Dmitry Fuchs

NOTES ON MANIFOLDS (239)

1. Generalities.

1.1 Definition of a manifold.

1.1.1. Charts.

Let M be a set. An *n-dimensional chart* on M is a pair (U,φ) where U is an open subset of $\mathbb{R}^n$ and φ is a 1–1 map of U into M.

Two n-dimensional charts on M, (U,φ) and (V,ψ), are called *compatible*, if

(1) the set $\varphi^{-1}(\psi(V))\subset U$ is open (in $U\Rightarrow$ in $\mathbb{R}^n$);

(2) the set $\psi^{-1}(\varphi(U))\subset V$ is open (in $V\Rightarrow$ in $\mathbb{R}^n$);

(3) the map $\varphi^{-1}(\psi(V))\xrightarrow{u\mapsto\psi^{-1}\varphi(u)}\psi^{-1}(\varphi(U))$ is smooth*.

(4) the map $\psi^{-1}(\varphi(U))\xrightarrow{v\mapsto\varphi^{-1}\psi(v)}\varphi^{-1}(\psi(V))$ is smooth.

In particular, (U,φ) and (V,ψ) are compatible, if $\varphi(U)\cap\psi(V)=\emptyset$.

1.1.2. Atlases.

A set $\{(U_\alpha,\varphi_\alpha)\mid\alpha\in A\}$ of n-dimensional charts on M is called an (n-dimensional) *atlas* if

(1) $\bigcup\limits_{\alpha\in A}\varphi_\alpha(U_\alpha)=M$;

(2) For any $\alpha,\beta\in A$ the charts $(U_\alpha,\varphi_\alpha),(U_\beta,\varphi_\beta)$ are compatible.

Two n-dimensional atlases on M, $\mathcal{A}$ and $\mathcal{B}$ are called equivalent, if their union $\mathcal{A}\cup\mathcal{B}$ is also an atlas (in other words, if any chart of $\mathcal{A}$ is compatible with any chart of $\mathcal{B}$).

[Notice that this relation is obviously reflexive and symmetric, but it is also transitive because of the following

Lemma. *Let $\mathcal{A}$ be an atlas, and $(U,\varphi),(V,\psi)$ be charts (not assumed to belong to $\mathcal{A}$). If (U,φ) and (V,ψ) are compatible with every chart of $\mathcal{A}$, then they are compatible with each other.*

Proof. Let $u\in\varphi^{-1}(\psi(V))\subset U$, and let $x=\varphi(u)\in M$. Then $x\in\chi(W)$ for some chart $(W,\chi)\in\mathcal{A}$, $x=\chi(w)$ for some $w\in W$. Thus,

$$\psi^{-1}(\varphi(u))=\psi^{-1}(x)=\psi^{-1}(\chi(w))=\psi^{-1}(\chi(\chi^{-1}(\varphi(u)))),$$

and $\psi^{-1}\circ\varphi$ is smooth at (a neighborhood of) u since it is a composition of (appropriate restrictions of) smooth maps $\psi^{-1}\circ\chi$ and $\chi^{-1}\circ\varphi$.

Similarly, $\varphi^{-1}\circ\psi$ is smooth.

Tis proves Lemma, and we see that if atlases $\mathcal{B}$ and $\mathcal{C}$ are equivalent to the atlas $\mathcal{A}$, then every charts of $\mathcal{B}$ and $\mathcal{C}$ are compatible with every chart of $\mathcal{A}$, hence, they are compatible with each other, and hence the atlases $\mathcal{B}$ and $\mathcal{A}$ are equivalent.]

* The word *smooth* will always mean $\mathcal{C}^\infty$; in particular, smooth maps are continuous.

1.1.3. Topology.

Let M be a set with an n-dimensional atlas $\mathcal{A}$. A subset B of M is called open (with respect to $\mathcal{A}$), if for any chart $(U, \varphi) \in \mathcal{A}$ the set $\varphi^{-1}(B)$ is open (in $U \Rightarrow$ in $\mathbb{R}^n$). (In particular, the sets $\varphi(U)$ are open.)

Proposition. *If the atlases $\mathcal{A}$ and $\mathcal{B}$ are equivalent, then a set $B \subset M$ is open with respect to $\mathcal{A}$ if and only if it is open with respect to $\mathcal{B}$.*

Proof. For any $B \subset M$,

$$B = \bigcup_{(V,\psi)\in\mathcal{B}} (B \cap \psi(V)) = \bigcup_{(V,\psi)\in\mathcal{B}} \psi(\psi^{-1}(B)).$$

If B is open with respect to $\mathcal{B}$, then all the sets $\psi^{-1}(B)$ are open. Let (U, φ) be a chart of $\mathcal{A}$. Then

$$\varphi^{-1}(B) = \varphi^{-1} \bigcup_{(V,\psi)\in\mathcal{B}} \psi(\psi^{-1}(B)) = \bigcup_{(V,\psi)\in\mathcal{B}} \varphi^{-1}\psi(\psi^{-1}(B)) = \bigcup_{(V,\psi)\in\mathcal{B}} (\psi^{-1}\varphi)^{-1}(\psi^{-1}(B)).$$

Since $\psi^{-1}\varphi$ is continuous (see condition (3) in 1.1.1), the last formula shows that $\varphi^{-1}(B)$ is a union of open set; hence it is also open. Since this is true for any chart $(U, \varphi) \in \mathcal{A}$, the set B is open with respect to $\mathcal{A}$. It is easy to check (left to the reader) that sets, open with respect to an atlas, form a *topology* (that is, $\emptyset$ and M are open, any union and any finite intersection of open sets is open).

This proposition shows that an equivalence class of atlases on M makes M a topological space, and we can speak of its purely topological properties like compactness or connectedness*. Actually, the two axioms given below in 1.1.4 are of topological nature: they are called in topology Second Countability Axiom and Hausdorff axiom.

1.1.4. Manifolds.

A class $\mathfrak{D}$ of equivalent n-dimensional atlases on M is called an *n-dimensional differential structure* on M, if the following two additional conditions hold:

(1) the class $\mathfrak{D}$ contains an at most countable atlas;

(2) for any different $p, q \in M$ there exist disjoint open $U, V \subset M$ such that $p \in U, q \in V$.

Charts of atlases from $\mathfrak{D}$ are called simply charts of $\mathfrak{D}$.

A set M with n-dimensional differential structure is called a (smooth) *n-dimensional manifold*.

Note that instead of a class of equivalent atlases we can speak of a *maximal* atlas. An atlas $\mathcal{A}$ is maximal if it contains all charts compatible with all its charts, in other words,

---

* M is *compact*, if every atlas contains a finite subatlas; M is *connected*, if for every two points $p, q \in M$ there exists a finite sequence of charts, $\{(U_i, \varphi_i),\ i = 1, \dots, n\}$ such that $p \in \varphi_1(U_1)$, $q \in \varphi_n(U_n)$, U_i is connected for $i = 1, \dots, n$, $\varphi_i(U_i) \cap \varphi_{i+1}(U_{i+1})$ is not empty for $i = 1, \dots, n-1$.

if it contains any atlas which is equivalent to it. Any class of equivalent atlases contains precisely one maximal atlas: the union of all its atlases. Thus we can define a differential structure as a maximal atlas (and replace the condition (1) above by the statement that our maximal atlas contains an at most countable subatlas).

1.1.5. Local coordinates.

Usually, speaking of manifolds, we will not refer explicitly to charts, but rather we will speak of *local coordinates*. If (U, φ) is a chart of the differential structure of M, then each point $p \in \varphi(U)$ is characterized by the n coordinates of $\varphi^{-1}(p) \in U \subset \mathbb{R}$. Thus, within $\varphi(U)$ there arises a (local) coordinate system. If p belongs to the domains of two different local coordinate systems, then, in some neighborhood of p, we have two coordinate systems, $\{x_1, \dots, x_n\}$ and $\{y_1, \dots, y_n\}$. Axioms (3) and (4) of 1.1.1 mean that, in this neighborhood, x_i's are smooth functions of $y_1, \dots, y_n$, and y_i's are smooth functions of $x_1, \dots, x_n$. Moreover, the Jacobian matrices $\left\| \frac{\partial x_i}{\partial y_j} \right\|$ and $\left\| \frac{\partial y_i}{\partial x_j} \right\|$ are inverse to each other and, in particular, invertible (non-degenerate).

1.1.6. Orientations.

Let M be a smooth manifold of dimension $n > 0$, let $(U, \varphi), (V, \psi)$ be two charts of M, and let $p \in \varphi(U) \cap \psi(V)$. We say that *the orientations of the two charts agree at p* if the determinant of the Jacobian matrix of the transformation $\varphi^{-1}\psi(V) \to \psi^{-1}\varphi(U)$, $x \mapsto \psi^{-1}(\varphi(x))$ at $\varphi^{-1}(p)$ is positive; if this determinant is negative, we say that the orientations *disagree*. Obviously, the set of all charts that cover p splits into two classes: the orientations of charts agree at p within each class and disagree between the classes. These classes are called *orientations at p*. Thus, for each point $p \in M$ there are two orientations at p.

An atlas is called *oriented* if the orientations of any two charts $(U, \varphi), (V, \psi)$ of this atlas agree at any common point of $\varphi(U)$ and $\psi(V)$. We say that two oriented atlases determine the same orientation, if their union is also an oriented atlas. An *orientation* of a manifold is a class of atlases that determine the same orientation. (In other words, an orientation is a *maximal oriented atlas*, that is an oriented atlas which contains any chart whose orientation agrees with that of any of its charts.) A manifold is called *orientable*, if it possesses an orientation, and is called oriented if an orientation is chosen for it. For any orientation of an oriented manifold there is the *opposite* orientation: if the first orientation is given by an atlas $\{(U, \varphi)\}$, then the opposite orientation is given by the atlas $\{(\rho(U), \varphi \circ \rho)\}$ where ρ is the reflection $(x_1, x_2, \dots, x_n) \mapsto (-x_1, x_2, \dots, x_n)$. Obviously, the two orientations are different.

An orientation of a manifold may be regarded as a simultaneous choice of orientations at all its points, provided that this choice is made in a coherent way; the latter means that any chart (U, φ) with connected U either belongs to the chosen orientation at every point of $\varphi(U)$ or does not belong to the chosen orientation at any of these points. (The latter means also "belongs to the opposite orientation.")

PROPOSITION. *A connected orientable manifold has precisely two different orientations, and they are opposite to each other. A disconnected manifold is orientable if and only if all its components are orientable; a choice of an orientation of a disconnected manifold is the same as (independent) simultaneous choices of orientations for all its components.*

We leave the proof to the reader.

1.2. Examples.

1.2.1. Euclidean spaces.

$\mathbb{R}^n$ itself is an n-dimensional manifold. The differential structure is determined by the one-chart atlas $\{(\mathbb{R}^n, \mathrm{id})\}$.

1.2.2. Spheres.

By definition, the n-dimensional sphere S^n is $\{(x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} \mid x_1^2 + \dots + x_{n+1}^2 = 1\}$. Define the maps $\varphi_+, \varphi_-\colon \mathbb{R}^n \to S^n$ by the formula

$$\varphi_\pm(x_1, \dots, x_n) = \left(\frac{2x_1}{1+r^2}, \dots, \frac{2x_n}{1+r^2}, \pm\frac{1-r^2}{1+r^2} \right)$$

where $r^2 = x_1^2 + \dots + x_n^2$ (geometrically, $\varphi_\pm(x_1, \dots, x_n)$ is the point of intersection of the sphere and the line passing through the points $(x_1, \dots, x_n, 0)$ and $(0, \dots, 0, \pm 1)$ different from $(0, \dots, 0, \pm 1)$; in geometry, the map $\varphi_+^{-1}\colon [S^n - (0, \dots, 0, 1)] \to \mathbb{R}^n$ is called the *stereographic projection.*). Since the maps $\varphi_\pm$ are, obviously, 1–1, we obtain two n-dimensional charts, $(\mathbb{R}^n, \varphi_\pm)$. Since $\varphi_\pm(\mathbb{R}^n) = S^n - (0, \dots, 0, \pm 1)$, it is true that $\varphi_+(\mathbb{R}^n) \cup \varphi_-(\mathbb{R}^n) = S^n$. The two charts are compatible: $\varphi_-^{-1}(\varphi_+(\mathbb{R}^n)) = \varphi_+^{-1}(\varphi_-(\mathbb{R}^n)) = \mathbb{R}^n - (0, \dots, 0)$ and each of the maps $\varphi_-^{-1} \circ \varphi_+, \varphi_+^{-1} \circ \varphi_-\colon \mathbb{R}^n - (0, \dots, 0) \to \mathbb{R}^n - (0, \dots, 0)$ acts as

$$(x_1, \dots, x_n) \mapsto \left(\frac{x_1}{r^2}, \dots, \frac{x_n}{r^2} \right)$$

(geometrically, it is the inversion)* and hence is smooth. Thus, the two charts form an atlas. We do not check the Hausdorff axiom (actually, the topology arising is the usual topology of the sphere). Thus, the sphere S^n becomes a smooth n-dimensional manifold.

1.2.3. Projective spaces.

The n-dimensional (real) projective space $\mathbb{R}P^n$ is defined as the set of all straight lines in $\mathbb{R}^{n+1}$ passing through the origin. Define the map $\varphi_i\colon \mathbb{R}^n \to \mathbb{R}P^n$, $1 \le i \le n+1$ in the following way: $\varphi_i(x_1, \dots, x_n)$ is the line in $\mathbb{R}^{n+1}$ passing through the points $(0, \dots, 0)$ and $(x_1, \dots, x_{i-1}, 1, x_i, \dots, x_n)$. Obviously, φ_i is 1–1, hence $(\mathbb{R}^n, \varphi_i)$ are n-dimensional charts on $\mathbb{R}P^n$. Also, the sets $\varphi_i(\mathbb{R}^n)$ cover $\mathbb{R}P^n$ (if a line $\ell \in \mathbb{R}P^n$ contains a point $(y_1, \dots, y_{n+1})$ with $y_j \ne 0$, then $\ell \in \varphi_j(\mathbb{R}^n)$). Let $V_k = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_k \ne 0\}$, $1 \le k \le n$.

---

* Indeed, $\left(\frac{x_1}{r^2}\right)^2 + \dots + \left(\frac{x_n}{r^2}\right)^2 = r^{-2}$, and

$$\varphi_\pm\left(\frac{x_1}{r^2}, \dots, \frac{x_n}{r^2}\right) = \left(\frac{2x_1}{r^2(1+r^{-2})}, \dots, \frac{2x_n}{r^2(1+r^{-2})}, \pm\frac{1-r^{-2}}{1+r^{-2}} \right)$$
$$= \left(\frac{2x_1}{1+r^2}, \dots, \frac{2x_n}{1+r^2}, \mp\frac{1-r^2}{1+r^2} \right) = \varphi_\mp(x_1, \dots, x_n)$$

.

Obviously, V_k are open in $\mathbb{R}^n$ and $\varphi_i^{-1}\varphi_j(\mathbb{R}^n) = V_{j-1}$, if $i < j$, and $= V_j$, if $i > j$. The map $\varphi_j^{-1}\varphi_i\colon \varphi_i^{-1}\varphi_j(\mathbb{R}^n) \to \varphi_j^{-1}\varphi_i(\mathbb{R}^n)$ is

$$V_{j-1} \to V_i,$$
$$(x_1,\dots,x_n) \mapsto \left(\frac{x_1}{x_{j-1}},\dots,\frac{x_{i-1}}{x_{j-1}},\frac{1}{x_{j-1}},\frac{x_i}{x_{j-1}},\dots,\frac{x_{j-2}}{x_{j-1}},\frac{x_j}{x_{j-1}},\dots,\frac{x_n}{x_{j-1}}\right), \text{ if } i < j,$$
$$V_j \to V_{i-1}, (x_1,\dots,x_n) \mapsto \left(\frac{x_1}{x_j},\dots,\frac{x_{j-1}}{x_j},\frac{x_{j+1}}{x_j},\dots,\frac{x_{i-1}}{x_j},\frac{1}{x_j},\frac{x_i}{x_j},\dots,\frac{x_n}{x_j}\right), \text{ if } i > j.$$

This is smooth, hence the charts $(\mathbb{R}^n, \varphi_i)$ are mutually compatible and thus they form an atlas. We skip checking the Hausdorff axiom and conclude that $\mathbb{R}P^n$ acquires a structure of a smooth n-dimensional manifold.

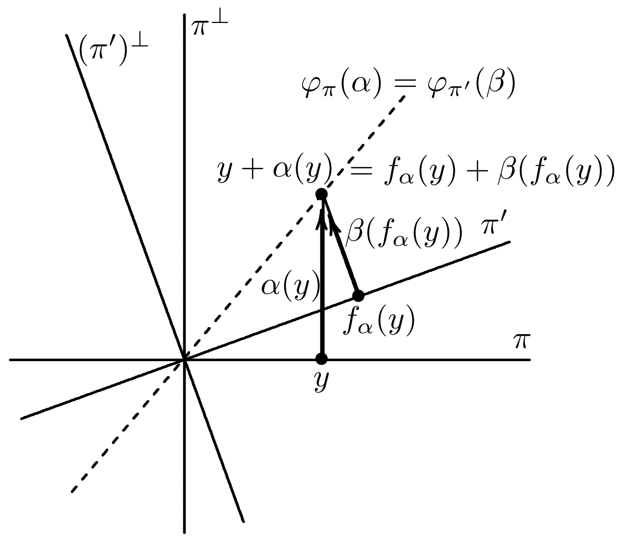
1.2.4. Grassmann manifolds.

For non-negative integers m, n, denote as $G(n,m)$ the set of all n-dimensional subspaces of the space $\mathbb{R}^{n+m}$. (In particular, $G(0,m)$ and $G(n,0)$ are one-point sets, and $G(1,n)$ is $\mathbb{R}P^n$.) Our goal is to equip $G(n,m)$ with the structure of an mn-dimensional manifold. This manifold is known as the *Grassmann manifold*.

For a $\pi \in G(n,m)$, define the map $\varphi_\pi\colon \mathcal{L}(\pi,\pi^\perp) \to G(n,m)$ (where $\mathcal{L}(\pi,\pi^\perp)$ is the (mn-dimensional) space of all linear maps from π to $\pi^\perp$) by the formula

$$\varphi_\pi(\alpha) = (\mathrm{id}_\pi \oplus \alpha)(\pi)$$

(here $\mathrm{id}_\pi \oplus \alpha$ is regarded as a map $\pi \to \pi \oplus \pi^\perp = \mathbb{R}^{n+m}$; in other words, $\varphi_\pi(\alpha)$ is the graph of α in $\pi \oplus \pi^\perp = \mathbb{R}^{n+m}$). Obviously, φ_π is one-to-one (equal graphs $\Rightarrow$ equal maps), thus, $(\mathcal{L}(\pi,\pi^\perp), \varphi_\pi)$ is an nm-dimensional chart of $G(n,m)$. The images $\varphi_\pi(\mathcal{L}(\pi,\pi^\perp))$ cover $G(n,m)$ (for example, $\pi = \varphi_\pi(0) \in \varphi_\pi(\mathcal{L}(\pi,\pi^\perp))$). Let us prove that these charts are mutually compatible.



Let $\pi, \pi' \in G(n,m)$, and let p, p' be orthogonal projections of $\mathbb{R}^{n+m}$ onto π, π'. We want to check that the map $F = \varphi_{\pi'}^{-1}\varphi_\pi\colon \varphi_\pi^{-1}(\varphi_{\pi'}(\mathcal{L}(\pi',(\pi')^\perp))) \to \varphi_{\pi'}^{-1}(\varphi_\pi(\mathcal{L}(\pi,\pi^\perp)))$ is smooth. For $\alpha \in \mathcal{L}(\pi, \pi^\perp), \beta \in \mathcal{L}(\pi', (\pi')^\perp$, the equality $F(\alpha) = \beta$ means that $\varphi_\pi(\alpha) = \varphi_{\pi'}(\beta)$. Let $f_\alpha : \pi \to \pi'$ be defined by the formula $f_\alpha = p' \circ (\mathrm{id}\, \pi + \alpha)$. We need two things: f_α is invertible and for every $y \in \pi,\ y + \alpha(y) = f_\alpha(y) + \beta(f_\alpha(y))$ (the condition $\det f_\alpha \neq 0$ gives an exact description of the subset $\varphi_\pi^{-1}(\varphi_{\pi'}(\mathcal{L}(\pi', (\pi')^\perp)))$ of $\mathcal{L}(\pi, \pi^\perp)$ which is therefore open). For β we have $(\mathrm{id}_{\pi'} \oplus \beta) \circ f_\alpha = \mathrm{id}_\pi \oplus \alpha$, or

$$\beta = F(\alpha) = (\mathrm{id}_\pi \oplus \alpha) \circ f_\alpha^{-1} - \mathrm{id}_{\pi'}$$

(it follows from our construction that the image of β is contained in $(\pi')^\perp$).

Notice that this construction provides $G(n, m)$ with an infinite atlas with charts labeled by $\pi \in G(n, m)$; but actually, we do need all this charts: it is sufficient to consider only $\binom{n+m}{n}$ charts corresponding to subspaces π spanned with n coordinate axes. After this reduction, the atlas generalizes the atlas constructed in 1.2.3 for projective spaces.

1.2.5. Atlases not satisfying additional axioms of 1.1.4.

There are many examples of atlases which do not satisfy the topological restrictions of 1.1.4 (second countability and Hausdorff's axiom). Some of this examples are artificial, but some represent important mathematical notions (like the example considered in 1.2.5.3 below).

