integral over M is defined. We can now set

$$\int_M \omega = \sum_i \int_M \omega_i$$

assuming the sum is defined, i.e. it is either finite or an absolutely convergent series. The absolute convergence is important because there is no distinguished order for the values of $i \in I$.

7. The exterior derivative and Stokes Theorem

The *exterior derivative* of an m-form on M is an $(m+1)$-form on M defined in local cooordinates by

$$d\omega = d(\omega_I dx^I) = (d\omega_I) \wedge dx^I,$$

where $d\omega_I$ is the differential of the function ω_I (seen as a 1-form).

A more geometric definition can be obtained from the formula

$$d\omega(\xi_1, \dots, \xi_{m+1}) = \lim_{h \to 0} \frac{1}{h^{m+1}} \int_{\partial P(h\xi_1, \dots, h\xi_{m+1})} \omega,$$

where $\partial P(h\xi_1, \dots, h\xi_{m+1})$ is the oriented boundary of the parallelorgram spanned by the vectors $h\xi_1, \dots, h\xi_{m+1}$ (with respect to some local coordinates). The parallelogram is obtained (an oriented) via the parametrization

$$(t^1, \dots, t^{m+1}) \in [0,1]^{m+1} \mapsto t^1 h\xi^1 + \cdots + t^{m+1} h\xi^{m+1} \in P(h\xi_1, \dots, h\xi_{m+1}).$$

Stokes theorem relates the integral of ω over the boundary of a manifold to the integral of $d\omega$ over the manifold itself. The standard setting here is that of a manifold with boundary. The definition of a *manifold M with boundary ∂M* is obtained from the definition of a manifold by allowing charts $\varphi_\alpha \colon U_\alpha \to \mathbb{R}^n$ to be homeomorphisms onto open sets in the closed half-space

$$H^n = \{(x^1, \dots, x^n) \in \mathbb{R}^n : x^n \geq 0\}.$$

The boundary ∂M consists precisely of the points of U_α (for some α) that correspond (under φ_α) to boundary points of H^n.

The following is easy to see (an exercise):

LEMMA 3.26. *Given an atlas $(U_\alpha, \varphi_\alpha)$ as above, the collection $(V_\alpha, \varphi_\alpha|_{V_\alpha})$, where $V_\alpha := \varphi_\alpha^{-1}(\{x^n = 0\})$, defines an $(n-1)$-manifold structure on the boundary ∂M. If $(U_\alpha, \varphi_\alpha)$ defines an orientation on M, the corresponding atlas $(V_\alpha, \varphi_\alpha|_{V_\alpha})$ also defines an orientation on ∂M that is said to be induced by the orientation of M.*

THEOREM 3.27 (Stokes theorem). *Let M be an n-manifold with boundary ∂M and ω be a differentiable $(n-1)$-form with compact support on M. Then*

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

CHAPTER 4

Riemannian geometry

1. Riemannian metric on a manifold

DEFINITION 4.1. A *(smooth) Riemannian metric* on a manifold M is an association to every $p \in M$ a symmetric positive definite bilinear form $g_p \colon T_pM \times T_pM \to \mathbb{R}$ (hence an inner product) such that in every local coordinates $(x^1, \dots, x^n)$, g_p is given by

$$g_p(a^i \frac{\partial}{\partial x^i}, b^j \frac{\partial}{\partial x^j}) = g_{ij}(p) a^i b^j \tag{1.1}$$

with smooth coefficients $g_{ij}(p)$. The pair (M, g) is called *Riemannian manifold.*

Formula (1.1) is often rewritten as $g = g_{ij} dx^i \otimes dx^j$. (Note that g is not a 2-form because it is symmetric in contrast to skew-symmetric 2-forms.)

EXAMPLE 4.2. The standard metric on $\mathbb{R}^n$ is defined by setting $g_{ij} = \delta_{ij}$, i.e.

$$g(a^i \frac{\partial}{\partial x^i}, b^j \frac{\partial}{\partial x^j}) = \sum_i a^i b^i \tag{1.2}$$

which is the standard inner product (or scalar product) on $\mathbb{R}^n$. More generally, if M is a submanifold in $\mathbb{R}^n$, a metric g on M can be obtained by restricting the standard metric (1.2) to $T_pM \times T_pM$ for every $p \in M$. This induced metric is classically called the *first fundamental form of* M.