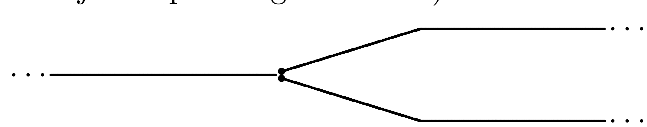
1.2.5.1. Uncountable unions.

A disjoint union $M = \coprod_{\alpha \in A} M_\alpha$ where A is an uncountable set and each M_α is a (non-empty) smooth n-dimensional manifold has a natural n-dimensional atlas consisting of all charts of all M_α's. This atlas satisfies the Hausdorff axiom, but not the countability axiom.

There are also examples of atlases which give rise to uncountable Hausdorff topology satisfying the property of *connectedness*.

1.2.5.2. A double line.

Let $M = (\mathbb{R} \times \{0, 1\})/(t, 0) \sim (t, 1)$, if $t < 0$. The two charts $\varphi_0, \varphi_1 : \mathbb{R} \to M,\ \varphi_i(t) = (t, i)$ form a 1-dimensional atlas of M, but the Hausdorff axiom does not hold (the points $(0, 0), (0, 1)$ do not have disjoint open neighborhoods):



1.2.5.3. An example from the sheaf theory.

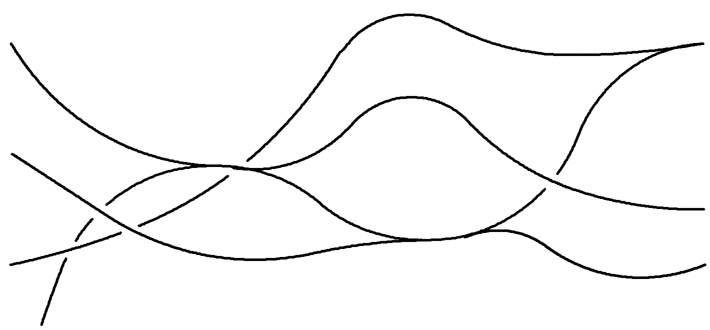
Let $p \in \mathbb{R}$. Call two smooth functions $f, g : \mathbb{R} \to \mathbb{R}$ equivalent at p, if, for some $\varepsilon > 0$, $f(x) = g(x) \forall x \in [p - \varepsilon, p + \varepsilon]$. Equivalence classes of functions with respect to this relation are called *germs* of functions at p, and the set of germs of functions at p is denoted as $\mathcal{S}_p$. Let $\mathcal{S}$ be the union $\coprod_{p \in \mathbb{R}} \mathcal{S}_p$. Every smooth function $f : \mathbb{R} \to \mathbb{R}$ belongs to a certain germ, f_p, for every $p \in \mathbb{R}$ (this germ f_p is referred to as the *germ of f at p*). The map $\varphi_f : \mathbb{R} \to \mathcal{S},\ \varphi_f(p) = f_p \in \mathcal{S}_p \subset \mathcal{S}$, is $1-1$ and may be regarded as a 1-dimensional chart of $\mathcal{S}$. These charts cover $\mathcal{S}$ (every germ is a germ of some function).

For two smooth functions, $f, g : \mathbb{R} \to \mathbb{R}$ the conditions $\psi_f(p) \in \psi_g(\mathbb{R})$ and $\psi_f(p) \in \psi_g(\mathbb{R}$ are equivalent to each other and are equivalent to the equality $f_p = g_p$, that is, to the fact that f and g agree in some neighborhood of p. It is clear that this set is open, that it is both $\psi_g^{-1}\psi_f(\mathbb{R})$ and $\psi_f^{-1}\psi_g(\mathbb{R})$, and that the maps from conditions (3) and (4) of 1.1.1 are both equal to id. Thus, the charts $(\mathbb{R}, \varphi_f)$ form a 1-dimensional atlas of $\mathcal{S}$. However, neither of coditions (1) and (2) in 1.1.4 is satisfied. First, $\mathcal{S}$ is not second countable, since the sets $\varphi_c(\mathbb{R})$ for constant functions c form an uncountable set of mutually disjoint open subsets of $\mathcal{S}$. Second, $\mathcal{S}$ is not Hausdorff; moreover, for every $s \in \mathcal{S}$ there exists an $s' \in \mathcal{S}$ such that s and s' have no disjoint open neighborhoods: if $s = f_p$ then we can put $s' = g_p$ where $g\colon \mathbb{R} \to \mathbb{R}$ is such smooth function that $g(x) = f(x)$ for $x \geq p$ and $g(x) \neq f(x)$ for $x < p$ (examples of such functions are well known; we will need a precise description of them below, see 1.3.3.2.1).

One can think of the topological space $\mathcal{S}$ as of the union of *all* graphs of all smooth functions $\mathbb{R} \to \mathbb{R}$; if two functions agree on an open interval, we glue the graphs together (but do not glue them at the endpoints of the interval); if, say, two functions agree at an isolated value of the argument, we consider the graphs as disjoint at this point. And the intersection of $\mathcal{S}$ with any vertical line $x = \text{const}$ is discrete.

The picture below (see next page) shows a small portion of $\mathcal{S}$ (the union of four graphs).

The topology of $\mathcal{S}$ is called the *sheaf topology*. The construction presented here has multiple generalizations which are important in the so called *sheaf theory*.



1.2.6. Constructions.

In 1.2.1–1.2.4, we considered some important but still sporadic examples of smooth manifold. Our supply of examples will grow immensely, if we apply to existing examples various constructions which can create manifolds from other manifolds. Some of these constructions are discussed in this section.

1.2.6.1. Products.

In M and N are smooth manifolds of dimension m and n, then $M \times N$ is a smooth manifold of dimension $m+n$. The differential structure of $M \times N$ consists of all products $(U \times V, \varphi \times \psi)$ where (U, φ) and (V, ψ) are charts from the differential structures of M and N correspondingly. (We consider $U \times V$ as an open subset of $\mathbb{R}^{m+n} = \mathbb{R}^m \times \mathbb{R}^n$.) We can also consider multiple (finite) products $M_1 \times \ldots \times M_k$.

1.2.6.2. Open sets.

If M is an n-dimensional smooth manifold and $B \subset M$ is an open set, then B also has a natural structure of an n-dimensional smooth manifold: the differential structure in B consists of charts $\left(\varphi^{-1}(B), \varphi\,|_{\varphi^{-1}(B)}\right)$ for all charts (U, φ) of M.

1.2.6.3. Submanifolds.

1.2.6.3.1. Definition.

Let N be an n-dimensional smooth manifold. A subset $M \subset N$ is called an m-dimensional submanifold of M (where $m \leq n$) if for any point $p \in M$ there exists a chart (U, φ) of N such that $\varphi^{-1}(M) = U \cap \mathbb{R}^m$ (we consider $\mathbb{R}^m$ as a subset of $\mathbb{R}^n$, $(x_1, \ldots, x_m) \in \mathbb{R}^m$ is identified with $(x_1, \ldots, x_m, 0 \ldots, 0) \in \mathbb{R}^n$).

An m-dimensional submanifold M of an n-dimensional smooth manifold N is an m-dimensional smooth manifold: the charts $\left(U \cap \mathbb{R}^m, \varphi |_{U \cap \mathbb{R}^m}\right)$ considered for all charts (U, φ) of N such that $\varphi^{-1}(M) = U \cap \mathbb{R}^m$ constitute an m-dimensional differential structure on M. It is not difficult to check that the conditions (1) and (2) from 1.1.4 hold for this differential structure, so M acquires a structure of an m-dimensional smooth manifold.

Notice that the topology of a submanifold (in the sense of 1.1.3) is the same as the topology induced by the topology of the manifold as defined in the point set topology.

Example. An n-dimensional submanifold of an n-dimensional smooth manifold N is an open subset of N. Indeed, if $m = n$, then $\mathbb{R}^m = \mathbb{R}^n, U \cap \mathbb{R}^m = U$, and the condition $\varphi^{-1}(M) = U \cap \mathbb{R}^m$ becomes $\varphi(U) \subset M$.

1.2.6.3.2. Submanifolds of Euclidean spaces (a preliminary definition).

Let $M \subset \mathbb{R}^n$ is described by a system of equations $F_i(x_1, \ldots, x_n) = 0,\ i = 1, \ldots, n - m$ where $F_1, \ldots, F_{m-n}$ are smooth real-valued functions in an open neighborhood of M in $\mathbb{R}^n$ which satisfy the following non-degeneracy condition: if, for some $(x_1^0, \ldots, x_n^0) \in \mathbb{R}^n,\ F_i(x_1^0, \ldots, x_n^0) = 0$ for $i = 1, \ldots, n-m$, then the Jacobian matrix

$$\left\| \begin{matrix} \dfrac{\partial F_1}{\partial x_1} & \ldots & \dfrac{\partial F_1}{\partial x_n} \\ \ldots & \ldots & \ldots \\ \dfrac{\partial F_{n-m}}{\partial x_1} & \ldots & \dfrac{\partial F_{n-m}}{\partial x_n} \end{matrix} \right\|$$

at the point $(x_1^0, \ldots, x_n^0)$ has maximal rank (that is, $n - m$). Then M is an m-dimensional submanifold of $\mathbb{R}^n$. The proof is based on a result from standard multi-variable calculus known under the name of Inverse Function Theorem. Here it is.

Lemma (Inverse Function Theorem). *Let $h_1, \ldots, h_n$) be smooth functions of n variables $x_1, \ldots, x_n$, defined in some neighborhood of a point $P = (x_1^0, \ldots, x_n^0) \in \mathbb{R}^n$ and such that the Jacobian matrix $\left\| \dfrac{\partial h_i}{\partial x_j} \right\|(P)$ is non-singular. Then there exists a neighborhood W of P (within the domain of h_i's) such that the map $h\colon W \to \mathbb{R}^n$ with the coordinate functions $h_1, \ldots, h_n$, that is,*

$$h(x_1, \ldots, x_n) = (h_1(x_1, \ldots, x_n), \ldots, h_n(x_1, \ldots, x_n)),$$

is invertible in the sense that there is a smooth map $\varphi(= h^{-1})\colon h(W) \to W$ such that $\varphi \circ h = \mathrm{id}_W,\ h \circ \varphi = \mathrm{id}_{\varphi(W)}$.

Proof of this Theorem can be found in calculus textbooks.

Now, let $(x_1^0, \dots, x_n^0) \in M$, and let $k_1, \dots k_{n-m}$, $1 \le k_1 < \dots < k_{n-m} \le n$ be such numbers that $\det \left\| \frac{\partial F_i}{\partial x_{k_j}} \right\| \ne 0$. Let $\ell_1, \dots, \ell_m$, $1 \le \ell_1 < \dots < \ell_m \le n$ be the remaining subscripts (that is, $\{k_1, \dots, k_{n-m}\} \cup \{\ell_1, \dots, \ell_m\} = \{1, \dots, n\}$). Consider the functions $y_1 = x_{\ell_1}, \dots, y_m = x_{\ell_m}, y_{m+1} = F_1, \dots, y_n = F_{n-m}$. The Jacobian matrix $\left\| \frac{\partial y_i}{\partial x_j} \right\|$ is

$$\begin{array}{c} \qquad\quad \ell_1 \quad\ \ \ell_2 \qquad\quad \ell_m \\ \left\| \begin{array}{ccccccc} & \overset{}{1} & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ \hline \frac{\partial F_1}{\partial x_1} & \dots & \dots & \dots & \dots & \dots & \frac{\partial F_1}{\partial x_n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\partial F_{n-m}}{\partial x_1} & \dots & \dots & \dots & \dots & \dots & \frac{\partial F_{n-m}}{\partial x_n} \end{array} \right\| \end{array}$$

(with zeroes in the blank spots). The determinant of this matrix is, obviously, $\pm \det \left\| \frac{\partial F_i}{\partial x_{k_j}} \right\|$ which is not zero at P. By Inverse Function Theorem, there exists a neighborhood W of P, an open set $U \subset \mathbb{R}^n$ (denoted as $h(W)$ in Inverse Function Theorem), and a smooth invertible 1–1 map $\varphi: U \to W \subset \mathbb{R}^n$, $\varphi(x_1, \dots, x_n) = (y_1, \dots, y_n)^*$. Then $\varphi^{-1}(M) = \varphi^{-1}(\{F_1 = \dots = F_{n-m} = 0\}) = \{y_{m+1} = \dots = y_n = 0\} = U \cap \mathbb{R}^m$. Thus, (U, φ) is a chart from the definition of a submanifold.

Example. S^n is defined in $\mathbb{R}^{n+1}$ by one equation $F(x_1, \dots, x_{n+1}) = x_1^2 + \dots + x_{n+1}^2 - 1 = 0$. The matrix $\left\| \frac{\partial F}{\partial x_1}, \dots, \frac{\partial F}{\partial x_{n+1}} \right\| = \|2x_1, \dots, 2x_{n+1}\|$ has rank 1 (that is, nowhere vanishes) in S^n. Hence, S^n is an n-dimensional submanifold of $\mathbb{R}^n$, and in this capacity, it acquires a structure of an n-dimensional smooth manifold. It is true (the proof is left to the reader) that this is the same structure as the one described in 1.2.2.

1.2.6.3.3. A generalization.

We do not need to assume that the functions F_i are defined globally. It is sufficient to have, for any point $q \in M$ a neighborhood W of q in $\mathbb{R}^n$ and functions $F_1, \dots, F_{n-m}: W \to \mathbb{R}$, satisfying the non-degeneracy condition as above and such that $\{p \in W \mid F_1(p) = \dots = F_{n-m}(p) = 0\} = W \cap M$. In this case, arguing precisely as before, we can show that M is a submanifold of $\mathbb{R}^n$.

1.2.6.3.4. A further generalization.

We can replace $\mathbb{R}^n$ by an arbitrary n-dimensional smooth manifold N. The neighborhood W from the last construction should be replaced by a chart (W, φ) of N. Details are left to the reader.

---

* In different words, $y_1, \dots, y_n$ are liocal coordinates in W.

1.2.6.3.5. A final remark.

The notion of a submanifold expands our stock of examples of manifolds immensely since now we count count among them all curves and surfaces in the plane and in space, and also all multidimensional surfaces in Euclidean spaces.

1.2.7. Orientability.

Since a one-chart atlas is automatically oriented, $\mathbb{R}^n$ is orientable and has a canonical orientation.

The orientations of the charts $(\mathbb{R}^n, \varphi_\pm)$ of S^n (see 1.2..2) *disagree* at all common points (it is clear geometrically, but can also be confirmed by a computation: the determinant of the Jacobian matrix of the map $\varphi_-^{-1} \circ \varphi_+$ is equal to $-r^{2n}$, which is negative. Hence, the atlas $\{(\mathbb{R}^n, \varphi_+0, (\mathbb{R}^n, \varphi_- \circ \rho)\}$ (see 1.1.6) is oriented, and S^n is an orientable manifold.

The case of $\mathbb{R}P^n$ is more interesting. The determinant of the Jacobian matrix of the transformation $V_{j-1} \to V_i$ described in 1.2.3 is easy to compute and equals $(-1)^{j-i} x_{j-1}^{-(n+1)}$. If n is odd, then the orientations of the charts $(\mathbb{R}^n, \varphi_i)$, $(\mathbb{R}^n, \varphi_j)$ agree if i, j have the same parity and disagree otherwise; thus, in this case, the charts $(\mathbb{R}^n, \varphi_1), (\mathbb{R}^n, \varphi_2 \circ \rho), (\mathbb{R}^n, \varphi_3), (\mathbb{R}^n, \varphi_4 \circ \rho), \dots, (\mathbb{R}^n, \varphi_n), (\mathbb{R}^n, \varphi_{n+1} \circ \rho)$ form and oriented atlas. On the contrary, if n is even, then the orientations any two of our charts agree at some points and disagree at some other points. Since all the charts are connected (their common domain is $\mathbb{R}^n$), this contradicts to the existence of any orientation at all: no choice of orientations at all points of $\mathbb{R}P^n$ can be coherent (see 1.1.6) within any two of the charts. Thus, $\mathbb{R}P^n$ (with $n > 0$) is orientable if and only if n is odd.

Products of manifolds is orientable if and only if all the factors are are orientable; moreover, orientations of factors determine an orientation of the product (which may depend, however, on the order of factors). Open subsets of an orientable manifold are orientable and inherit an orientation from the ambient manifold. No definite relation exists in general between the orientabilities of a manifold and a submanifold; we can only state that a submanifold given in an orientable manifold by a single non-degenerate system of equations is also orientable (the proof is left to the reader).

1.3. Maps and functions.

1.3.1. Definition of a smooth map.

Let M and N be two manifolds of dimensions m and n. A map $f\colon M \to N$ is called smooth, if for any charts, (U, φ) of M and (V, ψ) of N, the set $W = \varphi^{-1}(f^{-1}(\psi(V))) \subset U$ is open in U (if we stop here, our condition would mean precisely that f is continuous) and the map $g\colon W \to V, g(w) = \psi^{-1}(f(\varphi(w)))$ is smooth. We can obtain an equivalent definition of a smooth map if restrict the conditions above to charts belonging to any two specific atlases representing the differential structures of M and N, or if we require that for any point $p \in M$ there exist charts (U, φ) of M and (V, ψ) of N such that $p \in \varphi(U), f(p) \in \psi(V)$ and the conditions above hold.

As we have already mentioned above, smooth maps are continuous.

1.3.2. Diffeomorphisms, embeddings and immersions.

A smooth map $f\colon M \to N$ is called a *diffeomorphism*, if f is invertible (1–1 and onto) and the inverse map $f^{-1}\colon N \to M$ is also smooth. Obviously, identities are diffeomorphisms; furthermore, compositions of diffeomorphisms and inverses to diffeomorphisms are diffeomorphisms as well.

If a diffeomorphism $M \to N$ exists, then the manifolds M and N are called diffeomorphic. It is obvious that diffeomorphic manifolds are homeomorphic [1]; in particular if one of two diffeomorphic manifolds is compact, or connected, or simply connected, then so is the other one. Also, diffeomorphic manifolds have equal dimensions. (Proof. Let $f\colon M \to N$ be a diffeomorphism, $p \in M$ and $q = f(p) \in N$. Let $U, \varphi)$ and $V, \psi)$ be the charts of M and N containing p and q. The the maps $\psi^{-1} f \varphi\colon \varphi^{-1} f^{-1} \psi(V) \to \psi^{-1} f \varphi(U)$ and $\varphi^{-1} f^{-1} \psi\colon \psi^{-1} f \varphi(U) \to \varphi^{-1} f^{-1} \psi(V)$ are smooth, inverse to each other maps between open sets in Euclidean spaces; their Jacobian matrices, correspondingly at $\varphi^{-1}(p)$ and $\psi^{-1}(q)$ are inverse to each other, hence they are square matrices.)[2]

A map $f\colon M \to N$ is called a *(smooth) embedding*, if it is a diffeomorphism of M onto a submanifold of N[3]. A map $f\colon M \to N$ is called an immersion, if it is a local embedding, that is, if for any point $p \in M$ there exists a neighborhood $B \subset M$ such that $f|_B\colon B \to N$ is an embedding. Informally speaking, immersions are embeddings with self-intersections. An enhanced definition of immersions will be discussed below, in 2.2.4.

1.3.3. Functions.

1.3.3.1. The ring of functions.

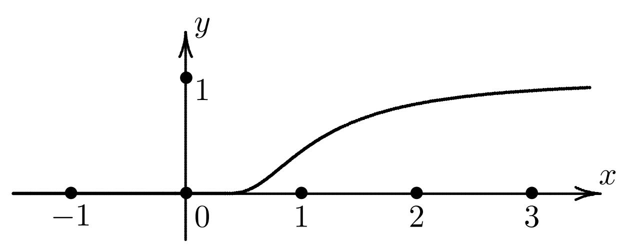
Smooth maps $M \to \mathbb{R}$ (where $\mathbb{R}$ is regarded as a 1-dimensional manifold) are called smooth functions. Smooth functions of M form a ring; this ring is denoted as $\mathcal{C}^\infty(M)$.[4]

1.3.3.2. Some auxiliary functions.

All functions considered below in this Section are smooth and take values in the closed interval $[0, 1]$.

1.3.3.2.1. Three functions on the line.

$$\rho(x) = \begin{cases} e^{-x^{-2}}, & \text{if } x > 0, \\ 0, & \text{if } x < 0 \end{cases}$$



[1] Homeomorphic smooth manifolds are not necessarily diffeomorphic. Examples of smooth manifolds homeomorphic, but not diffeomorphic, to S^7 were first constructed by J.Milnor in 1956. Examples of smooth manifolds homeomorphic, but not diffeomorphic, to $\mathbb{R}^4$ were constructed by S.Donaldson and M.Friedman in 1982.

[2] It is also true that homeomorphic manifolds have equal dimensions, but the proof requires a bit of algebraic topology.

[3] There exists a variety of embedding theorems. In particular, any smooth manifold may be embedded into an Euclidean space as a closed submanifold. Hence, the construction of 1.2.6.3.3 gives, actually, all existing manifolds, up to a diffeomorphism.

[4] There exists a theorem (due to Shields) that if, for two smooth manifolds M and N, the rings $\mathcal{C}^\infty(M)$ and $\mathcal{C}^\infty(N)$ are isomorphic, then the manifolds M and N are diffeomorphic.