Given a metric g on M, the *norm (length)* of a tangent vector $\xi \in T_pM$ is given by $\|\xi\| := \sqrt{g_p(\xi, \xi)}$ and the *angle* between two tangen vectors $\xi, \eta \in T_pM$ is given by

$$\alpha = \cos^{-1} \left(\frac{g(\xi, \eta)}{\|\xi\| \cdot \|\eta\|} \right).$$

Furthermore, the length of a smooth arc $c \colon [a, b] \to M$ is given by the integral $\int_a^b \|c'(t)\| \, dt$, where the norm of the tangent vector $c'(t) \in T_{c(t)}M$ is calculated as explained above. In fact, it follows from the transformation rule that the length of the arc c coincides with the length of any reparametrized arc $\widetilde{c} = c \circ \varphi$ with $\varphi \colon [\widetilde{a}, \widetilde{b}] \to [a, b]$ being a diffeomorphism. Hence one can talk about the length of a curve (which is an equivalence class of arcs).

Using length of curves one can define the distance $d(p, q)$ between two points $p, q \in M$ to be the infimum of lengths of curves connecting p and q. Without proof we quote here:

PROPOSITION 4.3. *The distance d associated to a Riemannian metric g satisfies metric space axioms. The obtained metric induces the same topology as the manifold topology.*

Given a metric on an **oriented** manifold, one can also compute volumes (measures) of subsets using the *volume form* which is defined in oriented local coordinates by

$$\omega := \sqrt{\det(g_{ij})}\, dx^1 \wedge \cdots \wedge dx^n. \tag{1.3}$$

It is a direct consequence of the transformation rule of g_{ij} and of the wedge product that the above form ω is well-defined and independent of the choice of oriented local coordinates.

The following general statement is a consequence of the existence of partitions of unity:

PROPOSITION 4.4. *On every manifold M there exists a Riemannian metric g.*

2. The Levi-Civita connection

A connection on a manifold is an additional structure permitting to differentiate vector fields in directions of tangent vectors.

Definition 4.5. An *(linear) connection* ∇ on a manifold M associates to every vector field X on M and every tangent vector $\xi \in T_pM$ another tangent vector $\nabla_\xi X \in T_pM$, called the *covariant derivative* of X in the direction of ξ, such that $\nabla_\xi X$ is bilinear in ξ and X and satisfies the following Leibniz rule:

$$\nabla_\xi(fX) = df(\xi)X + f\nabla_\xi X. \tag{2.1}$$

Choosing local coordinates $(x^1, \dots, x^n)$ on M and writing $\xi = \xi^i \frac{\partial}{\partial x^i}$ and $X = X^j \frac{\partial}{\partial x^j}$ we obtain, using the the linearity and (2.1):

$$\nabla_\xi X = \xi^i \frac{\partial X^j}{\partial x^i}\frac{\partial}{\partial x^j} + \xi^i X^j \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \xi^i \frac{\partial X^j}{\partial x^i}\frac{\partial}{\partial x^j} + \Gamma^k_{ij}\xi^i X^j \frac{\partial}{\partial x^k}, \tag{2.2}$$

where $\Gamma^k_{ij}(p)$ are coefficients of the vector $\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}$:

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \Gamma^k_{ij}(p)\frac{\partial}{\partial x^k}. \tag{2.3}$$

The connection ∇ is said to be smooth if its coefficients Γ^k_{ij}, called *Christoffel symbols* are smooth functions with respect to any choice of local coordinates. Given any connection ∇ and any choice of local coordinates, the functions $\Gamma^k_{ij}(p)$ are uniquely determined by (2.3). Vice versa, given any collection of functions $\Gamma^k_{ij}(p)$, a connection in the given coordinate chart can be defined by (2.2) and is uniquely determined by that formula.

An important geometric interpretation of a connection is that of a *parallel transport* of tangent vectors along curves. Given a differentiable (or piecewise differentiable) curve $c\colon [a, b] \to M$, a vector field X on M is called *parallel* along c if $\nabla_{c'(t)}X = 0$ for all t. It follows from the theory of ordinary differential equations that, given a curve c and a vector $X_0 \in T_{c(t_0)}$, there exists a parallel vector field X near $c(t_0)$ with $X(c(t_0)) = X_0$, which is unique along c.

Definition 4.6. Given a connection ∇, its torsion *torsion* is given by

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

for any vector fields X and Y. A connection is called *symmetric* or *torsion free* if $T(X, Y) \equiv 0$ for all X and Y.

A computation in local coordinates $(x^1, \dots, x^n)$ gives

$$T\left(X^i \frac{\partial}{\partial x^i}, Y^j \frac{\partial}{\partial x^j}\right) = X^i Y^j (\Gamma^k_{ij} - \Gamma^k_{ji})\frac{\partial}{\partial x^k}, \tag{2.4}$$

showing that the value $T(X, Y)(p) \in T_pM$ depends only on the values of $X(p), Y(p) \in T_pM$ and hence defines, for every p, a bilinear skew-symmetric map $T_pM \times T_pM \to T_pM$. It follows that ∇ is symmetric if and only if $\Gamma^k_{ij} = \Gamma^k_{ji}$ explaining the name.

For connections on a Riemannian manifold, there is a natural compatibility condition.

Definition 4.7. A connection ∇ on a Riemannian manifold (M, g) is called *metric* if

$$D_\xi g(X,Y) = g(\nabla_\xi X, Y) + g(X, \nabla_\xi Y) \tag{2.5}$$

for any tangent vector ξ and vector fields X and Y, where on the left-hand side D_ξ denotes the directional derivative of a smooth function.

A connection is metric if and only if for any parallel vector fields X, Y along any curve, $g(X,Y)$ is constant along that curve.

In local coordinates, choosing $X = \partial/\partial x^i$, $Y = \partial/\partial x^j$, $\xi = \partial/\partial x^k$, we can rewrite (2.5) as

$$D_k g_{ij} = g_{jl}\Gamma^l_{ki} + g_{il}\Gamma^l_{kj}. \tag{2.6}$$

Theorem 4.8 (Levi-Civita). *On a Riemannian manifold there exists an unique connection that is both symmetric and metric, called the* Levi-Civita connection*.*

Proof. Assuming ∇ is a Levi-Civita connection and permuting indices in (2.6) we have

$$D_k g_{ij} = g_{jl}\Gamma^l_{ki} + g_{il}\Gamma^l_{kj} \tag{2.7}$$

$$D_i g_{jk} = g_{kl}\Gamma^l_{ij} + g_{jl}\Gamma^l_{ik} \tag{2.8}$$

$$D_j g_{ki} = g_{il}\Gamma^l_{jk} + g_{kl}\Gamma^l_{ji}. \tag{2.9}$$

Adding the first two identities, subtracting the third and using the symmetry $\Gamma^l_{ij} = \Gamma^l_{ji}$ we obtain:

$$2g_{jl}\Gamma^l_{ki} = D_k g_{ij} + D_i g_{jk} - D_j g_{ki}. \tag{2.10}$$

Denoting by (g^{ij}) the inverse matrix of (g_{ij}) and solving (2.10) for Christoffel symbols, we obtain the classical formula

$$\Gamma^m_{ki} = \frac{g^{mj}}{2}(D_k g_{ij} + D_i g_{jk} - D_j g_{ki}) \tag{2.11}$$

showing the uniqueness. To show the existence, use (2.11) to define Γ^m_{ki} and hence ∇. Then ∇ is obviously symmetric and (2.5) can be directly verified. □

In the important case when M is a submanifold of $\mathbb{R}^n$ with its standard metric, the Levi-Civita connection ∇ on M can be obtained from the following general statement.

Proposition 4.9. *Let $(\widetilde{M}, \widetilde{g})$ be a Riemannian manifold and M a submanifold of $\widetilde{M}$ with induced metric. Denote by $\widetilde{\nabla}$ and ∇ the Levi-Civita connection of $\widetilde{M}$ and M respectively. Then, for $\xi \in T_pM$ and X tangent to M,*

$$\nabla_\xi X = \pi_p(\widetilde{\nabla}_\xi X), \tag{2.12}$$

where $\pi_p\colon T_p\widetilde{M} \to T_pM$ is the orthogonal projection.

Proof. One verifies directly that the connection given by the right-hand side of (2.12) satisfies all the assumptions of the Levi-Civita connection. The statements follows then by the uniqueness of the Levi-Civita connection. □

Denote now by $\pi_p^\perp \colon T_p\widetilde{M} \to T_p^\perp M$ the complementary orthogonal projection onto the orthogonal complement $T_p^\perp M$ of T_pM in $T_p\widetilde{M}$.