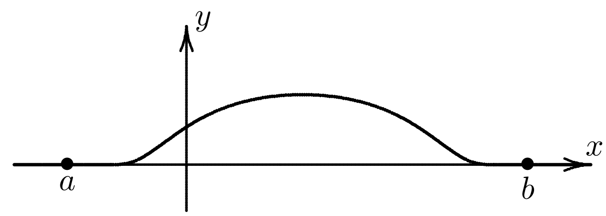
Notice that

$$\rho(x) = 0 \text{ for } x \leq 0,$$
$$\rho(x) > 0 \text{ for } x > 0,$$

and $\rho(0) = \rho'(0) = \rho''(0) = \rho'''(0) = \ldots = 0$.

For $a < b$,

$$\sigma_{a,b}(x) = \rho(x-a)\rho(b-x)$$

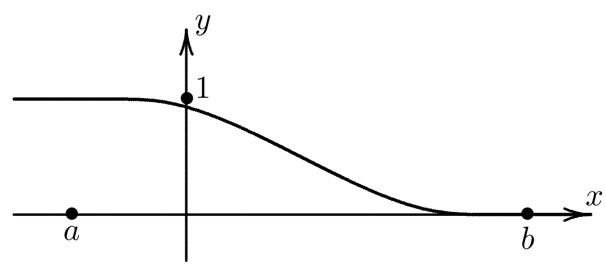


Notice that

$$\sigma_{a,b}(x) = 0 \text{ for } x \leq a \text{ and } x \geq b,$$
$$\sigma_{a,b}(x) > 0 \text{ for } a < x < b.$$

Also for $a < b$,

$$\tau_{a,b}(x) = \frac{\int_x^b \sigma_{a,b}(u)\,du}{\int_a^b \sigma_{a,b}(u)\,du}$$



Notice that

$$\tau_{a,b}(x) = 1 \text{ for } x \leq a,$$
$$0 < \tau_{a,b}(x) < 1 \text{ for } a < x < b,$$
$$\tau_{a,b}(x) = 0 \text{ for } t \geq b.$$

1.3.3.2.2. A function on $\mathbb{R}^n$.

For $R > r > 0$,

$$\lambda_{r,R}(x_1, \ldots, x_n) = \tau_{r,R}\left(\sqrt{x_1^2 + \ldots + x_n^2}\right).$$

This function is smooth, is equal to 1 within the ball of radius r centered at $0 \in \mathbb{R}^n$ and is equal to 1 in the complement of the ball of radius R centered at 0.

1.3.3.2.3. Functions on manifolds.

Proposition 1. *Let M be a smooth manifold, $p \in M$ be a point and $B \subset M$ be a neighborhood of p. There exists a smooth function on M taking values in $[0,1]$ and such that $f = 1$ in some neighborhood of p and $f = 0$ in the complement of B.*

Proof. Let (U, φ) be a chart of M such that $p = \varphi(0)$, let R be a positive number such that $\varphi^{-1}(B)$ contains the closed ball of radius R centered at 0, and let r be any number such that $0 < r < R$. The function $f\colon M \to \mathbb{R}$,

$$f(q) = \begin{cases} \lambda_{r,R}(\varphi^{-1}(q)), & \text{if } q \in \varphi(U), \\ 0, & \text{if } q \notin \varphi(U), \end{cases}$$

satisfies the conditions in Proposition.

Proposition 2. *Let M be a smooth manifold, $K \subset M$ be a compact set and $B \subset M$ be a neighborhood of K, that is, an open subset of M containing K. There exists a function on M such that $f = 1$ in some neighborhood of K and $f = 0$ in the complement of B.*

Remark. We will see below (in 1.3.3.4) that the assumption of compactness for K may be replaced by a weaker assumption of closedness.

Proof. For each p find a function $f_p: M \to \mathbb{R}$ as in Proposition 1, that is, equal to one in some neighborhood of p and equal to 0 in the complement of B. Let A_p be this neighborhood of p. Open sets A_p cover K, and, by the definition of a compact set, there exist a finite set $\{p_1, \dots, p_N\}$ of points of K such that the sets $A_{p_1}, \dots, A_{p_k}$ also cover K. The function $f = 1 - (1 - f_{p_1}) \dots (1 - f_{p_N})$ is equal to 1 in $\bigcup_{i=1}^N A_{p_i} \supset K$ and equal to 0 in the complement of B.

Corollary. *Let M be a smooth manifold, $K \subset M$ be a compact set and $B \subset M$ be a neighborhood of K. There exists a neighborhood A of K such that $\overline{A} \subset B$.*

Proof. Take a function f from Proposition 2 and put $A = \{q \in M \mid f(q) > 1/2\}$. Or take $A = \bigcup_{i=1}^N A_{p_i}$ from the previous proof.

1.3.3.3. Open covers.

1.3.3.3.1. Base of topology.

Below, we denote by $d^n_{x,r}$, or simply by $d_{x,r}$, an open ball in $\mathbb{R}^n$ with radius r centered at $x \in \mathbb{R}^n$. In the case when $x = 0$, the notation $d_{0,r}$ is abbreviated to d_r.

Let M be a smooth n-dimensional manifold. According to condition (1) in 1.1.4, M has a countable atlas, say $\{(U_i, \varphi_i)\}$. Consider all sets $\varphi_i(d_{x,r})$ where x has rational coordinates, r is rational and $d_{x,r} \subset U_i$. Obviously, these balls $d_{x,r}$ cover U_i and hence these sets $\varphi_i(d_{x,r})$ cover M. We will denote this cover by $\mathcal{W}$.

It is important that and any open subset U of M is the union of some sets from $\mathcal{W}$; namely, $U = \bigcup_{W \in \mathcal{W}, W \subset U} W$. (In point-set topology, a family of open sets with this property is called a base of topology.) Another important property of these sets is that they all have compact closures.

1.3.3.3.2. Some terminology.

Let $\mathcal{U}$ and $\mathcal{V}$ be two open covers of M. We say that $\mathcal{V}$ is a *subcover* of $\mathcal{U}$ if $\mathcal{V} \subset \mathcal{U}$; we say that $\mathcal{V}$ is *inscribed* in $\mathcal{U}$ if for any $V \in \mathcal{V}$ there exists a $U \in \mathcal{U}$ such that $V \subset U$; we say that $\mathcal{V}$ is *strongly inscribed* in $\mathcal{U}$ if for any $V \in \mathcal{V}$ there exists a $U \in \mathcal{U}$ such that $\overline{V} \subset U$.

An open cover $\mathcal{U}$ of M is called *compact* if the closures $\overline{U}$ are compact for all $U \in \mathcal{U}$ and is called *locally finite* if any point $p \in M$ has a neighborhood V such that only finite number of intersections $V \cap U, U \in \mathcal{U}$ is non-empty.

1.3.3.3.3. Countable open covers.

Proposition 1. *For any open cover $\mathcal{U}$ of M there exists a compact countable open cover of M strongly inscribed in $\mathcal{U}$.*

Proof. Let $\mathcal{W}_0$ be the subset of $\mathcal{W}$ consisting of those $W \in \mathcal{W}$ whose closures are contained in sets from $\mathcal{U}$:

$$\mathcal{W}_0 = \{W \in \mathcal{W} \mid \exists U \in \mathcal{U}, \overline{W} \subset U\}.$$

Let us prove that $\mathcal{W}_0$ is a cover of M. Let $p \in M$. Then $p \in U$ for some $U \in \mathcal{U}$. But since U is a union of sets from $\mathcal{W}$, there exists a $\widetilde{W} \in \mathcal{W}$ such that $p \in \widetilde{W} \subset U$. Let $\widetilde{W} = \varphi(d_{x,\tilde{r}})$ (see the definition of $\mathcal{W}$). Then $\varphi^{-1}(p) \in d_{x,\tilde{r}}$, and hence $\varphi^{-1}(p) \in d_{x,r}$ for some positive rational $r < \tilde{r}$. Put $W = \varphi(d_{x,r})$. Then $p \in W, W \in \mathcal{W}, \overline{W} \subset U$, that is $W \in \mathcal{W}_0$. Hence, $\mathcal{W}_0$ is a cover. It is countable, compact and strongly inscribed in $\mathcal{U}$ by construction.

Proposition 2. *Any open cover $\mathcal{U}$ of M contains a countable subcover.*

Proof. According to Proposition 1, there exists a countable open cover $\mathcal{V}$ of M inscribed in $\mathcal{U}$. For $V \in \mathcal{V}$ fix a $U_V \in \mathcal{U}$ such that $U_V \supset V$. Then $\bigcup_{V\in\mathcal{V}} U_V \supset \bigcup_{V\in\mathcal{V}} V = M$, that is, $\{U_V\}$ is a cover of M. It is a countable subcover of $\mathcal{U}$.

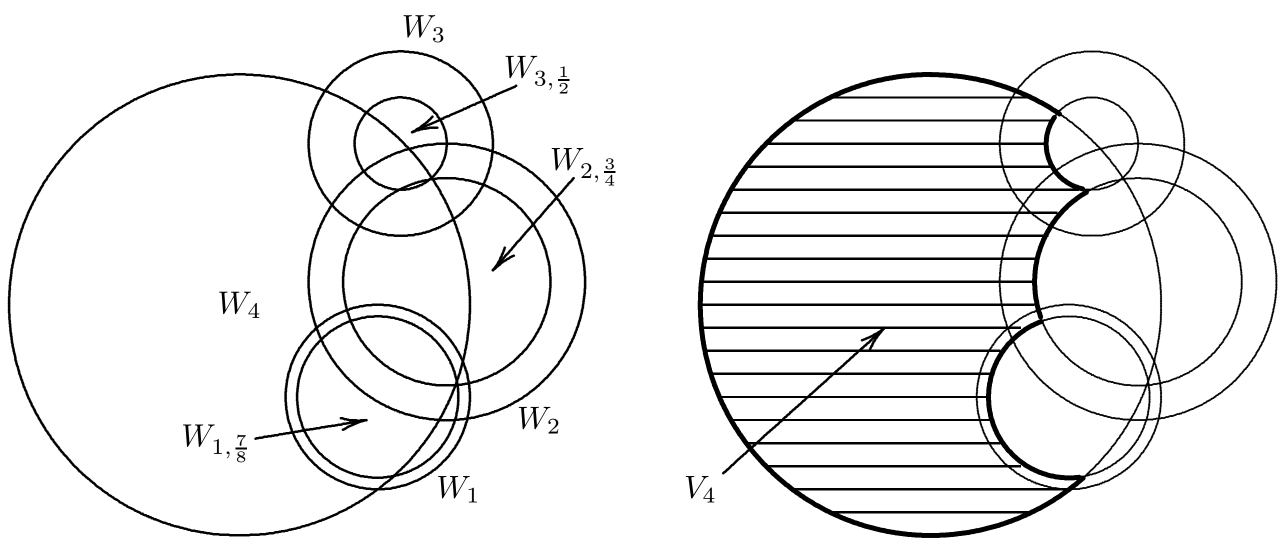
1.3.3.3.4. Paracompactness.

Theorem. *For any open cover $\mathcal{U}$ of M there exists a compact locally finite cover strongly inscribed in $\mathcal{U}$.*

This property of M (less the assumption of compactness for the subcover) is called the *paracompactness*.

Proof of Theorem. Let $\mathcal{W}_0 = \{W_i = \varphi_i(d_{x_i,r_i})\}$ be the compact countable cover of M strongly inscribed in $\mathcal{U}$ as constructed in the proof of Proposition 1 in 1.3.3.3.3. For $0 < t < 1$, put $W_{i,t} = \varphi_i(d_{x_i,tr_i})$. Furthermore, put

$$V_i = W_i - \left(\overline{W}_{i-1,\frac{1}{2}} \cup \overline{W}_{i-2,\frac{3}{4}} \cup \ldots \cup \overline{W}_{1,1-2^{1-i}}\right) = W_i - \bigcup_{k=1}^{i-1} \overline{W}_{k,1-2^{k-i}}$$



(this construction is shown schematically on the picture above). We will prove that $\mathcal{V} = \{V_i\}$ is a compact locally finite cover strongly inscribed in $\mathcal{U}$. We have four statements to prove (of which three are totally trivial).

(1) *$\mathcal{V}$ is a cover.* Let $p \in M$ and let $i = \min\{j \mid p \in W_j\}$. Then $p \in W_i$, $p \notin W_k$ for $k < i$, hence, $p \notin \overline{W}_{k,1-2^{k-i}}$ for $k < i$, hence, $p \in V_i$.

(2) *$\mathcal{V}$ is compact.* $\overline{V}_i \subset \overline{W}_i$. Since $\overline{W}_i$ is compact, $\overline{V}_i$ is also compact.

(3) $\mathcal{V}$ *is locally finite.* Let $p \in M$. Choose a $W_i = \varphi_i(d_{x_1,r_i}) \ni p$; let $y = \varphi_i^{-1}(p)$. Then there exist a neighborhood $G \subset d_{x_i,r_i}$ of y and a positive integer j such that $G \subset d_{x_i,(1-2-j)r_i}$. Let $V = \varphi_i(H)$; then V is a neighbourhood of p and $V \subset W_{i,1-2^{-j}}$. Hence, if $m \geq i + j$, then

$$V_m = W_m - (\ldots \cup \overline{W}_{i,1-2^{i-m}} \cup \ldots)$$

is disjoint from $V \subset W_{i,1-2^{-j}}$ since $1 - 2^{-j} \leq 1 - 2^{i-m}$ $(i - m \leq i - (i + j) = -j)$.

(4) $\mathcal{V}$ *is strongly inscribed in* $\mathcal{U}$, because it is inscribed in $\mathcal{W}$ which is strongly inscribed in $\mathcal{U}$.

1.3.3.3.5. Shrinking covers.

Proposition. *Let $\mathcal{U} = \{U_i\}$ be a countable compact locally finite open cover of M. There exists a (countable compact locally finite) open cover $\mathcal{V} = \{V_i\}$ of M such that $\overline{V}_i \subset U_i$ for all i.*

Remark. Actually this is true for any open cover of M, neither of the assumptions of compactness, countability, and local finiteness is needed.

Proof. Assume that for some $n \geq 1$ we already have open sets $V_1, \ldots, V_{n-1}$ such that $\overline{V}_i \subset U_i$ and $V_1, \ldots, V_{n-1}, U_n, U_{n+1}, \ldots$ cover M. (For $n = 1$ this assumption is empty and holds automatically.) Let $F = M - \left(\left(\bigcup_{i<n} V_i\right) \cup \left(\bigcup_{i>n} U_i\right)\right)$. Since the sets $V_i, i < n$, $U_i, i > n$, and U_n cover M, the set F is contained in U_n. Since $\overline{U}_n$ is compact, F is also compact. Hence, there exists an open set V such that $F \subset V, \overline{V} \subset U_n$ (see Proposition 2 of 1.3.3.2.3); take this set for V_n. Since

$$M = F \cup \left(\left(\bigcup_{i<n} V_i\right) \cup \left(\bigcup_{i>n} U_i\right)\right) \subset \left(\left(\bigcup_{i\leq n} V_i\right) \cup \left(\bigcup_{i>n} U_i\right)\right),$$

the sets $V_1, \ldots, V_{n-1}, V_n, U_{n+1}, U_{n+2}, \ldots$ cover M, which completes our induction: we obtain the sets $V_1, V_2, \ldots$ which satisfy the above assumption for each n. All we need to check is that these sets cover M. Let $p \in M$, and let $n = \max\{i \mid p \in U_i\}$ (the cover U_i is locally finite). Since the sets $V_1, \ldots, V_n, U_{n+1}, U_{n+2}, \ldots$ cover M and p does not belong to any of $U_{n+1}, U_{n+2}, \ldots$, it should belong to one of $V_1, \ldots, V_n$, and, thus, to some V_i.

1.3.3.4. Partitions of unity.

Let $f\colon M \to \mathbb{R}$ be a non-negative ($f(p) \geq 0$ for all $p \in M$) smooth function on M. Define the *support* of f as the set $\operatorname{supp} M = \overline{f^{-1}(0, \infty)}$; in other words, $p \in M$ *does not belong* to $\operatorname{supp} M$ if and only if f is equal to 0 in some neighborhood of p.

A family $\{F_\alpha\}$ of closed subsets of M is called locally finite, if every point $p \in M$ possesses a neighborhood V such that at most finitely many intersections $V \cap F_\alpha$ may be non-empty.

Lemma. *The union $F = \bigcup_\alpha F_\alpha$ of a locally finite family of closed sets is closed.*

Proof. Let $p \in M - F$. By definition of a locally finite set, there exists a neighborhood V of p and a finite set $\alpha_1, \ldots, \alpha_N$ such that $V \cap F_\alpha = \emptyset$, if $\alpha \neq \alpha_1, \ldots, \alpha_N$. Put $U = V - \bigcup_{i=1}^N F_{\alpha_i}$. The set U is open and $p \in U \subset M - F$. Thus, $M - F$ is open and F is closed.

A family $\{f_\alpha \mid \alpha \in A\}$ of non-negative functions on M is called locally finite, if the family $\operatorname{supp} f_\alpha$ is locally finite; in other words, the family f_α is locally finite, if any point $p \in M$ possesses a neighborhood V such that all the functions f_α, with a possible exception of finitely many of them, are equal to 0 within V. Obviously, if $\{f_\alpha \mid \alpha \in A\}$ is a locally finite family and $B \subset A$ then the sum $f_B = \sum_{\alpha \in B} f_\alpha$ is well-defined and is a non-negative smooth function.

THEOREM. *Let $\mathcal{U}$ be an arbitrary open cover of M. There exists a locally finite family of non-negative smooth functions $\{f_U \mid U \in \mathcal{U}\}$ such that*

(1) $\operatorname{supp} f_U \subset U$ *for any* $U \in \mathcal{U}$;
(2) $\sum_{U \in \mathcal{U}} f_U = 1$.

A family of functions with these properties is called a *partition of unity subordinated to* $\mathcal{U}$.

PROOF OF THEOREM. CASE 1: *$\mathcal{U}$ is countable, compact and locally finite.* According to Proposition in 1.3.3.3.5 (applied twice), there exist open covers $\mathcal{V} = \{V_U \mid U \in \mathcal{U}\}, \mathcal{W} = \{\mathcal{W}_{\mathcal{U}} \mid \mathcal{U} \in \mathcal{U}\}$ such that $\overline{V}_U \subset U, \overline{W}_U$ for all $U \in \mathcal{U}$. According to Proposition 2 in 1.3.3.2.3, there exist smooth functions g_U on M with values in $[0,1]$ such that g_U is 1 on W_U and 0 outside of V_U. The family $\{g_U\}$ is locally finite ($\operatorname{supp} g_U \subset \overline{V}_U \subset U$, and the cover $\mathcal{U}$ is locally finite), and the sum $G = \sum_U g_U$ is positive (actually, $G \geq 1$, since for each $p \in M$, $p \in W_U \Rightarrow f_U(p) = 1$ for at least one U). Put $f_U = \frac{g_U}{G}$; the functions f_U satisfy conditions (1) and (2) of Theorem.

CASE 2: *general.* Let $\mathcal{U}$ be an arbitrary open cover of M. According to Theorem in 1.3.3.3.4, there exists a countable compact locally finite cover $\mathcal{V}$ inscribed in $\mathcal{U}$. For $V \in \mathcal{V}$ fix $U_V \in \mathcal{U}$ that contains $\overline{V}$. As we already know (Case 1), there exists a partition of unity, $\{h_V\}$ subordinated to $\mathcal{V}$. For $U \in \mathcal{U}$, put $f_U = \sum_{V, U_V = U} h_V$. These are smooth functions taking values in $[0,1]$. Also, $\operatorname{supp} f_U \subset \overline{\bigcup_{V, U_V = U} \operatorname{supp} g_V} = \bigcup_{V, U_V = U} \operatorname{supp} g_V \subset \bigcup_{V, U_V = U} V \subset U$ (we use here Lemma) and $\sum_U f_U = \sum_U \sum_{V, U_V = U} g_V = \sum_V g_V = 1$.

COROLLARY 1. *Let M be a smooth manifold, $F \subset M$ be a closed set and $U \subset M$ be a neighborhood of F, that is, an open subset of M containing F. There exists a function on M such that $f = 1$ in some neighborhood of F and $f = 0$ in the complement of U.*

REMARK. Compare with Proposition 2 in 1.3.3.2.3.

PROOF OF COROLLARY 1. Let $V = M - F$. Then $\{U, V\}$ is an open cover of M. Let $\{f, g\}$ be a partition of unity subordinated to this cover. Then f is 0 outside U and g is 0 outside a closed set contained in V, hence is 0 in some neighborhood of F. Therefore, $f = 1 - g$ is 1 in some neighborhood of F.

COROLLARY 2. *Let U, V be open subsets of a manifold M such that $\overline{V} \subset U$. For any smooth function g on U there exists a smooth function h on M such that g and h agree on V.*

Proof. Corollary 1 yields a smooth function f on M which is 1 on $\overline{V}$ and 0 outside U. Take $h = \begin{cases} fg & \text{in } U, \\ 0 & \text{outside } U. \end{cases}$

2. Tangent vectors and vector fields.

2.1. Tangent vectors.

2.1.1. Definition.

Let M be a smooth manifold, and let $p \in M$. A *tangent vector* to M at p is, by definition, a linear map $\xi: \mathcal{C}^\infty(M) \to \mathbb{R}$ such that, for any $f, g \in \mathcal{C}^\infty(M)$, $\xi(fg) = \xi(f) \cdot g(p) + f(p) \cdot \xi(g)$. Tangent vectors to M at p form a real vector space (with respect to natural operations of addition and multiplication by real numbers); this vector space is called the *tangent space* to M at p and is denoted as T_pM.