Definition 4.10. The *second fundamental form* of the submanifold $M \subset \widetilde{M}$ at $p \in M$ is given by

$$\mathrm{II}_p(X,Y) = \pi_p^\perp(\widetilde{\nabla}_X Y(p)), \tag{2.13}$$

where X and Y are vector fields on $\widetilde{M}$ tangent to M.

Lemma 4.11. *The second fundamental form $\mathrm{II}_p(X,Y)$ depends only on the values $X(p)$, $Y(p)$ and defines a symmetric bilinear map $\mathrm{II}_p \colon T_pM \times T_pM \to T_p^\perp M$.*

Proof. Since $\widetilde{\nabla}$ is torsion-free, one has

$$\widetilde{T}(X,Y) = \widetilde{\nabla}_X Y - \widetilde{\nabla}_Y X - [X,Y] = 0.$$

For X and Y tangent to M, their Lie bracket $[X,Y]$ is also tangent to M. Hence, projecting on $T_p^\perp M$, we obtain the symmetry $\mathrm{II}_p(X,Y) = \mathrm{II}_p(Y,X)$. The fact that, for a fixed Y, the value of $\mathrm{II}_p(X,Y)$ depends only on $X(p)$ follows directly from the definition (2.13). Now the symmetry implies that the roles of X and Y can be exchanged, completing the proof of the lemma. $\square$

Examples 4.12. If M is a surface in $\mathbb{R}^3$ (with the standard metric), $\mathrm{II}(X,Y)$ is the classical second fundamental form of M obtained by differentiating the coefficients of Y along X ($\widetilde{\nabla}_X T$) and projecting to the normal direction. If M is a curve in $\mathbb{R}^2$, its second fundamental form is classically called the curvature of the curve. More generally, if M is a curve in a 2-dimensional Riemannian manifold, its second fundamental form (more precisely, its value on the oriented unit tangent vector) is called the *geodesic curvature* of M.

3. Geodesics and the exponential map

A *geodesic* is a curve whose tangent vector is parallel with respect to the Levi-Civita connection:

Definition 4.13. A parametrized curve $c\colon (a,b)\subset\mathbb{R}\to M$ in a Riemannian manifold M is a *geodesic* if

$$\nabla_{c'}c' = 0, \tag{3.1}$$

i.e. the vector field c' (its local extension) is parallel along c.

The main fact about geodesics is given by the following statement.

Proposition 4.14. *For every $p\in M$ and $\xi\in T_pM$ there exists a unique geodesic $c_\xi\colon I\to M$ with $0\in I$, $c(0)=p$, $c'(0)=\xi$ and $I\subset\mathbb{R}$ maximal.*

Proof. In local coordinates $(x^1,\dots,x^n)$ the curve $c(t)$ is given by its coordinate functions $(x^1(t),\dots,x^n(t))$. Then, using (2.2), the equation (3.1) can be rewritten as

$$\left(\frac{d}{dt}\frac{dx^k}{dt}+\Gamma^k_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt}\right)\frac{\partial}{\partial x^k}=0 \tag{3.2}$$

which is a system of second order ODEs. The statement now follows from the general existence and uniqueness result for solutions of ODE systems. □

Examples 4.15. In $\mathbb{R}^n$ with the standard metric we have $\Gamma^k_{ij}=0$ and hence the geodesics are the straight lines parametrized by $t\mapsto at+b$ with $a,b\in\mathbb{R}^n$. On the sphere S the geodesics are the great circles parametrized by the arc length and by any scalar multiple of the arc length. Indeed, the derivative $c'(t)$ for such a curve c in this case is orthogonal to TS and hence the covariant derivative is 0 in view of (2.12).

We mention without proof the fundamental *local distance minimizing property* of geodesics.

Theorem 4.16. *For any $p\in M$ there exists a neighborhood U of p in M such that for any $x,y\in U$, there exists unique geodesic connecting x and y whose length is equal to the distance $d(x,y)$ (defined in Section 1) and thus not greater than the length of any other curve connecting x and y.*

Given p and ξ as in Proposition 4.14, denote by $c_{p,\xi}(t)$ the corresponding unique geodesic defined for $t\in I$. Obviously $c_{p,0}(t)=p$ for all t and hence $c_{p,0}(1)$ is defined. Furthermore, it follows from the general facts about smooth dependence of solutions of ODE systems on the initial data, that $c_{p,\xi}(1)$ is defined for $\xi\in T_pM$ in a neighborhood U of 0 in T_pM and smoothly depends on $\xi\in U$.