PROPOSITION 1. *For any $\xi \in T_pM$ and any constant (function) c, $\xi(c) = 0$.*

Proof. $\xi(c) = \xi(c \cdot 1) = \xi(c) \cdot 1 + c \cdot \xi(1)$. On the other hand, the linearity of ξ shows that $\xi(c \cdot 1) = c \cdot \xi(1)$. Hence, $\xi(c) \cdot 1 = \xi(c) = 0$.

PROPOSITION 2. *If $f = 0$ in some neighborhood of p, then $\xi(f) = 0$ for any $\xi \in T_pM$.*

Proof. Let U be a neighborhood of p in which $f = 0$. Let $g \in \mathcal{C}^\infty(M)$ be a function such that $g(p) = 0$ and $g = 1$ in $M - U$. Then $fg = f$ and $\xi(f) = \xi(fg) = \xi(f)g(p) + f(p)\xi(g) = 0$ since $f(p) = g(p) = 0$.

COROLLARY. *If f and g agree within a neighborhood of p, then $\xi(f) = \xi(g)$.*

This follows immediately from Proposition 2.

In particular, we can apply ξ to functions defined only locally, in a neighborhood of p: Corollary 2 in 1.3.3.4 shows that such a function, restricted to a smaller neighborhood of p, can be extended to M, and Proposition 2 above shows that the value of ξ on the extended function does not depend on the extension.

2.1.2. Some lemmas from Analysis.

LEMMA. *Let $f \in \mathcal{C}^\infty(\mathbb{R}^n)$ and $f(0) = 0$. Then there exist functions $f_1, \ldots, f_n \in \mathcal{C}^\infty(\mathbb{R}^n)$ such that*

$$f(x_1, \ldots, x_n) = x_1 f_1(x_1, \ldots, x_n) + \ldots + x_n f_n(x_1, \ldots, x_n).$$

Proof. Put

$$f_i(x) = \int_0^1 \frac{\partial f}{\partial x_i}(tx)\, dt$$

($x = (x_1, \ldots, x_n)$). The standard theorems of analysis show that f_i are smooth functions. Furthermore,

$$\sum x_i f_i(x_1, \ldots, x_n) = \int_0^1 \sum x_i \frac{\partial f}{\partial x_i}(tx_1, \ldots, tx_n)\, dt = \int_0^1 \frac{d}{dt} f(tx_1, \ldots, tx_n)\, dt$$

$$= f(tx_1, \ldots, tx_n)\Big|_{t=0}^{t=1} = f(x_1, \ldots, x_n) - f(0, \ldots, 0) = f(x_1, \ldots, x_n).$$

GENERALIZED LEMMA. *Let U be an open subset of $\mathbb{R}^n$, $p = (p_1, \ldots, p_n) \in U$, $f \in \mathcal{C}^\infty(U)$, $f(p) = C$. Then there exist functions $f_1, \ldots, f_n \in \mathcal{C}^\infty(\mathbb{R}^n)$ such that, in some neighborhood of p in U,*

$$f(x_1, \ldots, x_n) = C + (x_1 - p_1) f_1(x_1, \ldots, x_n) + \ldots + (x_n - p_n) f_n(x_1, \ldots, x_n).$$

Proof. Using Corollary 2 in 1.3.3.4, find a function $g \in \mathcal{C}^\infty$ which agrees with f in some neighborhood of p. Then apply Lemma to the function

$$h(x_1, \dots, x_n) = g(x_1 + p_1, \dots, x_n + p_n) - C$$

(obviously, $h(0, \dots, 0) = 0$):

$$h(x_1, \dots, x_n) = \sum x_i h_i(x_1, \dots, x_n),$$
$$g(x_1, \dots, x_n) = C + h(x_1 - p_1, \dots, x_n - p_n) = C + \sum (x_i - p_i) h_i(x_1 - p_1, \dots, x_n - p_n)$$

and we can put $f_i(x_1, \dots, x_n) = h_i(x_1 - p_1, \dots, x_n - p_n)$.

2.1.3. Coordinates of tangent vectors.

Let $\{x_1, \dots, x_n\}$ be a local coordinate system in a neighborhood U of p with p having coordinates $(p_1, \dots, p_n)$. A smooth function on M may be regarded in U as a smooth function of $x_1, \dots, x_n$. Obviously, the map $\partial_i \colon \mathcal{C}^\infty(M) \to \mathbb{R}$ where $\partial_i(f)$ is the value of the partial derivative $\dfrac{\partial f}{\partial x_i}$ at p, is a tangent vector to M at p.

THEOREM. *The vectors $\partial_1, \dots, \partial_n \in T_pM$ form a basis of T_pM; in particular, $\dim T_pM = n = \dim M$.*

Proof. (1) $\partial_1, \dots, \partial_n$ are linear independent. Indeed, the coordinates $x_1, \dots, x_n$ are smooth functions in U. Obviously, $\partial_i(x_j) = \delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$ If $a_1\partial_1 + \ldots + a_n\partial_n = 0$, then $0 = (a_1\partial_1 + \ldots + a_n\partial_n)(x_j) = a_1\delta_{1j} + \ldots + a_n\delta_{nj} = a_j$ for each j.

(2) Any $\xi \in T_pM$ is a linear combination of ∂_i's. Put $a_i = \xi(x_j)$ and prove that $\xi = \sum a_i\partial_i$. Indeed, any $f \in \mathcal{C}^\infty(M)$ locally is $f(x_1, \dots, x_n)$, and, according to Generalized Lemma above,

$$f(x_1, \dots, x_n) = f(p) + \sum (x_i - p_i) f_i(x_1, \dots, x_n).$$

Hence

$$\begin{aligned} \xi(f) &= \xi(f(p)) + \sum \xi((x_i - p_i) f_i(x_1, \dots, x_n)) \\ &= 0 + \sum (\xi(x_i - p_i) f_i(p) + (p_i - p_i)\xi(f_i)) \\ &= \sum ((a_i - 0) f_i(p) + 0 \cdot \xi(f_i)) = \sum a_i f_i(p). \end{aligned}$$

On the other hand,

$$\begin{aligned} \partial_i(f) &= \partial_i(f(p)) + \sum_j \partial_i((x_j - p_j) f_j(x_1, \dots, x_n)) \\ &= 0 + \sum (\partial_i(x_j - p_j) f_j(p) + (p_j - p_j)\partial_i(f_j)) = \sum_j \delta_i^j f_j(p) = f_i(p) \end{aligned}$$

and

$$\left(\sum\nolimits_i a_i\partial_i\right)(f) = \sum_i a_i f_i(p).$$

Hence, $\xi(f) = \left(\sum_i a_i\partial_i\right)(f)$ for any $f \in \mathcal{C}^\infty(M)$, that is, $\xi = \sum a_i\partial_i$.

For $\xi \in T_pM$, the numbers a_i, such that $\xi = \sum a_i\partial_i$, are called *coordinates* of ξ with respect to the local coordinate system $x_1, \dots, x_n$.

2.1.4. Coordinate change.

Let $x_1, \dots, x_n$ and $x'_1, \dots, x'_n$ be two local coordinate systems in neighborhoods of p, let $\xi \in T_pM$, and let, correspondingly, $a_1, \dots, a_n$ and $a'_1, \dots, a'_n$ be the coordinates of ξ with respect to the two coordinate systems. Then $a'_i = \xi(x'_i) = \sum_j a_j \frac{\partial}{\partial x_j} x'_i\Big|_{x=p} = \sum_j \frac{\partial x'_i}{\partial x_j}(p)a_j$. In other words, the Jacobian matrix $\left\| \frac{\partial x'_i}{\partial x_j}(p) \right\|$ is the transition matrix from $\{a_1, \dots, a_n\}$ to $\{a'_1, \dots, a'_n\}$.

This observation gives rise to the classical, coordinate definition of a tangent vector. A tangent vector to M at p is a function assigning to each local coordinate system a sequence of n numbers, $a_1, \dots, a_n$, which are transformed according to the formula when one local coordinate system is replaced by another one.

2.1.5. The manifold of tangent vectors.

Let M be a smooth n-dimensional manifold, and let $TM = \bigcup_{p \in M} T_pM$ be the set of all tangent vectors of M. Let (U, φ) be a chart of M, and let $x_1, \dots, x_n$ be corresponding local coordinates in $\varphi(U)$. The product $U \times \mathbb{R}^n \subset \mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$ is an open set in $\mathbb{R}^{2n}$. Consider $\tilde\varphi\colon U \times \mathbb{R}^n \to TM$, $\tilde\varphi(q, (a_1, \dots, a_n)) = \sum_{i=1}^n a_i \frac{\partial}{\partial x_i}(p) \in T_pM$ where $p = \varphi(q)$. The pair $(U \times \mathbb{R}^n, \tilde\varphi)$ is a $2n$-dimensional chart of TM. Moreover, these charts are compatible: if $(U, \varphi), (V, \psi)$ are two charts of M and J is the Jacobian matrix of the map $\psi^{-1}\varphi\colon \varphi^{-1}\psi(V) \to \psi^{-1}\varphi(U)$, then $\tilde\varphi^{-1}\tilde\psi(V \times \mathbb{R}^n) = \varphi^{-1}\psi(V) \times \mathbb{R}^n$, $\tilde\psi^{-1}\tilde\varphi(U \times \mathbb{R}^n) = \psi^{-1}\varphi(U) \times \mathbb{R}^n$, and the map $\tilde\varphi\tilde\psi^{-1}$ acts as $(q, a) \mapsto (\varphi\psi^{-1}(q), Ja)$; obviously, it is smooth. Thus, TM becomes a smooth $2n$-dimensional manifold. Obviously, the projection $\pi\colon TM \to M$, $\pi(T_pM) = p$, is a smooth map.

2.2. Differentials of smooth maps.

2.2.1. Definition.

Let $h\colon M \to N$ be a smooth map between smooth manifolds, let $p \in M$, and let $q = h(p) \in N$. Define a map $d_ph\colon T_pM \to T_qN$ by the formula

$$((d_ph)(\xi))(f) = \xi(f \circ h) \text{ where } \xi \in T_pM, f \in \mathcal{C}^\infty(N).$$

Certainly, we need to check that $\eta = (d_ph)(\xi)$ belongs to T_qN; obviously, in is a linear map $\mathcal{C}^\infty(N) \to \mathbb{R}$, and

$$\begin{aligned}
\eta(fg) &= \xi(fg \circ h) = \xi((f \circ h) \cdot (g \circ h)) \\
&= \xi(f \circ h) \cdot (g \circ h)(p) + (f \circ h)(p) \cdot \xi(g \circ h) \\
&= \xi(f \circ h) \cdot (g(h(p)) + (f(h(p)) \cdot \xi(g \circ h) \\
&= \eta(f) \cdot g(q) + f(q) \cdot \eta(g).
\end{aligned}$$

Obviously, the map $d_p h\colon T_pM \to T_qN$ is linear. It is called the *differential* of the smooth map h at $p \in M$. It is also obvious that $d_p \,\mathrm{id}_M = \mathrm{id}_{T_pM}$ and, for any smooth maps $h\colon M \to N$ and $k\colon N \to P$, $d_p(k \circ h) = d_{h(p)}k \circ d_p h$.

The maps $d_p h$ compose a smooth map $dh\colon TM \to TN$, and again $d\,\mathrm{id}_M = \mathrm{id}_{TM}$ and $d(k \circ h) = dk \circ dh$.

2.2.2. Differentials of diffeomorphisms and open embeddings.

The last statements imply that if h is a diffeomorphism, then $d_p h$ is an isomorphism for any p. It is clear also that if U is an open subset of a manifold M, then $T_pU = T_pM$ for any $p \in U$. Together, these two propositions show that if $h\colon U \to M$ is an embedding with the image open in M, then $d_p h$ is an isomorphism for any $p \in U$. This can be applied, in particular, to the case of charts of M. First of all, for any $p \in \mathbb{R}^n$, we may identify $T_p\mathbb{R}^n$ with $\mathbb{R}^n$ using the correspondence $a_1\partial_1 + \ldots + a_n\partial_n \leftrightarrow (a_1, \ldots, a_n)$. This implies the identification $T_pU = \mathbb{R}^n$ for any open $U \subset \mathbb{R}^n$ and any $p \in U$. Now, if (U, φ) is a chart of M, then $d_p\varphi\colon T_pU \to T_{\varphi(p)}M$ becomes an isomorphism $\mathbb{R}^n \to T_{\varphi(p)}M$, which yields a basis in $T_{\varphi(p)}M$. This is the basis in the tangent space corresponding to a chart (a local coordinate system) as was described in 2.1.3.

For the manifolds of tangent vectors, the first statement means that if $h\colon M \to N$ is a diffeomorphism, that $dh\colon TM \to TN$ is also a diffeomorphism. Then we observe that $T\mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$ and, for an open subset U of $\mathbb{R}^n$, $TU = U \times \mathbb{R}^n$. For a chart (U, φ) of M, the differential $d\varphi$ is an open embedding $U \times \mathbb{R}^n \to TM$; this is a chart of TM as it was described in 2.1.5.

2.2.3. Differentials of embeddings.

Since for $m \leq n$, $\mathbb{R}^m$ is a subspace of $\mathbb{R}^n$, we may regard, for any $p \in \mathbb{R}^m$, the tangent space $T_p\mathbb{R}^m$ as a subspace of the tangent space $T_p\mathbb{R}^n$. Passing to open subsets, we see that for any open subset U of $\mathbb{R}^n$ and for any $p \in U \cap \mathbb{R}^m$, $T_p(U \cap \mathbb{R}^m)$ is a subspace of T_pU. Applying the differential of any chart of a smooth n-dimensional manifold N covering a point q of an m-dimensional submanifold M of N, we conclude that T_pM is a subspace of T_pN. This inclusion of T_pM into T_pN is, obviously, the differential of the inclusion map $M \to N$, and hence it does not depend of the choice of charts.

Also, TM is a submanifold of TN.

A vector $\xi \in T_pN$ is called *tangent to* M, if $\xi \in T_pM$. There are two convenient ways to identify vectors tangent to M.

Proposition 1. *A vector $\xi \in T_pN$ is tangent to M, if and only if $\xi(f) = 0$ for any function $f \in \mathcal{C}^\infty(N)$ such that $f|_M = 0$.*

Proof. Let $i\colon M \to N$ be the inclusion map. A vector $\xi \in T_pN$ is tangent to M, if and only if $\xi = (d_p i)(\eta)$ for some $\eta \in T_pM$, that is, if and only if $\xi(f) = \eta(f \circ i) = \eta(f|_M)$. Thus, if ξ is tangent to M and $f|_M = 0$, then $\xi(f) = 0$. Conversely, let $\xi(f) = 0$, if $f|_M = 0$. Any function $g \in \mathcal{C}^\infty(M)$ can be extended locally, in a neighborhood of p, to a smooth function on N (because any smooth function on $\mathbb{R}^m$ can be extended to a smooth function on $\mathbb{R}^n$). Let $\tilde{g}$ be this extended function. The value $\xi(\tilde{g})$ does not depend on the choice of extension (because if $\tilde{g}'$ is another extension, then $(\tilde{g}' - \tilde{g})|_{\mathbb{R}^m} = 0$ and $\xi(\tilde{g}') - \xi(\tilde{g}) = \xi(\tilde{g}' - \tilde{g}) = 0$). Put $\eta(g) = \xi(\tilde{g})$ obviously, $\eta \in T_pM$ and $(d_p i)(\eta) = \xi$.

Proposition 2. *Assume that M is described, at a neighborhood of p, by a non-degenerate system of equations $F_i = 0, i = 1, \ldots, n-m$ (see 1.2.4.2 – 1.2.4.4). A vector $\xi \in T_pN$ is tangent to M, if and only if $\xi(F_1) = \ldots = \xi(F_{n-m}) = 0$.*

Proof. Let $V = \{\xi \in T_pN \mid \xi(F_1) = \ldots = \xi(F_{n-m}) = 0\}$. We need to show that $V = T_pM$. Since $F_i|_M = 0$ for all i, Proposition 1 implies that $T_p(M) \subset V$. Let $\{x_1, \ldots, x_n\}$ be a local coordinate system on N in a neighborhood of p. The non-degeneracy of the system $\{F_i = 0\}$ means that there exist numbers $1 \leq k_1 < \ldots < k_{n-m} \leq n$ such that $\det \left\| \dfrac{\partial F_i}{\partial x_{k_j}} \right\| \neq 0$. This shows that no linear combination of $\partial_{k_1}, \ldots, \partial_{k_{n-m}}$ belong to V. Hence, $\operatorname{codim} V \geq n-m$, $\dim V \leq m$, and $V = T_pM$.

Example. A vector $(a_1, \ldots, a_{n+1}) = a_1\partial_1 + \ldots + a_{n+1}\partial_{n+1}$ is tangent to the sphere $S^n \subset \mathbb{R}^{n+1}$ at a point $(p_1, \ldots, p_{n+1}) \in S^n$ if and only if the function $(a_1\partial_1 + \ldots + a_{n+1}\partial_{n+1})((x_1)^2 + \ldots + (x_{n+1})^2 - 1) = 2(a_1x_1 + \ldots + a_{n+1}x_{n+1}$ vanishes at $(p_1, \ldots, p_{n+1})$, that is, if $a_1p_1 + \ldots + a_{n+1}p_{n+1} = 0$, that is, if the vector $(a_1, \ldots, a_{n+1})$ is orthogonal to the radius-vector $(p_1, \ldots, p_{n+1})$. We can also describe TS^n: it is a submanifold of $T\mathbb{R}^{n+1} = \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} = \mathbb{R}^{2n+2}$ composed of $(p, a) \in \mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$ for which $\|p\| = 1, \langle p, a\rangle = 0$.

Notice in conclusion, that since any embedding is a composition of a diffeomorphism and an inclusion of a submanifold, the differential d_ph of any embedding $h\colon M \to N$ is a monomorphism (one-to-one) and the differential $dh\colon TM \to TN$ is an embedding.

2.2.4. Differentials of immersions.

Since immersions locally are embeddings, differentials of immersions are also monomorphisms. Actually, we have more.

Proposition. *A smooth map $h\colon M \to N$ is an immersion, if and only if d_ph is a monomorphism for each $p \in M$.*

Proof. We need to prove that if $f\colon M \to N$, $\dim M = m, \dim N = n$ is a smooth map and d_pM is a monomorphism for each $p \in M$, then f is an immersion. Take a point $p \in M$ and choose local coordinates (aka charts) $\{x_1, \ldots, x_m\}$ and $\{y_1, \ldots, y_n\}$ in neighborhoods of p and $q = h(p)$. With respect to these coordinates, h is given by functions $y_j = h_j(x_1, \ldots, x_m)$, $j = 1, \ldots, n$. The fact that d_ph is a monomorphism means that the $m \times n$ matrix $J = \left\| \dfrac{\partial h_j}{\partial x_i} \right\|$ has a $m \times m$ submatrix $\widetilde{J}$ which has a non-zero determinant at p and hence in some neighborhood of p. Without loss of generality, we can assume that $\widetilde{J}$ is formed by first m columns of J. By Inverse Function Theorem, $x_1, \ldots, x_m$ can be locally expressed as functions of $y_1, \ldots, y_m$: $x_i = k_i(y_1, \ldots, y_m)$, $i = 1, \ldots, m$. Consider functions $z_1, \ldots, z_n$ of $y_1, \ldots, y_n$ in a neighborhood of q:

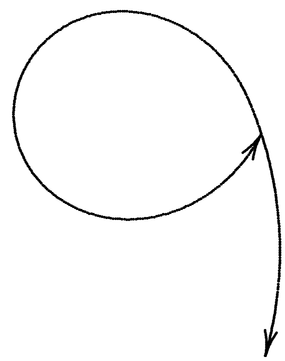
$$z_j = k_j(y_1, \ldots, y_m) \text{ for } j \leq m,\ z_j = y_j - h_j(k_1(y_1, \ldots, y_m), \ldots, k_m(y_1, \ldots, y_m)) \text{ for } j > m.$$

The Jacobian matrix of $z_, \ldots, z_n$ with respect to $y_1, \ldots, y_n$ has the form $\begin{bmatrix} \widetilde{J}^{-1} & 0 \\ * & I \end{bmatrix}$, and it has, in a neighborhood of q, a non-zero determinant. Hence, $z_1, \ldots, z_n$ provide a new local coordinate system (a new chart of N) in a neighborhood of q. Within this chart, a point $(y_1, \ldots, y_n) \leftrightarrow (z_1, \ldots, z_n)$ belongs to $h(M)$ if and only if $y_j =$

$h_j(k_1(y_1,\ldots,y_m),\ldots,k_m(y_1,\ldots,y_m))$, $j=1,\ldots,n$. But for $j\le m$ this holds automatically, while for $j>m$ this means precisely that $z_{m+1}=\ldots=z_n=0$. Thus, h locally is a diffeomorphism onto a submanifold, that is, h is an immersion.

2.2.5. Immersions and embeddings.

We know that embeddings are 1−1 immersions. On the other hand, multiple examples, like the one below, show that $1-1$ immersions are not necessarily embeddings.



Still the following is true.

Proposition. *A $1-1$ immersion of a compact manifold into any other manifold is an embedding.*

Proof. Let h be a $1-1$ immersion of a compact m-dimensiolnal manifold M into an n-dimensional manifold N. By definition of an immersion, for every point $p\in M$, there exist a neighborhood $U\subset M$ of p and a chart (V,ψ) of N such that $\psi^{-1}(h(U))=V\cap\mathbb{R}^m$. Since $M-U$ is compact, $h(M-U)$ is closed. Let $V'=V-\psi^{-1}(h(M-U))$. It is open, and $\psi^{-1}(h(M))\cap V'=\psi^{-1}(h(m-U)\cup f(U))\cap V'=\psi^{-1}(h(U))\cap V'=V'\cap\mathbb{R}^m$. Thus, $V',\psi|_{V'}$ is a chart required by definition of embedding.

2.2.6. Velocity vectors.

A (parametrized) *curve* γ in a manifold M is a smooth map of a finite or infinite open interval $I\subset\mathbb{R}$ into M. If $t_0\in I$, t is the coordinate on the line, and $p=\gamma(t_0)$, then the vector $d_{t_0}\gamma(d/dt)\in T_pM$ is called the *velocity vector* of γ at p. Notation: $\dot\gamma(t_0)$.

2.3. Vector fields.

2.3.1. Definition.

A *vector field* X on a manifold M is a linear map $X{:}\,\mathcal{C}^\infty(M)\to\mathcal{C}^\infty(M)$ satisfying the condition $X(fg)=X(f)g+fX(g)$ for any $f,g\in\mathcal{C}^\infty(M)$. (In the algebraic language: a vector field on M is a derivation of the ring $\mathcal{C}^\infty(M)$.) Obviously, vector fields on M compose a (generally, infinite dimensional) vector space; it is denoted as Vect(M).

It is also possible to multiply vector fields by functions. In the algebraic language, Vect(M) is a module over the ring $\mathcal{C}^\infty(M)$.

2.3.2. Values at points.

Let $X\in\mathrm{Vect}(M)$, and let $p\in M$. The map $X_p{:}\,\mathcal{C}^\infty(M)\to\mathbb{R}$, $X_p(f)=(X(f))(p)$ is a vector in T_pM. This vector is called the value of X at p. It is clear that if $X_p=Y_p$ for all $p\in M$, then $X=Y$. Thus, the function $p\mapsto X_p$ determines the vector field X.

Proposition. *For any vector field X, the map $M \to TM$, $p \mapsto X_p \in T_pM \subset TM$ is smooth. Moreover, any smooth map $M \to TM$ sending each point $p \in M$ into T_pM is determined by a vector field in the above way. Thus, vector fields on M are the same as smooth maps $X\colon M \to TM$ such that $\pi \circ X = id_M$.*

The proof is left to the reader who is advised to postpone it until Section 2.3.4 has been read.

2.3.3. Locality.

Lemma. *If a function $f \in \mathcal{C}^\infty(M)$ vanishes in an open set $U \subset M$ then $X(f)$ vanishes in U for any $X \in \mathrm{Vect}(M)$.*

Proof. Let $p \in U$. Choose a function $g \in \mathcal{C}^\infty(M)$ such that $g(p) = 0, g|_{M-U} = 1$. Then $fg = f$ and $X(f) = X(fg) = X(f)g + fX(g)$ which vanishes at p since $f(p) = g(p) = 0$.

Proposition. *The restriction $X(f)|_U$ is determined by the restriction $f|_U$.*

Proof. If $f|_U = g|_U$, then $(f-g)|_U = 0$, hence $X(f-g)|_U = 0$, hence $(X(f) - X(g))|_U = 0$, hence $X(f)|_U = X(g)_U$.

This shows that we can apply vector fields to functions defined only in open subsets of M, and the domain of $X(f)$ will be the same as the domain of f. Indeed, let $f \in \mathcal{C}^\infty(U)$. For a $p \in U$ choose open sets V, W such that $p \in V, \overline{V} \subset W, \overline{W} \subset U$, and then choose a function $g \in \mathcal{C}^\infty(M)$ such that $g = 1$ in V and $g = 0$ in $M - W$. Then the product fg is defined (and smooth) in the whole manifold M, and we put $(X(f))(p) = (X(fg))(p)$. The proposition above implies that this does not depend on any choices, so $X(f)$ becomes a valid function in U. Moreover, the equality $X(f) = X(fg)$ holds in V, so $X(f)$ is smooth. This map $\mathcal{C}^\infty(U) \to \mathcal{C}^\infty(U)$ is a vector field on U. We denote it as $X|_U$.

2.3.4. A coordinate presentation of vector fields.

Let $X \in \mathrm{Vect}(M)$, and let $x_1, \dots, x_n$ be local coordinates in some $U \subset M$. Then $x_1, \dots, x_n$ are smooth functions on U, and hence $X(x_i)$ are also well defined smooth function on U (see 2.3.3). Put $X(x_i) = X_i$.

Proposition. *For any $f \in \mathcal{C}^\infty(U)$, $X(f) = \displaystyle\sum_{i=1}^n X_i \frac{\partial f}{\partial x_i}$.*

Proof. For any $p \in U$, $(X(f))(p) = X_p(f) = \sum_i X_p(x_i) \dfrac{\partial f}{\partial x_i}(p)$ (see 2.1.3).

This shows that any vector field is presented, in local coordinates $x_1, \dots, x_n$, as $\sum_i X_i \dfrac{\partial}{\partial x_i}$. If $x'_1, \dots, x'_n$ is another local coordinate system, then, in the intersection of the domains of the two systems, X is presented both as $\sum_i X_i \dfrac{\partial}{\partial x_i}$ and as $\sum_i X'_i \dfrac{\partial}{\partial x'_i}$. In this intersection, $X'_i = \sum_j \dfrac{\partial x'_i}{\partial x_j} X_j$.

This observation gives rise to a coordinate description of vector fields, similar to that of tangent vectors given in 2.1.4. A vector field on M is a function assigning to each local coordinate system $x^1, \dots, x^n$ a sequence of n smooth functions in the domain of the coordinates, $X^1, \dots, X^n$, which are transformed by the formula $X'_i = \sum_j \dfrac{\partial x'_i}{\partial x^j} X^j$ under the coordinate change.

2.3.5. Commutators.

PROPOSITION 1. *Let X, Y be vector fields on M. Then*

$$[X,Y] = X \circ Y - Y \circ X \colon \mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M)$$

is also a vector field.

Proof.

$$\begin{aligned}
[X,Y](fg) &= X(Y(fg)) - Y(X(fg)) \\
&= X(Y(f)g + fY(g)) - Y(X(f)g + fX(g)) \\
&= X(Y(f))g + Y(f)X(g) + X(f)Y(g) + fX(Y(g)) \\
&\qquad - Y(X(f))g - X(f)Y(g) - Y(f)X(g) - fX(Y(g)) \\
&= (X(Y(f)) - Y(X(f)))g + f(X(Y(g)) - X(Y(g))) \\
&= [X,Y](f)g + f[X,Y](g)
\end{aligned}$$

REMARK. In general, the compositions $X \circ Y$ and $Y \circ X$ are not vector fields; only their difference is.

PROPOSITION 2 (coordinate presentation of commutators). *If, with respect to some local coordinate system $\{x_1, \dots, x_n\}$, $X = \sum_i X_i \dfrac{\partial}{\partial x_i}$ and $Y = \sum_j Y_j \dfrac{\partial}{\partial x_j}$, then $[X,Y] = \sum_k Z_k \dfrac{\partial}{\partial x_k}$ where $Z_k = \sum_i \left(X_i \dfrac{\partial Y_k}{x_i} - Y_i \dfrac{\partial X_k}{x_i} \right)$.*

Proof: a direct check (left to the reader).

PROPOSITION 3. *Commutators have the following properties.*

(i) *The operation $X, Y \mapsto [X,Y]$ is bilinear (over $\mathbb{R}$).*
(ii) $[X,Y] = -[Y,X]$.
(iii) $[[X,Y],Z] + [[Y,Z],X] + [[Z,X],Y] = 0$.
(iv) $[fX,Y] = f[X,Y] - Y(f)X$.

The proof is again left to the reader.

REMARK. Property (iii) is called *Jacobi identity*. It serves as a substitute to the associativity property which, in general, does not hold for commutators. In the algebraic language Properties (i)–(iii) mean that $\mathrm{Vect}(M)$ is, with respect to the vector space operation and the commutator operation, a *Lie algebra*.

2.3.6. Integral curves and flows.

2.3.6.1. Integral curves.

Let X be a vector field on M. A curve γ on M is called an *integral curve* (or *trajectory*) of X if for any t in the domain of γ the velocity vector of γ at $\gamma(t)$ is $X_{\gamma(t)}$. In the language of Analysis, a vector field is a (system of) differential equation(s) and an integral curve is a solution. Indeed, in local coordinates $x_1, \dots, x_n$, the vector field X is presented as $\sum_i X_i \dfrac{\partial}{\partial x_i}$ and the curve γ has parametric equations $x_i = x_i(t)$. The velocity vector of

γ at $\gamma(t_0)$ is $d_{t_0}\gamma(d/dt) = \dot{x}_i(t_0)$, and γ is an integral curve of X if and only if $\dot{x}_i(t_0) = X_i(x_1(t_0), \dots, x_n(t_0))$ for $i = 1, \dots, n$, in other words, if $x_1(t), \dots, x_n(t)$ is a solution of the system $\dot{x}_i = X_i$.

Proposition 1. *Let X be a vector field, and let $p \in M$. For some $\varepsilon > 0$, there exists a unique integral curve $\gamma: (-\varepsilon, \varepsilon) \to M$ of X with $\gamma(0) = p$.*

This is the theorem of the existence and uniqueness for solutions of ODE.

Notice that Proposition 1 is local: if we try to expand the integral curve beyond the interval $(-\varepsilon, \varepsilon)$, we encounter many different possibilities: we may get a curve $\mathbb{R} \to M$ whose image is a one-dimensional submanifold on M closed in topology of M and diffeomorphic to $\mathbb{R}$; or this curve may go to infinity in finite time (in any direction); or it may be closed (diffeomorphic to S^1; or it may be winding around a “limit cycle”; or it may end up at a point where X vanishes; or the closure of its image may be a torus of some dimension; or something else.

The choice of ε in Proposition 1 may be partially controlled.

Proposition 2. *Any $q \in M$ has a neighborhood U such that for all $p \in U$ the number ε from Proposition 1 may be taken the same. If M is compact, then ε may be taken the same for all $p \in M$.*

The first is obvious, the second follows from the definition of compactness. Actually, as we will see in next Section (Proposition 2 of 2.3.6.2), in the compact case ε can be taken arbitrary (or infinite).

2.3.6.2. Flows.

Let M be a smooth manifold and X be a vector field on M.

Proposition 1. *For each point $q \in M$ there exist an $\varepsilon > 0$, a neighborhood U of q, and a family of open embeddings $\{\varphi_t: U \to M \mid -\varepsilon < t < \varepsilon\}$ such that*

(1) *the map $\Phi: U \times (-\varepsilon, \varepsilon) \to M$, $\Phi(p, t) = \varphi_t(p)$ is smooth;*

for all $p \in M$

(2) *$\varphi_0(p) = p$;*
(3) *the curve $\gamma_p: (-\varepsilon, \varepsilon) \to M$, $\gamma_p(t) = \varphi_t(p)$, is an integral curve of X;*
(4) *if $\varphi_t(p) \in U$ and $|t + u| < \varepsilon$ then $\varphi_u(\varphi_t(p)) = \varphi_{t+u}(p)$.*

Proof. The existence of integral curves $\gamma_p: (-\varepsilon, \varepsilon) \to M$ for p in some neighborhood of q with $\gamma_p(0) = p$ is stated by Proposition 2 of 2.3.6.1. Put $\varphi_t(p) = \gamma_p(t)$. We already have Properties (2) and (3). Property (4) holds because of the uniqueness of an integral curve: both $u \mapsto \varphi_u(\varphi_t(p))$ and $u \mapsto \varphi_{t+u}(p)$ are integral curves of X with value $\varphi_t(p)$ for $u = 0$. The smoothness of the map $(p, t) \mapsto \varphi_t(p)$ (Property (1))is the theorem of the smooth dependence of solutions of ODE on the initial conditions. At last, to prove that φ_t's are open embeddings, we may need to shrink both U and ε: take a neighborhood $V \subset U$ of q and a positive $\varepsilon' \leq \varepsilon$ such that $\varphi_t(V) \subset U$ for $|t| < \varepsilon'$. Then for all such t the maps $\varphi_t: V \to \varphi_t(V)$ and $\varphi_{-t}: \varphi_t(V) \to V$ are mutually inverse smooth maps. Hence, they are diffeomorphisms, and $\varphi_t(V)$ is open.

Proposition 2. *Let M be compact. Then for any $X \in \mathrm{Vect}(M)$ there exists a family of diffeomorphisms $\{\varphi_t: M \to M \mid t \in \mathbb{R}\}$ such that*

(1) *the map* $\Phi: M \times \mathbb{R} \to M$ *is smooth;*

(2) $\varphi_t \circ \varphi_u = \varphi_{t+u}$*;*

for all $p \in M$*,*

(3) $\varphi_0(p) = p$*;*

(4) *the curve* $\gamma_p: \mathbb{R} \to M$*,* $\gamma_p(t) = \varphi_t(p)$*, is an integral curve of* X*.*

Proof. The existence of integral curves $\gamma_p: (-\varepsilon, \varepsilon) \to M$ with $\gamma_p(0) = p$ is stated by Proposition 2 of 2.3.6.1, and the only thing we need to add to the arguments of the previous proof is the extendability of the family φ_t to all real t. To do this we simply define φ_t with arbitrary t as $(\varphi_{t/n})^n$ where $|t/n| < \varepsilon$.

The family $\{\varphi: M \to M\}$ is called the *flow generated by* X. We will also refer to the family φ_t of open embeddings from Proposition 1 as to a local flow.

3. Critical points and critical values.

The word "generic" is almost as common (and almost as informal) in Mathematics as the word "obvious". We routinely state that a generic quadratic equation has two complex roots, that a generic curve in space has no inflection points, that a generic line (in space again) has precisely one common point with a given plane, and so on. Sometimes this has a formal sense (say, something occurs with a probability one), sometimes this simply discards certain events as unlikely. In differential Topology, the most trusted tool of rigorous establishing of genericity is provided by results related to the notions of critical and regular points and values.

3.1. Definitions.

Let M, N be smooth manifolds, let $\dim M = m$, $\dim N = n$, and let $f: M \to N$ be a smooth map. A point $p \in M$ is called *regular* with respect to f if $\operatorname{rank} d_p f = n$; a point is *critical*, if it is not regular. For example, $p \in M$ is a critical point for a function $f: M \to \mathbb{R}$ if $d_p f = 0$. Remark, that if $m < n$, then *every* point of M is critical with respect to f; to avoid this, visibly ugly, feature of the notion, people sometimes replace the condition $\operatorname{rank} d_p f = n$ by $\operatorname{rank} d_p f = \min(m, n)$; but for our purposes, the definition above is the best.

A point $q \in N$ is called a *critical value* of f, if $q = f(p)$ for some critcal point of f. A point $q \in N$ is called a *regular value* of M, if it is not a critical value of f. It may seem ridiculous, but according to this definition, every point of $N - f(M)$ becomes a regular value of f (that is, a regular value of f is not necessarily a value of f).

THEOREM. *If* $q \in N$ *is a regular value of a map* $f: M \to N$*, then* $f^{-1}(q)$ *is a submanifold of* M *of dimension* $\dim M - \dim N$*.*

Proof. Let (V, ψ) be a chart of N such that $0 \in V$ and $\psi(0) = q$. Then the coordinate functions of the map $F: f^{-1}(\psi(V)) \to V$, $F(p) = \psi^{-1}(f(p))$ form a system of $\dim N$ equations in a neighborhood of $f^{-1}(q)$ such that the set of solutions is $f^{-1}(q)$ and the gradients of left hand sides (with respect to local coordinates in M) are linearly independent at every point of $f^{-1}(q)$.

3.2. Sard's theorem.

This is one of the best known "general position theorems."

For an arbitrary map $f\colon M \to N$, can we expect that the set of regular point is in some sense ample? No, if, say, $\dim M < \dim N$, then there are no regular points at all. Another example: for a constant map $M \to N$, whatever the dimensions of M and N are, all points are critical. But for both these examples, the set of critical value is very thin: the image of a smooth map of a lower-dimensional manifold into a higher-dimensional one cannot be expected too big; the more so, for the constant map, there is only one value.

Sard's theorem provides a strong restriction for the set of critical values of a smooth map, rather than for the set of critical points. There are several statements. The weakest one (however, sufficient for many applications) states that any smooth map $f\colon M \to N$ (we always assume that $\dim N > 0$) has at least one regular value. A stronger statement says that the set of regular values is dense in N. There are two textbook statements of Sard's theorem. We formulate both.

THEOREM 1. *For any smooth map $f\colon M \to N$, the set of critical values has measure zero.*

THEOREM 2. *If M is compact, then the set of regular values of f is a dense open subset of N; in general case, the set of regular values of f is a countable intersection of dense open subsets of N.*

REMARK. As usual, "smooth" means $\mathcal{C}^\infty$ for us. Actually Sard's Theorem is true for finitely differentiable maps, but the number of continuous derivatives must be specified. Theorems 1 and 2 hold for $\mathcal{C}^r$ maps, if $r > \max(\dim M - \dim N, 0)$. There exists an example (due to D. Men'shov) of a $\mathcal{C}^1$ function $\mathbb{R}^2 \to \mathbb{R}$ for which every real number is a critical value.

In the next section, we will give a proof of Theorem 1 in a trivial, but important, case of $\dim M \le \dim N$. A full proof of Theorem 1 (borrowed from the book "Topology from the differential viewpoint") is contained in Appendix.

3.3. Proof for $\dim M \le \dim N$.

Denote the set of critical value of a smooth map g as $\mathrm{cv}(g)$.

First of all, it is known that the union of countable many sets of measure zero is again a set of measure zero. Since both manifolds M and N have countable atlases, the set of critical values of f is a countable union of sets $C_{(U\varphi),(V,\psi)} = \psi(\mathrm{cv}(\psi^{-1} \circ f \circ \varphi\colon \varphi^{-1}(f^{-1}(\psi(V))) \to V))$ where $(U, \varphi), (V, \psi)$ are charts of M, N. Thus, we can restrict ourselves to smooth maps between open subsets of Euclidean spaces; moreover, we can assume both bounded. Let $g\colon W \to Z$ is a smooth map between bounded open sets $W \subset \mathbb{R}^m$, $Z \subset \mathbb{R}^n$. If $m = n$, then critical points of g are the points where the determinant of the Jacobian matrix J_g of g is zero. For $\varepsilon > 0$, let $B_\varepsilon = \{p \in W \mid |\det J_g| < \varepsilon\}$; this is an open subset of W. Then $\mathrm{meas}(\mathrm{cv}(g)) = \mathrm{meas}(g(\{\det J_g = 0\})) \le \mathrm{vol}(g(B_\varepsilon)) \le \varepsilon \,\mathrm{vol}(B_\varepsilon) \le \varepsilon \,\mathrm{vol}(W)$, which certainly means that $\mathrm{meas}(\mathrm{cv}(g)) = 0$. The case $m < n$ (essentially, much more simple than the case $m = n$) is reduced to the case $m = n$. We just replace the map $g\colon W \to Z$ by the composition $\widetilde{g} = g \circ \pi\colon W \times d \to Z$ where d is an open ball of dimension $m - n$ and π is the projection of $W \times d$ onto W. The critical value of g (same as values of g) is the same as critical values of $\widetilde{g}$ (same as values of $\widetilde{g}$), and we can apply the previous result.

It should be added that, at least in the case of $m < n$, it is very easy to prove Sard's theorem in the form of Theorem 2; we do not need this.

3.4. First application: the embedding theorem.

Theorem (H. Whitney). *Any compact n-dimensional manifold M can be embedded in $\mathbb{R}^{2n+1}$ and immersed into $\mathbb{R}^{2n}$. Moreover, any smooth map $M \to \mathbb{R}^m$ may be $\mathcal{C}^\infty$ approximated by an embedding, if $m \geq 2n+1$ and by immersion, if $m \geq 2n$.*

Proof. First, we will establish that M can be embedded into some Euclidean space. The proof of this result (which, actually, is the most important part of Whitney's theorem) does not use Sard's theorem.

Let $\{(U_i, \varphi_I),\ i = 1, \ldots, r\}$ be a finite atlas of M, and $x_{i1}, \ldots, x_{in}: \varphi_i(U_i) \to$ be the corresponding local coordinates (that is, $x_{ik}(\varphi_i(x_1, \ldots, X_n)) = x_k$ for all $(x_1, \ldots, x_n) \in U_i$). Choose some open cover $\{V_i\}$ of M such that $\overline{V}_i \subset \varphi_i(U_i)$ and functions $\widetilde{x}_{ik}: M \to \mathbb{R}$ such that $\widetilde{x}_{ik} = x_{ik}$ within V_i (see 1.3.3.3.5).

Let $V = \bigcup_{i=1}^r (V_i \times V_i) \subset M \times M$; obviously, V is an open neighborhood of the diagonal $\operatorname{diag} = \{(p,q) \in M \times M \mid p = q\} \subset M \times M$. For $p, q \in M, p \neq q$, choose a smooth function $h_{pq}: M \to [0,1]$ such that $H_{pq}(p) = 1$, $h_{pq}(q) = 1$. Let $S_{pq} = h_{pq}^{-1}\left(\frac{1}{2}, 0\right]$, $T_{pq} = h_{pq}^{-1}\left[0, \frac{1}{2}\right)$; notice that if $p' \in S_{pq}$ and $q' \in T_{p,q}$, then $h_{p,q}(p') \neq h_{p,q}(q')$.