Definition 4.17. The *exponential map* at $p\in M$ is given by

$$\exp_p(\xi) := c_{p,\xi}(1),\quad U\subset T_pM\to M. \tag{3.3}$$

Lemma 4.18. *The map $\exp_p$ is a local diffeomorphism at 0 with $d_0\exp_p=\mathsf{id}$.*

Proof. It follows directly from the definition of geodesics that with $c(t)$ being geodesic, any curve $c(\lambda t)$ is also a geodesic for any $\lambda \in \mathbb{R}$. Hence we obtain the *homogeneity* $c_{p,\lambda\xi}(t) = c_{p,\xi}(\lambda t)$ and hence

$$d_0 \exp_p(v) = \frac{d}{dt}\Big|_{t=0} \exp_p(tv) = \frac{d}{dt}\Big|_{t=0} c_{p,tv}(1) = \frac{d}{dt}\Big|_{t=0} c_{p,v}(t) = v$$

as required (the last equality follows from the construction of $c_{p,v}(t)$. The fact that $\exp_p$ is a local diffeomorphism at 0 follows from the inverse mapping theorem. $\square$

Using Lemma 4.18 one can defined local coordinates near p by taking as a chart the local inverse of $\exp_p$. These are the important *normal coordinates* at p in which every straight line through the origin is a geodesic.

4. Curvature and the Gauss equation

The *curvature* or the *(Riemannian) curvature tensor* of a Riemannian manifold M associates to vector fields X, Y, Z the new vector field given by

$$R(X,Y)Z := \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z, \tag{4.1}$$

where ∇ is the Levi-Civita connection. More generally, given any connection ∇, it curvature tensor is given by (4.1). The name "tensor" reflects the fact that the value $R(X,Y)Z$ at a point p depends only on the values of $X(p)$, $Y(p)$ and $Z(p)$. Indeed, a calculation in local coordinates shows that

$$R\left(X^i\frac{\partial}{\partial x^i}, Y^j\frac{\partial}{\partial x^j}\right)\left(Z^k\frac{\partial}{\partial x^k}\right) = R^s_{ijk}X^iY^jZ^k\frac{\partial}{\partial x^s} \tag{4.2}$$

with

$$R^s_{ijk} = \frac{\partial\Gamma^s_{jk}}{\partial x^i} - \frac{\partial\Gamma^s_{ik}}{\partial x^j} + \Gamma^s_{il}\Gamma^l_{jk} - \Gamma^s_{jl}\Gamma^l_{ik}. \tag{4.3}$$

The curvature tensor $R(X,Y)Z$ satisfies the Bianchi identity (similar to the Jacobi identity for Lie brackets):

$$R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = 0. \tag{4.4}$$

Indeed, since the value of $(R(X,Y)Z)$ at p depends only on $X(p)$, $Y(p)$ and $Z(p)$, we may choose X, Y, Z to have constant coefficients in some local coordinates and hence all Lie brackets to be zero. Then the symmetry of ∇ implies $\nabla_X Y = \nabla_Y X$, $\nabla_Y Z = \nabla_Z Y$ and $\nabla_Z X = \nabla_X Z$. Then (4.4) is obtained directly from (4.1) using these relations.

It is convenient to associate with $R(X,Y)Z$ the closely related 4-multilinear scalar form

$$R(X,Y,Z,V) := g(R(X,Y)Z,V). \tag{4.5}$$

Since g is positive definite, $R(X,Y,Z,V)$, in turn, uniquely determines $R(X,Y)Z$.

Lemma 4.19. *The form $R(X,Y,Z,V)$ has the following symmetry properties:*

$$R(X,Y,Z,V) = -R(Y,X,Z,V) \tag{4.6}$$

$$R(X,Y,Z,V) = -R(X,Y,V,Z) \tag{4.7}$$

$$R(X,Y,Z,V) = R(Z,V,X,Y). \tag{4.8}$$

In particular,

$$R(X,X,Z,V) = R(X,Y,Z,Z) = 0. \tag{4.9}$$

The proofs are standard and can be found in most books. (In case M is a submanifold in $\mathbb{R}^n$ these properties follow immediately from the Gauss equation (4.10) below).