The open sets V, $S_{pq} \times T_{pq}$ cover $M, \times M$ (if $p = q$, then $(p,q) \in V$; if $p \neq q$, then $(p,q) \in S_{p,q} \times T_{p,q}$). Since $M \times M$ is compact, this cover has a finite subcover, $\{V,\ S_{p_jq_j} \times T_{p_jq_j},\ j = 1, \ldots, s\}$. Claim: $nr + s$ functions

$$\widetilde{x}_{ik}\,(1 \leq i \leq r,\ 1 \leq k \leq n), h_{p_jq_j}\,(1 \leq j \leq s): M \to \mathbb{R}$$

determine an embedding $F: M \to \mathbb{R}^N$, $N = nr + s$.

In virtue of Propositions in 2.2.4 and 2.2.5, to proof this, we need to check that F is $1-1$ and d_pF is a monomorphism for every $p \in M$. *Proof of the first.* Let $p \neq q$. If $p, q \in V_i$ for some i, then $\widetilde{x}_{ik}(p) \neq \widetilde{x}_{ik}(q)$ for some k (local coordinates x_{ik} distinguish p and q). If $(p,q) \notin V$, then $(p,q) \in S_{p_jq_j} \times T_{p_jq_j}$ for some j, and $h_{p_jq_j}(p) \neq h_{p_jq_j}(q)$. *Proof of the second.* If $p \in V_i$, then the composition $G = \{M \xrightarrow{F} \mathbb{R}^{nr+s} \xrightarrow{\pi} \mathbb{R}^n\}$ where

$$\pi(x_1, \ldots, x_{nr+s}) = (x_{n(i-1)+1}, \ldots, x_{ni})$$

is determined by the fucntions $\widetilde{x}_{i1}, \ldots, \widetilde{x}_{in}$ and hence coincides within V_i with the composition $V_i \subset \varphi_i^{-1}(U_i) \xrightarrow{\varphi_i^{-1}} U_i \subset \mathbb{R}^n$. Since the latter is an embedding, $d_pG = d_{F(p)}\pi \circ d_pF$ is a monomorphism, and hence d_pF is a monomorphism.

Thus, we have an embedding $F: M \to \mathbb{R}^N$. Now we apply Sard's theorem to reduce N. We want to find a line ℓ in $\mathbb{R}^N$ such that if $\pi_\ell: \mathbb{R}^N \to \mathbb{R}^{N-1}$ is a projection parallel to this line, then $\pi_\ell \circ F: M \to \mathbb{R}^{N-1}$ is an embedding (or an immersion). Consider the map $M \times M - \operatorname{diag} \to \mathbb{RP}^{N-1}$ which sends (p,q), $p \neq q$, into the line parallel to the vector $\overrightarrow{F(p)F(q)}\,(\neq 0)$. If $N > 2n+1$, then $\dim(M \times M - \operatorname{diag}) = 2n < N - 1 = \dim \mathbb{RP}^{N-1}$ and, according to Sard, the image Σ_1 of this map has measure 0. Let $T_1M \subset TM$ consists of tangent vectors to $F(M)$ of length 1; this is a manifold of dimension $2n-1$. Consider the map $T_1M \to \mathbb{RP}^{N-1}$ which sends a vector $v \in T_1M$ into a line parallel to v. If $N > 2n$, then $\dim T_1M = 2n-1 < N-1 = \dim \mathbb{R}P^{N-1}$, and the image Σ_2 of this map has measure 0.

If $\ell \notin \Sigma_1$, then $\pi_\ell \circ F$ is $1-1$ ($\pi_\ell(F(p)) = \pi_\ell(F(q))$ if and only if the chord $[F(p), F(q)]$ is parallel to ℓ). If $\ell \notin \Sigma_2$, then $\pi_\ell \circ F$ is an immersion (if the line through $F(p)$ parallel to ℓ is not tangent to $F(M)$, then π_ℓ is $1-1$ on $T_{F(p)}F(M)$). Iterating such projections, we get an embedding $M \to \mathbb{R}^{2n+1}$ and an immersion $M \to \mathbb{R}^{2n}$.

To finish the proof, we need to check the "Moreover" statement. Let $h: M \to \mathbb{R}^m$ be some smooth map. Consider the map $h \times F: M \to \mathbb{R}^{m+N}$ where F is the map defined above. Applying projections along lines as above, we can reduce $h \times F$ to an embedding $M \to \mathbb{R}^m$, if $m \geq 2n+1$ or to immersion $M \to \mathbb{R}^m$, if $m \geq 2n$. But the lines used in these projections are chosen from dense subsets of projective spaces. Therefore we can make the composite projection $\mathbb{R}^{m+N} \to \mathbb{R}^m$ arbitrarily close to the projection of the product $\mathbb{R}^m \times \mathbb{R}^N$ onto the first factor ($(x, y) \mapsto x$). Then the composition of this map with $h \times F$ will be arbitrarily close (in the $\mathcal{C}^\infty$ sense) to h. This completes the proof.

REMARKS. The assumption of compactness of M is not needed: the whole statement holds for arbitrary manifolds. Only the first part of proof is significantly different for the non-compact case. The same can be said about the real analytic case.

The Sard arguments work even in algebraic geometry. For example, every (non-singular) projective complex algebraic variety of dimension n may be algebraically embedded into $\mathbb{C}P^{2n+1}$. In particular, every complex algebraic curve may be realized as a non-singular complex curve in the complex projective space, but not, in general, in the complex projective plane. This is expectable: a complex algebraic curve may have an arbitrary "genus" g (that is, be diffeomorphic to a sphere with g handles), while a non-singular curve determined in the complex projective plane by an equation of some degree d, and then the genus equals $\dfrac{(d-1)(d-2)}{2}$.

In differential topology, there are many results concerning embeddings in Euclidean spaces of smaller dimensions. For example, any manifold of a positive dimension n can be embedded in $\mathbb{R}^{2n}$ (but the approximation as stated above does not hold for such embeddings). If the (positive) dimension of a manifold is not a power of 2, or if is orientable, then an embedding in $\mathbb{R}^{2n-1}$ is possible. And so on.

3.5. Transversal regularity.

Sard's theorem is not the only general position theorem in differential topology. For an example of a different result (which is not an automatic corollary of Sard's theorem, although ultimately follows from it) we mention the "transversal regularity theorem" due to R. Thom.

Let P be a submanifold of a manifold N. A map $f: M \to N$ is called transversely regular (or t-regular) with respect to P, if for every point $p \in f^{-1}(P) \subset M$, $T_{f(p)}N = T_{f(p)}N + d_pf(T_pM)$. For example, if P is one point then f is transversely regular with respect to P if and only if P is a regular value of f.

THEOREM 1. *If f is transversely regular with respect to P, then $f^{-1}(P)$ is a submanifold of M of dimension* $\dim M + \dim P - \dim N$.

This theorem is similar to Theorem in 3.1, generalizes that theorem and is proved in a similar way.

THEOREM 2 (Thom). *Every smooth map $M \to N$ may be $\mathcal{C}^\infty$ approximated by smooth maps transversely regular with respect to P.*

For example, if $\dim M < \dim N - \dim P$, this means that every map $M \to N$ can be approximated by maps whose image is disjoint from P.

4. Manifolds with boundaries.

4.1. Main definition.

Important notation:

$$\mathbb{R}^n_- = \{(x_1, \dots, x_n) \mid x_n \leq 0\}.$$

Let us repeat all said in Section 1.1 replacing $\mathbb{R}^n$ by $\mathbb{R}^n_-$. With a minor exception concerning orientations in the case of $\dim M = 1$ (which we will discuss below, in 4.2.2), everything makes sense. We need only to adjust the terminology. The words *chart* and *atlas* may be replaced, if necessary, by the words *∂-chart* and *∂-atlas*. The word *manifold* becomes *manifold with boundary*, or *∂-manifold*.

4.2. Interior and boundary. Relations between orientations.

4.2.1. Definitions.

Let M be an n-dimensional manifold with boundary. Among the charts (U, φ) of (the maximal atlas of) M there are charts with U open in $\mathbb{R}^n$ (that is, disjoint from $\mathbb{R}^{n-1} \subset \mathbb{R}^n_-$). The union of images of such charts is denoted as $\operatorname{Int} M$; it is a dense open subset of M, and it is an n-dimensional manifold in the sense of our previous definition. It is called the *interior* of M. The difference $M - \operatorname{Int} M = \partial M$ is an $(n-1)$-dimensional manifold (if $\{(U, \varphi)\}$ is an atlas of M, then $\{(U \cap \mathbb{R}^{n-1}, \varphi|_{U \cap \mathbb{R}^{n-1}})\}$ is an atlas of ∂M. The manifold ∂M is called the *boundary* of M.

Notice that *a boundary of a ∂-manifold is a manifold without boundary*: $\partial\partial M = \emptyset$.

There is an alternative description of the boundary. A point $p \in M$ belongs to ∂M if and only if there is a chart (U, φ) of M such that $p = \varphi(q)$ for some $q \in U \cap \mathbb{R}^{n-1}$. (In particular, $\partial \mathbb{R}^n_- = \mathbb{R}^{n-1}$.)

REMARK. The equivalence of these two descriptions may seem obvious, but it requires a lemma that an open neighborhood in $\mathbb{R}^n_-$ of a $q \in \mathbb{R}^{n-1}$ is not homeomorphic to an open subset of $\mathbb{R}^n$. This follows from a more general fact that a subset of $\mathbb{R}^n$ homeomorphic to an open subset of $\mathbb{R}^n$ is open itself. The last proposition, however obvious it may seem, requires a bit of algebraic topology for a proof. If we replace homeomorphisms by diffeomorphisms, then the last will follow from standard theorems of calculus (what we need, is that diffeomorphisms are open maps). We leave details to the reader.

4.2.2. Orientations.

LEMMA. *Let $n > 1$, and let $(U, \varphi), (V, \psi)$ be two charts of a ∂-manifold M covering a point $q \in \partial M$. The charts $(U, \varphi), (V, \psi)$ of M are orientably compatible at q if and only if the charts $(U \cap \mathbb{R}^{n-1}, \varphi|_{U \cap \mathbb{R}^{n-1}}), (V \cap \mathbb{R}^{n-1}, \psi|_{V \cap \mathbb{R}^{n-1}})$ of ∂M are orientably compatible at q.*

Proof. It is convenient here to use the language of local coordinates. Let $(x_1, \dots, x_n)$, $(x'_1, \dots, x'_n)$ be local coordinates corresponding to the charts $(U, \varphi), (V, \psi)$ of M. Then $(x_1, \dots, x_{n-1}), (x'_1, \dots, x'_{n-1})$ are local coordinates corresponding to the charts $(U \cap \mathbb{R}^{n-1}$,

$\varphi|_{U\cap\mathbb{R}^{n-1}}), (V\cap\mathbb{R}^{n-1}, \psi|_{V\cap\mathbb{R}^{n-1}})$ of ∂M. We have the following: at q, $x_n = x'_n = 0$; within $\varphi(U)$, $x_n \le 0$; within $\psi(V)$, $x'_n \le 0$; within $\varphi(U)\cap\psi(V)$, if $x_n = 0$, then $x'_n = 0$ (and vice versa). Let J be the Jacobian matrix of transition from $(x_1, \dots, x_n)$ to $(x'_1, \dots, x'_n)$ at q, and J_∂ be the Jacobian matrix of transition from $(x_1, \dots, x_{n-1})$ to $(x'_1, \dots, x'_{n-1})$ at q. Then $J = \begin{bmatrix} J_\partial & 0 \\ * & a \end{bmatrix}, a > 0$ (meaning that $\dfrac{\partial x'_n}{\partial x_i} = 0$ for $i < n$ and $\dfrac{\partial x'_n}{\partial x_n} > 0$). Hence $\det J_\partial$ and $\det J = a \cdot \det J_\partial$ have the same sign.

This lemma shows that if M is orientable, then ∂M is also orientable, and every orientation of M gives rise to an orientation if ∂M. Also, M and $\operatorname{Int} M$ are orientable simultaneously, and there is a natural $1-1$ correspondence between their orientations.

The case $\dim M = 1$ requires a separate consideration. If one applies the definition given in 1.1.6 to 1-dimensional ∂-manifold, the results may be disastrous. For example, any (connected) charts of $M = [0,1]$ covering the points $0, 1$ induce opposite orientations on $\operatorname{Int} M = (0,1)$; so we have to admit that the manifold $[0,1]$ is not orientable at all! To avoid this nonsense, we have to modify the definition of orientation of 1-dimensional ∂-manifolds. We say that an orientation of a 1-dimensional ∂-manifold M is *by definition* an orientation of $\operatorname{Int} M$. Furthermore, in the case $\dim M = 1$, we cannot even speak of relations between orientations of M and ∂M since we never defined orientations for 0-dimensional manifolds. But we will need them, so let us introduce the following, maybe, unexpected, definition. An orientation of a 0-dimensional manifold N is, by definition, a function $N \to \{+, -\}$. (Thus, all 0-dimensional manifolds are orientable, and a manifold consisting of r points has 2^r orientations.) If M is an oriented 1-dimensional ∂-manifold then we define an orientation of ∂M as taking the value "+" on $q \in \partial M$, if a (connected) chart of M covering q is orientably compatible with the orientation of $\operatorname{Int} M$, and the value "−" otherwise.

4.3. Overview of previous results from the point of view of manifolds with boundary.

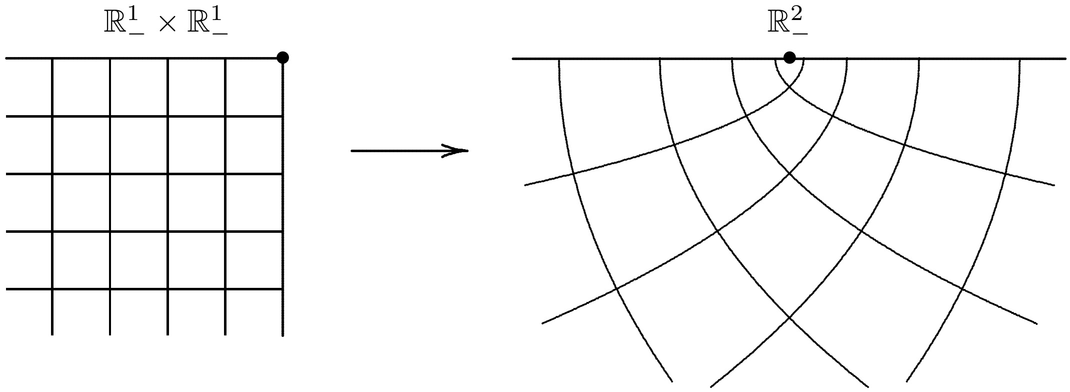
The most part of the theory developed above can be directly extended to the boundary case. We will provide a brief overview below and will point out the most essential differences.

4.3.1. Examples.

Let us begin with $D^n = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_1^2 + \ldots + x_n^2 = 1\}$. Obviously, $\partial D^n = S^{n-1}$. There are many examples of geometric nature; for instance, there is the *closed* Möbius band; its boundary is diffeomorphic to the circle S^1. If a codimension 1 submanifold N of M is given by the equation $F = 0$ where $F\colon M \to \mathbb{R}$ is a function such that no point of N is critical, that $M_+ = \{p \in M \mid F(p) \ge 0\}$ is a manifold with boundary N. For example, there is a "drilling a hole" construction: take an n-dimensional manifold (without a boundary) M, a chart (U, φ) of M, and a small open ball $d \subset U$; then $M - \varphi(d)$ is a ∂-manifold with the boundary diffeomorphic to S^{n-1}. Remark that if the manifold M is connected, then the result of the drilling a hole construction does not depend on the choice of d.

4.3.2. Products.

If M, N be manifolds one of which, say, M, has empty boundary, then $M \times N$ has an obvious structure of a ∂-manifold with $\partial(M \times N) = M\partial N$. Similarly, if $\partial N = \emptyset$, then $\partial(M \times N) = (\partial M) \times N$. The case when both M and N have non-empty boundaries is less automatic, since there is no diffeomorphism $\tau: \mathbb{R}^m_- \times \mathbb{R}^n_- \to \mathbb{R}^{m+n}_-$. Since $\mathbb{R}^m_- = \mathbb{R}^{m-1} \times \mathbb{R}^1_-$ and $\mathbb{R}^n_- = \mathbb{R}^{n-1} \times \mathbb{R}^1_-$, it is sufficient to fix an identification between $\mathbb{R}^1_- \times \mathbb{R}^1_-$ and $\mathbb{R}^2_-$. We do this by means of the map $(x, y) \mapsto (x^2 - y^2, -2xy)$ (originated from squaring of complex numbers), see the picture on the next page. For charts $(U, \varphi), (V, \psi)$ of M, N, we construct a chart $(\tau(U \times V), (\varphi \times \psi) \circ \tau^{-1})$ of $M \times N$, and these charts form an atlas which makes $M \times N$ a ∂-manifold. Furthermore, $\partial(M \times N)$ is the union of $(\partial M) \times N$ and $M \times \partial N$ with the intersection of these two ∂-manifold being their common boundary, $\partial M \times \partial N$.



4.3.3. Submanifolds.

The notion of a submanifold in the boundary case is not well established. It is clear how to define an m-dimensional ∂-submanifold N of a manifold M without boundary: for every point of $p \in N$ there must be a chart (U, φ) of M such that $p \in \varphi(U)$ and $\varphi^{-1}(N) = U \cap \mathbb{R}^m_-$. Similarly, N is a submanifold without boundary of a ∂-manifold M, if it is a submanifold of $\operatorname{Int} M$ in the sense of the definition in 1.2.6.3.1. If both M and N are ∂-manifolds, then we can either request that N is a submanifold of $\operatorname{Int} M$, or request that $\partial N \subset \partial M$ with appropriate conditions on charts of M coverings points in ∂N; we will skip a more detailed discussion.

4.3.4. Smooth maps.

All said in Section 1.3 can be repeated in the boundary case without serious modifications. For ∂-manifolds, there are notions of smooth maps, diffeomorphisms, embeddings, and immersions. Diffeomorphism must take boundaries into boundaries, while smooth maps do not have to do this (although one can make such an assumption). The property of paracompactness, as well as related properties, hold for ∂-manifolds.

4.3.5. Tangent vectors and vector fields.

The theory of Section 2 is also extended to the case of ∂-manifolds. One should notice that, for an n-dimensional manifold M with boundary, and the point $p \in \partial M$, the space T_pM is n-dimensional and contains $T_p\partial M$ as an $(n-1)$-dimensional subspace. Every non-zero vector from T_pM either is tangent to ∂M, or is directed *inside* M, or is directed *outside* M. TM is a manifold with boundary, and $\partial(TM) = \bigcup_{p \in \partial M} T_pM$. Vector fields are defined as before, but trajectories and flows are defined only for vector fields X such that for every $p \in \partial M$, the vector X_p is either tangent to ∂M, or is directed inside M.

4.3.6. Embeddings into Euclidean spaces.

PROPOSITION. *Let M be a compact manifold with boundary. Then, for some N, there exists an embedding $F: M \to \mathbb{R}^N_-$ such that $F(\partial M)$ is contained in $\mathbb{R}^{N-1}$ and for every $q \in \partial M$, $d_q(T_qM)$ is transverse to $\mathbb{R}^{N-1}$. Moreover, we can make $d_qF(T_qM)$ perpendicular to $\mathbb{R}^{N-1}$.*

Proof. First, let us construct a smooth function $h: M \to \mathbb{R}^1_-$ without critical points on ∂M, such that $h|_{\partial M} = 0$ and $h|_{\mathrm{Int}\, M} < 0$. For this purpose, we consider a finite atlas $\{(U_i, \varphi_i)\}$ of M, and fix a partition of unity $\{f_i: M \to \mathbb{R}\}$ subordinated to the covering $\varphi_i(U_i)$. Then the function $h = \sum_i f_i(x_n \circ \varphi_i^{-1})$ (where x_n is regarded as a function on U_i) satisfies our conditions.

Next, we consider an embedding of M into a Euclidean space as constructed in the first part of the proof in 3.4 and add h as an additional (last) coordinate function. we get an embedding of M into some $\mathbb{R}^N_-$ as required. To achieve *perpendicularity* of $d_qF(T_qM)$ to $\mathbb{R}^{N-1}$, it is sufficient to replace the function x_n in the previous construction by $-\sqrt{-x_n}$.

REMARK. It is possible to make this theorem a full replica of Theorem in 3.4 (regarding the dimensions and approximation); we will not need this, and will not do this.

4.4. Collars and attaching ∂-manifolds along boundaries.

Now we turn to properties and constructions which exist only in the boundary case.

4.4.1. Definition of a collar.

Let M be a ∂-manifold. A *collar* of M is a (smooth) embedding $c: \partial M \times [0, \varepsilon) \to M$ ($\varepsilon > 0$) such that $c(p, 0) = p$ $\forall p \in \partial M$. It is always possible to *narrow* a collar by restricting to $\partial M \times [0, \varepsilon')$, $0 < \varepsilon' < \varepsilon$.