The following important relation known as *Gauss equation* generalizes the famous Theorema Egregium of Gauss.

Theorem 4.20 (The Gauss equation). *Let M be a submanifold of a Riemannian manifold $(\widetilde{M}, \widetilde{g})$ with the induced metric. Denote by R and $\widetilde{R}$ the corresponding curvature tensors and by II the second fundamental form of M in $\widetilde{M}$. Then, for X, Y, Z, V tangent to M, one has the relation*

$$R(X,Y,Z,V) - \widetilde{R}(X,Y,Z,V) = \widetilde{g}\big(\mathrm{II}(Y,Z), \mathrm{II}(X,V)\big) - \widetilde{g}\big(\mathrm{II}(X,Z), \mathrm{II}(Y,V)\big). \tag{4.10}$$

In particular, if $\widetilde{M} = \mathbb{R}^n$ with the standard metric, $\widetilde{R} = 0$ and hence R is given by the right-hand side of (4.10)*.*

Proof. By defintion (4.1) we have

$$\widetilde{R}(X,Y,Z,V) = \widetilde{g}\big(\widetilde{\nabla}_X \widetilde{\nabla}_Y Z, V\big) - \widetilde{g}\big(\widetilde{\nabla}_Y \widetilde{\nabla}_X Z, V\big) - \widetilde{g}\big(\widetilde{\nabla}_{[X,Y]} Z, V\big). \tag{4.11}$$

Using the formula (2.12) and the definition (2.13) of the second fundamental form we write

$$\widetilde{\nabla}_X \widetilde{\nabla}_Y Z = \widetilde{\nabla}_X \nabla_Y Z + \widetilde{\nabla}_X \mathrm{II}(Y,Z). \tag{4.12}$$

Since $\nabla_Y Z$ is also tangent to M, we can again apply (2.12) and (2.13) to the first term:

$$\widetilde{\nabla}_X \widetilde{\nabla}_Y Z = \nabla_X \nabla_Y Z + \mathrm{II}(X, \nabla_Y Z) + \widetilde{\nabla}_X \mathrm{II}(Y,Z). \tag{4.13}$$

Since $\mathrm{II}(X, \nabla_Y Z)$ is normal to TM, we have

$$\widetilde{g}\big(\mathrm{II}(X, \nabla_Y Z), V\big) = 0. \tag{4.14}$$

For the same reason $\widetilde{g}(\mathrm{II}(Y,Z), V) = 0$ and using the fact that $\widetilde{\nabla}$ is metric (2.5), we obtain

$$0 = D_X \widetilde{g}\big(\mathrm{II}(Y,Z), V\big) = \widetilde{g}\big(\widetilde{\nabla}_X \mathrm{II}(Y,Z), V\big) + \widetilde{g}\big(\mathrm{II}(Y,Z), \widetilde{\nabla}_X V\big) \tag{4.15}$$

and hence, using again that $\mathrm{II}(Y,Z)$ is normal to TM and $\mathrm{II}(X,V)$ is the orthogonal projection of $\widetilde{\nabla}_X V$ to the normal space,

$$\widetilde{g}\big(\widetilde{\nabla}_X \mathrm{II}(Y,Z), V\big) = -\widetilde{g}\big(\mathrm{II}(Y,Z), \widetilde{\nabla}_X V\big) = -\widetilde{g}\big(\mathrm{II}(Y,Z), \mathrm{II}(X,V)\big). \tag{4.16}$$

We can now compute the scalar product of (4.13) with V using (4.14) and (4.16):

$$\widetilde{g}\big(\widetilde{\nabla}_X \widetilde{\nabla}_Y Z, V\big) = \widetilde{g}\big(\nabla_X \nabla_Y Z, V\big) - \widetilde{g}\big(\mathrm{II}(Y,Z), \mathrm{II}(X,V)\big). \tag{4.17}$$

Exchanging X and Y we obtain

$$\widetilde{g}\big(\widetilde{\nabla}_Y \widetilde{\nabla}_X Z, V\big) = \widetilde{g}\big(\nabla_Y \nabla_X Z, V\big) - \widetilde{g}\big(\mathrm{II}(X,Z), \mathrm{II}(Y,V)\big). \tag{4.18}$$