4.4.2. Existence.

PROPOSITION. *If ∂M is compact, M possesses a collar.*

Proof. Fix a locally finite atlas $\{(U_i, \varphi_i)\}$ of M, and in every $\varphi_i(U_i)$ take the vector field $X_i = -\dfrac{\partial}{\partial x_n}$. Then put $X = \sum_i f_i X_i$ where $\{f_i\}$ is a partition of unity subordinated to $\{\varphi_i(U_i)\}$. This is a vector field directed inside M at every point of ∂M. It generates a flow $\alpha_t : M \to M$; since ∂M is compact, there exists an ε such that in a neighborhood of α_t is defined on ∂M for $0 \le t < \varepsilon$. Put $c(p, t) = \alpha_t(p)$. This is a collar, at least on $\partial M \times [0, \varepsilon')$, possibly with $\varepsilon' < \varepsilon$.

REMARK. Collars exist without the assumption of compactness of ∂M.

4.4.3. Uniqueness.

PROPOSITION. *If ∂M is compact, then for every two collars $c_1, c_2: \partial M \times [0, \varepsilon) \to M$, there exist a diffeomorphism $f: M \to M$ which is the identity on ∂M and outside any given neighborhood of ∂M and such that for some $\varepsilon' < \varepsilon$, $f(c_1(p, t)) = c_2(p, t)$ $\forall p \in \partial M$, $t \le \varepsilon'$.*

Proof. We may assume that M is compact (if it is not, we can replace M by $c_1(\partial M \times [0, \delta])$ for some positive $\delta < \varepsilon$). Choose an embedding $F: M \to \mathbb{R}^N_-$ as stated in Theorem of 4.3.6 (including the perpendicularity condition). Then $(N - n)$-dimensional planes perpendicular to tangent planes to $F(M)$ are disjoint in some neighborhood of $F(M)$ in $\mathbb{R}^N_-$ and cover this neighborhood; hence they determine a projection π of this neighborhood onto

$F(M)$. Take small $\varepsilon_1, \varepsilon_2$ such that $0 < \varepsilon_1 < \varepsilon_2 < \varepsilon$ and put $\tau(t) = \tau_{\varepsilon_1,\varepsilon_2}(t)$ (see 1.3.3.2.1). Then put $f \circ F(c_1(p,t)) = \pi(\tau(t)F(c_2(p,t)) + (1-\tau(t))F(c_1(p,t))$. This is diffeomorphism defined in some neighborhood of $F(\partial M)$, it is the identity outside a smaller neighborhood of $F(\partial M)$ (that is, where $t > \varepsilon_2$) and takes $F(c_1(p,t))$ into $F(c_2(p,t))$ within a still smaller neighborhood of ∂M (where $t < \varepsilon_1$). (Certainly, ε_1 and ε_2 should be sufficiently small for that.) Then we extend f to the whole $F(M)$, and the diffeomorphism $f\colon F(M) \to F(M)$ regarded as a diffeomorphism $f\colon M \to M$ fits into Proposition.

4.4.4. Attachments.

4.4.4.1. Attachment along the boundary.

Let M' and M'' be n-dimensional ∂-manifolds with diffeomorphic boundaries, and let $h\colon \partial M' \to \partial M''$ be a diffeomorphism. In the disjoint union $M' \sqcup M''$, merge for all $p \in \partial M'$ points p and $h(p)$. We want to equip the resulting set M with a structure of an n-dimensional manifold without boundary. To do this, we will need collars $c' : \partial M' \times [0,\varepsilon) \to M'$ and $c'' : \partial M'' \times [0,\varepsilon) \to M''$.

Let $\{(U'_\alpha, \varphi'_\alpha)\}$, $\{(U''_\beta, \varphi''_\beta)\}$, and $\{(V, \psi)\}$ be atlases of M', M'', and $\partial M'$. We denote as ι_1, ι_2 the canonical inclusions $M' \to M$, $M'' \to M$ and compose an atlas of M of the following charts: $(U'_\alpha, \iota_1 \circ \varphi'_\alpha)$, $(U''_\beta, \iota_2 \circ \varphi''_\beta)$, and $(V \times (-\varepsilon,\varepsilon), \overline{\psi})$ where (for $q \in V$, $t \in (-\varepsilon,\varepsilon)$)

$$\overline{\psi}(q,t) = \begin{cases} \iota_1 \circ c'(\psi(q), t), & \text{if } t \geq 0, \\ \iota_2 \circ c''(h \circ \psi(q), -t), & \text{if } t \leq 0 \end{cases}$$

(obviously, the two formulas give the same for $t = 0$).

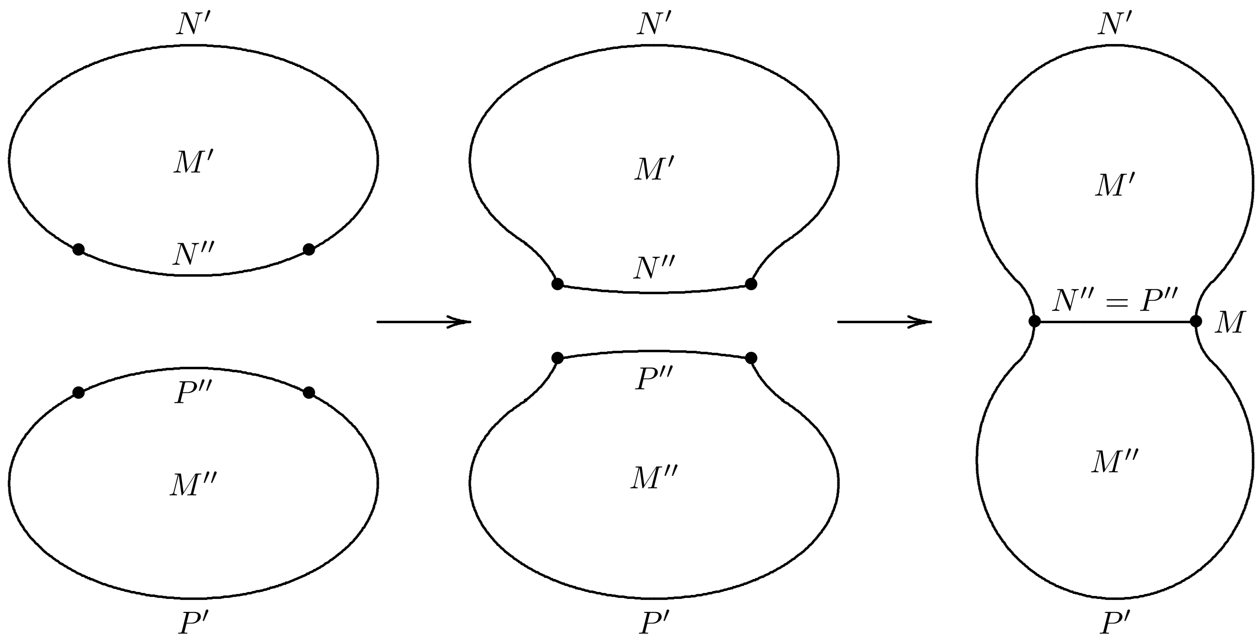
This manifold structure on M surely depends on the collars, but Proposition of 4.4.3 implies that up to a diffeomorphism it depends only on M', M'', and h. By this reason, we may use for the manifold M the notation $M' \cup_h M''$. Notice that $M' \cup_h M$ can be described axiomatically, as an n-dimensional manifold with an $(n-1)$-dimensional submanifold N diffeomorphic to $\partial M'$ such that $M - N$ is the union of two open and closed subsets, P' and P'', such that $P' \cup N$ is diffeomorphic to M' and $P'' \cup N$ is diffeomorphic to M''.

The attaching construction has obvious generalizations to the case when h connects subsets of $\partial M'$ and $\partial M''$ composed of whole components, or even disjoint subsets of the boundary of the same manifold, again composed of the whole components.

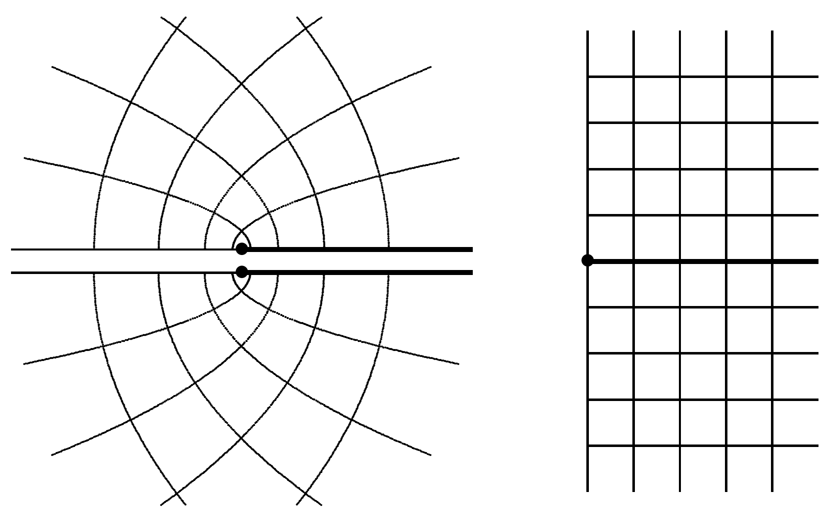
Examples. If $M' = M'' = D^n$ and $h = \mathrm{id}$, then $M' \cup_h M''$ is (diffeomorphic to) S^n. If M' is the closed Möbius band, $M'' = D^2$, and h is an arbitrary diffeomorphism between their boundaries (both are diffeomorphic to S^1), then $M' \cup_h M''$ is $\mathbb{R}P^2$. If $M' = M''$ is a Möebius band, and $h = \mathrm{id}$, then $M' \cup_h M''$ is a Klein bottle. If $M' = M'' = S^1 \times D^1$, then $M' \cup_h M''$, dependingly on h, is either a torus, or a Klein bottle.

4.4.4.2. Attachment along a piece of the boundary.

We will need the attachment operation in a more general setting. Let M', M'' be two manifolds (of the same dimension) with boundaries. Suppose that the both boundaries, $\partial M'$ and $\partial M''$, are obtained by the previous attachment operation: $\partial M' = N' \cup N''$, $\partial M'' = P' \cup P''$, $N' \cap N'' = \partial N' = \partial N''$, $P' \cap P'' = \partial P' = \partial P''$. Suppose also that there fixed a diffeomorphism $h\colon N'' \to P''$. The result of our construction is a manifold M with boundary $N' \cup_{\partial h} P'$ (where $\partial h\colon \partial N'' \to \partial P''$ is the restriction of h). The construction is schematically shown on the picture below (next page).



The smooth ∂-manifold structure on M is defined everywhere except $\partial N'' = \partial P'' \subset \partial M$ in the obvious way (we need collars for both M' and M''). To define a chart in a neighborhood of a point of $\partial N'' = \partial P''$, we need to fix an identification of the union of two copies of $\mathbb{R}^n_-$ glued along $\mathbb{R}^{n-1}_-$ onto $\mathbb{R}^n_-$. This shown schematically on the picture below. The formula from the map from the right to the left is $(x, y) \mapsto (x^2 - y^2, 2xy)$ (the squaring of complex numbers).



5. Morse theory.

5.1. Morse lemma.

Let $f(x_1, \ldots, x_n)$ be a smooth function defined in a neighborhood of a point $p \in \mathbf{R}^n$. We assume that p is a critical point of f, that is, $\dfrac{\partial f}{\partial x_1}(p) = \ldots \dfrac{\partial f}{\partial x_n}(p) = 0$.

5.1.1. Non-degenerate critical points. Indices.

We say that the critical point p is *non-degenerate* or *Morse*, if $\det \left\| \dfrac{\partial^2 f}{\partial x_i \partial x_j}(p) \right\| \neq 0$. The symmetric matrix $\left\| \dfrac{\partial^2 f}{\partial x_i \partial x_j}(p) \right\|$ is a matrix of a quadratic form $\sum_{i,j} \dfrac{\partial^2 f}{\partial x_i \partial x_j}(p) u_i u_j$, and it is known from linear algebra, that after an appropriate linear coordinate change $(u) \to (v)$, this form becomes $\pm v_1^2 \pm \ldots \pm v_m^2$ where the numbers of pluses and minuses do not depend on the choice of the coordinate change. Moreover, since the matrix of the quadratic form is non-singular, $m = n$. The number of the minus signs in the last expression is called the index of the quadratic form, and we call it also the index of the critical point p. It is known from calculus, that a critical point of index 0 is a point of a local minimum, a critical point of index n is a point of a local maximum, and critical points of other indices are (different kinds of) saddlepoints.

It is obvious (well known) that neither the non-degeneracy of a critical point, nor the index, depend on the coordinate choice.

5.1.2. The statement.

Theorem (Morse Lemma). *Let p be a non-degenerate critical point of index k of a smooth function f. There exists a diffeomorphism φ of a neighborhood of 0 onto a neighborhood of p such that*

$$f \circ \varphi(y_1, \ldots, y_n) = f(p) + y_1^2 + \ldots + y_{n-k}^2 - y_{n-k+1}^2 - \ldots - y_n^2.$$

Comments. In other words, after a (possibly, non-linear) coordinate change, the function f becomes a quadratic form plus a constant. If one adds to the left hand side of the last formula "plus higher order terms", then the statement becomes a conjunction of two standard theorems: the Taylor theorem and the theorem of the canonical form of quadratic forms (mentioned above). It is important, that after an appropriate coordinate change, the function f becomes quadratic not up to higher order terms, but precisely. Certainly, this would not be true without a non-degeneracy condition. For the simplest example, let $n = 1$ and $f(x) = x^3$. For this function, 0 is a degenerate critical point. Its quadratic approximation at 0 is just 0. But no coordinate change can make this function *equal* to 0.

5.1.3. Proof.

Our proof mimics the proof of the theorem of the canonical form of quadratic forms from old textbooks of Linear algebra. More modern textbooks, however, usually contain a different proof of the quadratic form theorem, so it is unlikely that the reader would readily recognize it.

We will assume (without any loss of generality) that $p = 0$ and $f(p) = 0$. All the formulas below valid in some neighborhood of 0, we never mention it, but always mean it.

According to lemma of 2.1.2, $f(x) = \sum_i x_i f_i(x)$ where f_i's are smooth functions defined in a neighborhood of 0. Obviously, $\dfrac{\partial f}{\partial x_i}(0) = f_i(0)$, hence $f_1(0) = \ldots = f_n(0) = 0$. Applying the same lemma to f_i, we get $f_i(x) = \sum_j x_j f_{ij}(x)$, and hence $f(x) =$

$\sum_{i,j} x_i x_j f_{ij}(x)$; we can assume that $f_{ij}(x) = f_{ji}(x)$ (if necessary, we can take $\dfrac{f_{ij}+f_{ji}}{2}$ for the new f_{ij}). Notice that $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(0) = 2f_{ij}(0)$.

First, consider the case when $f_{11}(0) > 0$. Put (for $(x_1, \ldots, x_n)$ in a neighborhood of the origin)

$$y_1 = x_1\sqrt{f_{11}} + \sum_{j=2}^{n} x_j \frac{f_{1j}}{\sqrt{f_{11}}},\ y_2 = x_2, \ldots, y_n = x_n.$$

The Jacobian $\det \left\| \dfrac{\partial y_i}{\partial x_j}(0) \right\|$ is equal to $\sqrt{f_{11}(0)} \neq 0$. Hence, we can consider $x_1, \ldots, x_n$ (in a neighborhood of the origin) as functions of $y_1, \ldots, y_n$. Also,

$$f(y_1, \ldots, y_n) = y_1^2 + g(y_1, \ldots, y_n) \text{ where } g(y_1, \ldots, y_n) = \sum_{i,j=2}^{n} y_i y_j g_{ij}(y_1, \ldots, y_n).$$

The function $f(y)$ still has a non-degenerate singularity at 0. Moreover, for every fixed y_1, the function $g(y_1, \ldots, y_n)$ (as a function of $y_2, \ldots, y_n$) has a non-degenerate singularity at $(y_2, \ldots, y_n) = (0, \ldots, 0)$. Indeed, $\dfrac{\partial g}{\partial y_j}(y_1, 0, \ldots, 0) = 0$ for $j \geq 2$ (obviously), and

$$\frac{\partial^2 f}{\partial y_1^2}(y_1, 0, \ldots, 0) = 2,\ \ \frac{\partial^2 f}{\partial y_1 \partial y_j}(y_1, 0, \ldots, 0) = 0, \text{if } j \geq 2,$$

$$\frac{\partial^2 f}{\partial y_i \partial y_j}(y_1, 0, \ldots, 0) = \frac{\partial^2 g}{\partial y_i \partial y_j}(y_1, 0, \ldots, 0) = g_{ij}(y_1, 0, \ldots, 0),\ \text{if } i, j \geq 2.$$

$$\text{Hence, } \det \left\| \frac{\partial^2 g}{\partial y_i \partial y_j}(y_1, 0, \ldots, 0) \right\| = \frac{1}{2} \det \left\| \frac{\partial^2 f}{\partial y_i \partial y_j}(y_1, 0, \ldots, 0) \right\| \neq 0.$$

The case when $f_{11}(0) < 0$ is similar: we take $\sqrt{-f_{11}}$ instead $\sqrt{f_{11}}$ and get $f(y_1, \ldots, y_n) = -y_1^2 + g(y_1, \ldots, y_n)$ with the same g as above. The case when $f_{11}(0) = 0$, but $f_{kk} \neq 0$ for some k, is reduced to the previous case by a renumbering of the coordinates. The last case is $f_{11}(0) = \ldots = f_{nn}(0) = 0$. Still, in this case some $f_{k\ell}(0)$ must be not zero (since $\det \|f_{ij}(0)\| \neq 0$; after a renumerating of coordinates, we may assume that $f_{12}(0) \neq 0$. In this case, we make a coordinate change $x_1 = x_1' + x_2',\ x_2 = x_1' - x_2',\ x_j = x_j'$ for $j \geq 3$ and get $f(x) = \sum_{i,j} x_i x_j f_{ij}(x) = \sum_{i,j} x_i' x_j' f_{ij}'(x'),\ f_{11}' = f_{11} + f_{22} + 2f_{12} \neq 0$; thus, this case is reduced to a case considered above.

Thus, in all cases, we find a coordinate change $(x) \to (y)$, after which the function $f(x_1, \ldots, x_n)$ becomes $\pm y_1^2 + g(y_1, \ldots, y_n)$ where for every fixed y_1 (close to 0!), the function $g(y_1, y_2, \ldots, y_n)$ of $n-1$ variables $y_2, \ldots, y_n$ satisfies the same conditions as the fucntion f of n variables. The same arguments as before show that there is a coordinate change $(y_2, \ldots, y_n) \to (z_2, \ldots, z_n)$ after which $g(y_1, y_2, \ldots, y_n)$ becomes $\pm z_2^2 + h(z_1, \ldots, z_n)$ (and $f(x_1, \ldots, x_n)$ becomes $\pm z_1^2 \pm z_2^2 + h(z_1, \ldots, z_n)$; we put $z_1 = y_1$) and for every fixed z_1, z_2

the function $h(z_1, z_2, z_3, \ldots, z_n)$ of $n-2$ variables $z_3, \ldots, z_n$ satisfies the same conditions as the fucntion f.

Proceeding this way, we arrive at the statement of Theorem.

(Notice that at every step, beginning from the second, we use a stronger form of Lemma of 2.1.2: if $f(x_1, \ldots, x_n, y_1, \ldots, y_m)$ is a smooth function in a neighborhood of $0 \in \mathbb{R}^{n+m}$, and $f(0, \ldots, 0, y_1, \ldots, y_m) = 0$ (for all y), then $f(x_1, \ldots, x_n, y_1, \ldots, y_m) = \sum_{i=1}^n x_i f_i(x_1, \ldots, x_n, y_1, \ldots, y_m)$; the proof is the same as in 2.1.2.)

5.1.4. Simple corollaries.

Proposition 1. *If p is a non-degenerate critical point of a function f, then in some neighborhood of p, f has no other critical points.*

Proof. We may assume that $p = 0$, $f(p) = 0$. Then, by the Morse Lemma, in some neighborhood of $0, f$ is $\pm x_1 \pm \ldots \pm x_n$. The latter has no critical points different from 0.

Proposition 2. *If all critical points of f are non-degenerate, then in any compact subset of the domain of f, the function f has finitely many critical points.*

Proof. The set of critical points is a closed (obviously) discrete (by Proposition 1) set. Hence, its intersection with any compact set is finite.

Proposition 3. *Suppose that for some compact set $K \subset \mathbb{R}^n$, all the critical points of a smooth function f in K are non-degenerate. Then there exists an $\varepsilon > 0$ such that if a smooth function g differs in K from f less that by ε, and the same is true for partial derivatives of f an g of order ≤ 2. Then all the critical points of g in K are also non-degenerate.*

Proof. The (finite) set of critical points of f in K has two neighborhoods, U and V, $\overline{U} \subset V$, with the following properties: for some $\delta > 0$, (1) in V, the absolute value of the determinant $\det \left\| \frac{\partial^2 f}{\partial x_i \partial x_j} \right\|$ is greater that δ; (2) in the complement of U in K, the sum $\left| \frac{\partial f}{\partial x_1} \right| + \ldots + \left| \frac{\partial f}{\partial x_n} \right|$ is greater than δ. If ε is small enough, then the same inequalities, with $\delta/2$ instead of δ, hold for g. This means that all the critical points of g in K belong to U, and all of them are non-degenerate.