Finally, since $[X,Y]$ and V are tangent to M and $\nabla_{[X,Y]} Z$ is the tangential component of $\widetilde{\nabla}_{[X,Y]} Z$, we have

$$\widetilde{g}\big(\widetilde{\nabla}_{[X,Y]} Z, V\big) = \widetilde{g}\big(\nabla_{[X,Y]} Z, V\big). \tag{4.19}$$

The required formula (4.10) follows directly from (4.17) – (4.19). □

Example 4.21. In case M is a surface in $\mathbb{R}^3$, the second fundamental form is scalar valued and can be diagonalized in an orthonormal basis (E_1, E_2) of vector fields, i.e.

$$g(E_i, E_j) = \delta_{ij}, \quad \mathrm{I\!I}(E_i, E_j) = \lambda_i \delta_{ij}. \tag{4.20}$$

The eigenvalues λ_1, λ_2 are classically called the *principal curvatures* of M. The Gauss equation (2.12) now yields

$$R(E_1, E_2, E_1, E_2) = \lambda_1 \lambda_2. \tag{4.21}$$

The right-hand side is the classical *Gaussian curvature* of M which, in view of (4.21) depends only on the intrinsic metric of M but not on the isometric embedding of M into $\mathbb{R}^3$. The latter fact is known as Theorema Egregium of Gauss. The formula (4.21) completely determines the tensor R via the symmetry relations in Lemma 4.19:

$$R(E_1, E_2, E_1, E_2) = R(E_2, E_1, E_2, E_1) = -R(E_2, E_1, E_1, E_2) = -R(E_1, E_2, E_2, E_1) = \lambda_1 \lambda_2 \tag{4.22}$$

and all other values of $R(E_i, E_j, E_k, E_l)$ are zero.

Example 4.21 shows that the curvature tensor of the surface is essentially given by the number $\lambda_1 \lambda_2$ in (4.21). For manifolds of higher dimension, one obtains a number called the sectional curvature for every 2-dimensional subspace of the tangent space.

Definition 4.22. The *sectional curvature* of a Riemannian manifold (M, g) in the direction of the plain defined by two linearly independent vectors $\xi, \eta \in T_pM$ is defined by

$$K(\xi, \eta) := \frac{R(\xi, \eta, \xi, \eta)}{g(\xi, \xi) g(\eta, \eta) - g(\xi, \eta)^2} \tag{4.23}$$

The quantity in the denominator of (4.23) is the squared area of the parallelogram spanned by ξ and η. The sectional curvature can be seen as the "normalized" value of $R(\xi, \eta, \xi, \eta)$. The main fact about the ratio in (4.23) is that it depends only on the plain spanned in T_pM by ξ and η but not on the actual choice of ξ and η. This can be seen by observing that any basis change in a two-plain can be obtained as composition of the elementary changes $(\xi, \eta) \mapsto (\eta, \xi)$, $(\xi, \eta) \mapsto (\lambda\xi, \eta)$ and $(\xi, \eta) \mapsto (\xi + \lambda\eta, \eta)$ and verifying that the ratio $K(\xi, \eta)$ does not change under these transformations. In particular, (ξ, η) can be chosen to be an orthonormal basis in which case we have $K(\xi, \eta) = R(\xi, \eta, \xi, \eta)$.

Other important quantities obtained from R are *Ricci tensor* and the *scalar curvature*, both obtained using *traces* or *contractions*.

The *Ricci tensor* at a point $p \in M$ is given by the trace (contraction)

$$\mathrm{Ric}(X, Y)(p) := \sum_i R(X, E_i, Y, E_i)(p) \tag{4.24}$$

with the summation is taken over an orthonormal basis $(E_1, \dots, E_n)$ of T_pM. The right-hand side is equal to the trace of the operator $Z \mapsto R(X, Z)Y$. The *Ricci curvature* is given by the ratio

$$\mathrm{Ric}(X) := \frac{\mathrm{Ric}(X, X)}{g(X, X)} \tag{4.25}$$

and clearly depends only on the direction of X. It follows from Lemma 4.19 that $\mathrm{Ric}(X, Y)$ is symmetric in X and Y.

The *scalar curvature* is a (scalar) function on M obtained by taking the trace one more time:

$$S(p) := \sum_i \mathrm{Ric}(E_i, E_i)(p). \tag{4.26}$$