5.2. Morse functions.

5.2.1. Definition.

A smooth function on a manifold (without boundary) is called a *Morse function*, if all its critical points are non-degenerate. Notice that the property of a function on a compact manifold to be a Morse function is "$\mathcal{C}^2$-open". This means that if a sequence $\{f_n\}$ of smooth function on M converges to a Morse function uniformly with partial derivatives of order ≤ 2 (with respect to local coordinates from some finite atlas), then the functions f_n, with finitely many exceptions, are Morse functions. This follows from Proposition 3 in 5.1.4.

The goal of the Morse theory is to derive as much information as possible about the manifold from he behavior of critical points/values of one Morse function. Example of a result: if a compact manifold possesses a Morse function with just two critical points (there should be at least two critical point: a maximum and a minimum), then the manifold is homeomorphic to a sphere.

5.2.2. Existence.

THEOREM. *Every compact manifold possesses a Morse function.*

We give below two different proofs of this theorem. Both use the Sard theorem (for maps between manifolds of the same dimension), and, within these notes, it is the most important application of the Sard theorem. Both proofs can be modified to the non-compact case.

5.2.2.1. The first proof of existence.

LEMMA. *Let $f: D^n \to \mathbb{R}$ be a smooth function. There exists a dense open set $A \subset \mathbb{R}^n$ such that if $a = (a_1, \ldots, a_n) \in A$, then the function $f_a: D^n \to \mathbb{R}$,*

$$f_a(x) = f(x) - \langle a, x \rangle = (x_1, \ldots, x_n) - a_1 x_1 - \ldots - a_n x_n$$

is a Morse function.

Proof. Take for A the set of regular values of the "gradient map" $\nabla f: D^n \to \mathbb{R}^n$ with coordinate functions $\dfrac{\partial f}{\partial x_1}, \ldots \dfrac{\partial f}{\partial x_n}$. It is open and dense because its complement is closed (obviously) and has measure zero by the Sard theorem. First, $x = (x_1, \ldots, x_n)$ is a critical point of f_a precisely if $(\nabla f)(x) = a$. Second, the matrix of second derivatives, $\left\| \dfrac{\partial^2 f}{\partial x_i \partial x_j} \right\|$ is precisely the Jacobian matrix of ∇f. Hence, f_a has a degenerate critical point if and only if a is a critical value of ∇f.

Proof of Theorem. We denote as d_r and open ball of radius r in $\mathbb{R}^n$ centered at 0. Let $\{(\mathbb{R}^n, \varphi_i)\ i = 1, \ldots, N\}$ be a sequence of charts of M such that $U_i = \varphi_i(d_1)$ cover M. We will construct inductively functions $f_0, f_1, \ldots, f_N: M \to \mathbb{R}$ such that f_k has no degenerate critical points in $\bigcup_{i=1}^k U_i$; then f_N will be a Morse function on M. We put $f_0 = 0$. Suppose that for some $k < n$, f_k has already been constructed. For $a \in \mathbb{R}^n$, define a smooth function $h_a: M \to$ by the formula

$$h_a(p) = \begin{cases} f_k(p), & \text{if } p \in M - \varphi_{k+1}(d_2), \\ f_k(p) - \lambda_{1,2}(\varphi_{k+1}^{-1}(p)) \cdot \langle a, \varphi_{k+1}^{-1}(p) \rangle, & \text{if } p \in \varphi_{k+1}(\mathbb{R}^n) \end{cases}$$

(see 1.3.3.2.2 for the definition of λ). Because of the openness property of the Morse condition (see 5.1.1), f a is sufficiently close to 0, then h_a has no degenerate critical points in $\bigcup_{i=1}^k U_i$; by Lemma, for a from a dense open subset of $\mathbb{R}^n$, h_a has no degenerate critical points in U_{k+1}. We choose an a satisfying the both conditions, and put $f_{k+1} = h_a$.

5.2.2.2. The second proof of existence.

In this proof, we assume that M is a submanifold of $\mathbb{R}^N$ (which is not a restrictive assumption, at least for compact manifolds, see Section XXX).

For a non-zero $\xi \in \mathbb{R}^N$, consider the function $f_\xi: M \to \mathbb{R}$, $f_\xi(x) = \langle x, \xi \rangle$ (where $\langle\ \rangle$ is the standard dot-product in $\mathbb{R}^N$). Obviously, this is a smooth function; this is our candidate to a Morse function on M.

Let $P = \{(\xi, x) \in (\mathbb{R}^N - 0) \times M \mid \xi \perp T_x M\}$. It is obvious that P is an N-dimensional submanifold of $(\mathbb{R}^N - 0) \times M$ (indeed, the perpendicularity condition may be expressed as a system of n independent equation; actually, we do not need this argument, since below we will introduce and use local coordinates on P). Let $\pi\colon P \to \mathbb{R}^N - 0,\ \rho\colon P \to M$ be the restrictions to P of the projections of the product $(\mathbb{R}^N - 0) \times M$ onto the factors. We will prove the following two propositions.

Proposition 1: *x is a critical point of the function f_ξ, if and only if $\xi \perp T_x M$, that is, if $(\xi, x) \in P$.*

Proposition 2. *A critical point x of the function f_ξ is degenerate if and only if (ξ, x) is a critical point of the map $\pi\colon P \to \mathbb{R}^N - 0$.*

(Proposition 1 may be regarded as obvious; still, its proof is contained in our computations below.)

The two propositions show that the function f_ξ has degenerate critical points if and only if ξ is a critical value of π; that is, f_ξ is a Morse function if and only if ξ is a regular value of π. Thus, the Sard theorem shows that there are (many) Morse functions on M.

Proof of Propositions. Let $x^0 \in M$. In some neighborhood U of x^0, M can be presented by a system of equations, $F_i(x_1, \ldots, x_N) = 0,\ 1 \le i \le N - n$ with rank $\left\| \dfrac{\partial F_i}{\partial x_j} \right\| = N - n$. We can assume, without loss of generality, that the leftmost $(N-n) \times (N-n)$ minor, $\det \left\| \dfrac{\partial F_i}{\partial x_j} \right\|_{1 \le j \le N-n}$ is not zero. Then, by the implicit function theorem, in some neighborhood of x^0 (which may be smaller than U, but which will be still denoted as U), the system $\{F_i(x_1, \ldots, x_N) = 0\}$ can be solved in $x_1, \ldots, x_{N-n}$: $x_i = f_i(x_{N-n+1}, \ldots, x_N)$. In other words, the system $\{F_i(x_1, \ldots, x_N) = 0\}$ is equivalent (locally, in a neighborhood U of x^0) to the system $f_i(x_{N-n+1}, \ldots, x_N) - x_i = 0$, $1 \le i \le N - n$, and we can redefine the functions F_i as $F_i(x_1, \ldots, x_N) = f_i(x_{N-n+1}, \ldots, x_N) - x_i$.

Obviously, $x_{N-n+1}, \ldots, x_N$ is a local coordinate system in the neighborhood U of x^0 in M. It is also clear that $(\xi, x) \in \rho^{-1}(U) \subset P$ if and only if ξ is a (non-zero) linear combination of the gradients ∇F_i, $\xi = \xi_1 \nabla F_1 + \ldots + \xi_{N-n} \nabla F_{N-n}$, $(\xi_1, \ldots, \xi_{N-n}) \ne (0, \ldots, 0)$. We see that $(\xi_1, \ldots, \xi_{N-n}, x_{N-n+1}, \ldots, x_N)$ is a local coordinate system in $\rho^{-1}(U) \subset P$ (confirming the fact that P is an N-dimensional manifold).

Now, let us find the critical points of the map π. We have:

$$\begin{aligned}
\pi(\xi_1, \ldots, \xi_{N-n}, x_{N-n+1}, \ldots, x_N) &= \sum_{i=1}^{N-n} \xi_i \nabla F_i \\
&= \sum_{i=1}^{N-n} \xi_i \left(\underbrace{0, \ldots, \overset{(i)}{-1}, \ldots, 0}_{N-n}, \frac{\partial f_i}{\partial x_{N-n+1}}, \ldots, \frac{\partial f_i}{\partial x_N} \right) \\
&= \left(-\xi_1, \ldots, -\xi_{N-n}, \sum_{i=1}^{N-n} \xi_i \frac{\partial f_i}{\partial x_{N-n+1}}, \ldots, \sum_{i=1}^{N-n} \xi_i \frac{\partial f_i}{\partial x_N} \right)
\end{aligned}$$

The Jacobian matrix of this map is $\begin{bmatrix} -I_{N-n} & D \\ 0 & \sum_i \xi_i H_i \end{bmatrix}$ where $D = \left\| \dfrac{\partial f_i}{\partial x_j} \right\|$ and $H_i = \left\| \dfrac{\partial^2 f_i}{\partial x_j \partial x_k} \right\|$. Thus we see that $(\xi, x) = (\xi_1, \ldots, \xi_{N-n}, x_{N-n+1}, \ldots, x_N) \in P$ is a critical point of the map π if and only if $\det(\sum_i \xi_i H_i(x_{N-n+1}, \ldots, x_N)) = 0$.

Next, we consider the function f_ξ. Let $x = (x_1, \ldots, x_N) \in U \subset M$ which means that $x_i = f_i(x_{N-n+1}, \ldots, x_N)$, $1 \le i \le N-n$, and let $\xi = (y_1, \ldots, y_N)$. Then $f_\xi(x) = \sum_{i=1}^{N-n} y_i f_i(x_{N-n+1}, \ldots, x_N) + \sum_{j=N-n+1}^{N} y_j x_j$ and for $N-n+1 \le j \le N$,

$$\frac{\partial f_\xi}{\partial x_j} = \sum_{i=1}^{N-n} y_i \frac{\partial f_i}{\partial x_j} + y_j \quad \text{or} \quad (y_1, \ldots, y_N) \cdot \left(\frac{\partial f_1}{\partial x_j}, \ldots, \frac{\partial f_{N-n}}{\partial x_j}, \underbrace{0, \ldots, \overset{(j)}{1}, \ldots, 0}_{n} \right)$$

Thus, x is a critical point of f_ξ, if and only if ξ is orthogonal to all the vectors $\eta_j = \left(\dfrac{\partial f_1}{\partial x_j}, \ldots, \dfrac{\partial f_{N-n}}{\partial x_j}, \underbrace{0, \ldots, \overset{(j)}{1}, \ldots, 0}_{n} \right)$, $N-n+1 \le j \le N$. But the vectors η_j are linearly independent, thus vectors orthogonal to all of them form a $(N-n)$-dimensional space. On the other side, the vectors $\nabla F_i = \left(\underbrace{0, \ldots, \overset{(i)}{-1}, \ldots, 0}_{N-n}, \dfrac{\partial f_i}{\partial x_{N-n+1}}, \ldots, \dfrac{\partial f_i}{\partial x_N} \right)$ are also linearly independent and obviously are all orthogonal to η_j. Thus, x is a critical point of f_ξ if and only if ξ is a linear combination of gradients of F_i, that is, if $\xi \perp T_x M$, that is, if $(\xi, x) \in P$. This settles Proposition 1.

Now, let, for some $x^0 \in U \subset M$, $\xi = \sum_{i=1}^{N-n} \xi i \nabla F_i(x^0) = \sum_{i=1}^{N-n} \xi_i \Bigg(\underbrace{0, \ldots, \overset{(i)}{-1}, \ldots, 0}_{N-n}, \dfrac{\partial f_i}{\partial x_{N-n+1}}(x^0), \ldots, \dfrac{\partial f_i}{\partial x_N}(x^0) \Bigg)$. Then $f_\xi(x_{N-n+1}, \ldots, x_N)) = -\sum_{i=1}^{N-n} \xi_i \Bigg[f_i(x_{N-n+1}, \ldots, x_N) + \sum_{j=N-n+1}^{N} x_j \dfrac{\partial f_i}{\partial x_j}(x^0) \Bigg]$. The second summand is linear in x_j and does not contribute to the matrix of second derivatives. Thus, $\left\| \dfrac{\partial^2 f_\xi}{\partial x_j \partial x_k} \right\| = \sum_{i=1}^{N-n} \xi_i \left\| \dfrac{\partial^2 f_i}{\partial x_j \partial x_k} \right\|$. Comparing this result with the previous result concerning the critical points of π, we observe that x is a degenerate critical point of f_ξ precisely if (ξ, x) is a degenerate critical point of f_ξ. This proves Proposition 2.

5.2.3. Generalizations.

5.2.3.1. Density.

Our proof of the existence theorem in 5.2.2 actually gives a stronger result: the set of Morse function (on a compact manifold) is dense in the space of all smooth functions with respect to the $\mathcal{C}^\infty$ topology. In other words, for every smooth function $h: M \to \mathbb{R}$, there exists a sequence $\{h_n\}$ of Morse functions which converges to h uniformly with all partial derivatives of all orders with respect to all charts (U, φ) with compact $\overline{U}$. To obtain this result, we modify the proof of Theorem 5.2.2 in two ways. First, for f_0 we take not 0, but h. Second every a used in construction of a Morse function h_n we take at the distance less than $\frac{1}{n}$ from the origin. Then the sequence $\{h_n\}$ will converge to h as stated.

5.2.3.2. Different critical values.

We can enhance the existence theorem with an additional property: every (compact) manifold M posesses a Morse function with all critical values different. To achieve that, we first consider some Morse function f on M. Then we take a critical point p and include it into two neighborhoods, U and V such that $\overline{V} \subset U$ and that f has no critical points in U different from p. In is convenient to assume that U is covered with some chart. For some positive δ, the gradient of the function f has length $> \delta$ at every point of $U - V$ (all this with respect to a chosen system of local coordinates in U). Take a smooth function $h: M \to \mathbb{R}$ which is equal to 1 within V and to 0 within $M - U$. Let C be the upper bound for the length of the gradient of h within U (with respect to the same local coordinate system). Then the function $g = f + \varepsilon h$ where $0 < \varepsilon < \frac{\delta}{C}$ has the same critical points as f and the same critical values with one exception: $g(p) = f(p) + \varepsilon$. In this way, we can vary critical values of a Morse function, and, in particular, can make them all different.

5.2.3.3. The boundary case.

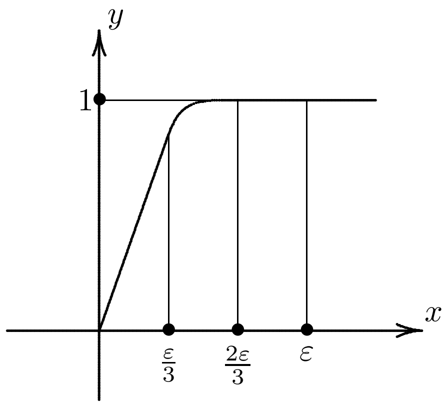
A smooth function $f: M \to \mathbb{R}$ is called a Morse function, if (1) $f\,|_{\partial M}$ is constant; (2) for a $p \in \operatorname{Int} M$, $f(p) \neq f(\partial M)$ (if M is connected, then $f\,|_{\operatorname{Int} M}$ has to be either $> f(\partial M)$, or $< f(\partial M)$; (3) f has no critical points on ∂M; (4) f has no degenerate critical points on M. Notice that condition (2) means that if $p \in \partial M$ and $\xi \in T_pM - T_p(\partial M)$ is directed inside the manifold (see 4.3.5), then $\xi(f) \neq 0$. Notice also that this is not the only possible definition of a Morse function on a manifold with boundary: for example, sometimes it is convenient to replace the condition (1) by the "opposite" condition that the restriction $f\,|_{\partial M}$ is a Morse function on ∂M (in which case we need to remove condition (2). We accept the definition of a Morse function on a manifold with boundary which better serves our purposes.

Theorem. *Every compact ∂-manifold has a Morse function.*

Remark. As in 5.2.2, compactness assumption is not necessary.

Proof of Theorem follows the lines of Proof in 5.2.2. First, using a collar $c: \partial M \times [0, \varepsilon) \to M$ we define a function $f_0: M \to \mathbb{R}$ by the formula

$$f_0(p) = \begin{cases} \mu(t), & \text{if } p = c(q,t) \text{ for some } q \in \partial M,\ 0 \le t < \varepsilon, \\ 1, & \text{if } p \in M - c(\partial M \times [0, \varepsilon)). \end{cases}$$



where μ is the function whose graph is shown on the right (the reader can write a formula for this function using the functions from 1.3.3.2.1).

The remaining construction is similar to that in 5.2.2. We fix charts $(\mathbb{R}^n, \varphi_i)$, $i = 1, \dots, N$ of $\operatorname{Int} M$ such that $\varphi_i(D^n)$ cover $M - c\left(\partial M - \frac{\varepsilon}{3}\right)$ and then define a sequence of functions $f_1, \dots, f_N$ precisely as it was done in 5.2.2 (with sufficiently small a). The function f_N is a Morse function on M.

Precisely as above, we can prove that any smooth function on M satisfying the conditions (1)–(3) of the definition of a Morse function on M can be $\mathcal{C}^\infty$ approximated by Morse functions and that a Morse function can be assumed having no equal critical values.

5.2.3.4. Further generalizations.

In conclusion, we will discuss some possibilities to request further properties from Morse functions; but we will not prove them, at least now (some indication how it can be done will be given later). For simplicity, we will restrict ourselves to manifolds without boundary (although everything can be extended to the boundary case).

First, it is well known that every smooth function on a compact manifold has at least one local maximum and at least one local minimum (one can take a *global* maximum and a *global* minimum). One can say more:

PROPOSITION 1. *A connected manifold possesses a Morse function with precisely one local minimum and precisely one local maximum.*

Second, it is certainly true that the critical point of a Morse function with the smallest value is a local (global) minimum. and the critical point with the greatest value is a local (global) maximum. Again, one can say more:

PROPOSITION 2. *An arbitrary manifold possesses a Morse function f with all critical values different, and such that if p, q are critical points and* $\operatorname{ind} p < \operatorname{ind} q$, *then* $f(p) < f(q)$.

Say, in dimension 2, we can request that all local minimum values are less than the values at saddle points and those are less than all values at local maxima.

5.3. A gradient-like vector field.

This is the last technical device we need to construct the Morse theory.

5.3.1. Definition.

Let M be a smooth manifold and let $f\colon M \to \mathbb{R}$ be a Morse function. A vector field X on M is called a gradient-like vector field (of f) if the function Xf is negative at all points of M not critical for f. (It is obvious that if $p \in M$ is a critical point of f, then $Xf(p) = 0$ for any vector field X.) So, in particular, if p is not a critical point of f, then $X_p \neq 0$. In the complement of the set of critical points, we can divide X by $-Xf$; then we get a vector field $\widetilde{X}$ (in this complement) such that $Xf = -1$.

From now on, we will assume the manifold M compact.

5.3.2. Existence.

Theorem. For a Morse function f on a compact manifold M, there is a gradient-like vector field.

Proof. Choose a finite set of charts (U_i, φ_i) of M such that $M = \bigcup_i \varphi_i(V_i)$ where $\overline{V}_i \subset U_i$. Let $h_i\colon M \to [0,1]$ be a smooth function such that $h_i(\varphi_i(V_i)) = 1,\ h_i(M - \varphi_i(U_i)) = 0$. Let then X_i be the image in $\varphi_i(U_i)$ the "minus gradient" vector field $-\nabla f \circ \varphi_i$ of the function $f \circ \varphi_i\colon U_i \to \mathbb{R}$. Then $X = \sum_i h_i x_i$ is a gradient-like vector field for f.

We can assume that in this construction, for every critical point of f, there is only one chart covering this point, and this is the chart from the Morse lemma.

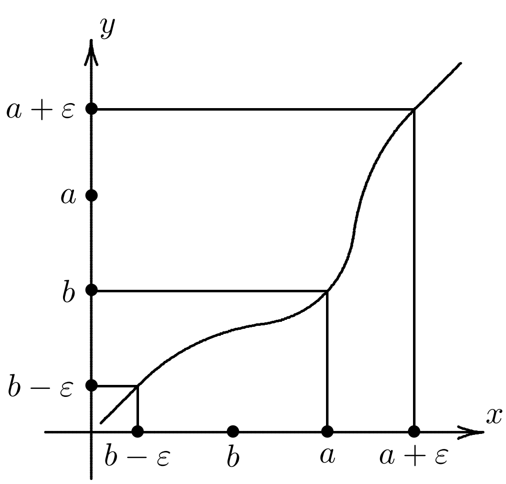
5.4. Main results.

5.4.1. Manifolds of smaller values.

Let $f\colon M \to \mathbb{R}$ be a Morse function on a compact manifold with all critical values different. Let $c_1 < c_2 < \ldots < c_N$ be all critical values of f. For a $c \in \mathbb{R}$, we consider the sets $M_c = \{p \in M \mid f(p) = c\}$ and $M_{\leq c} = \{p \in M \mid f(p) \leq c\}$. If c is not a critical value of f, then $M_{\leq c}$ is a manifold with the boundary M_c. Obviously, $M_c = M_{\leq c} = \emptyset$, if $c < c_1$, and $M_c = \emptyset$, $M_{\leq c} = M$, if $c > c_N$. Our goal is to study the dependence of $M_c$and $M_{\leq c}$ on c and to apply the results to describing the structure of M.

5.4.2. M_c and $M_{\leq c}$ are stable between critical value.

Theorem. *If, for some i, $c_i < b < a < c_{i+1}$, then $M_{\leq a}$ is diffeomorphic to $M_{\leq b}$ (and $M_a = \partial M_{\leq a}$ is diffeomorphic to $M_b = \partial M_{\leq b}$.*



Proof. Fix an $\varepsilon > 0$ such that $c_i < b - \varepsilon < a+\varepsilon < c_{i+1}$ and fix a gradient-like vector field X; we may assume that within $\{p \in M \mid b-\varepsilon \leq f(p) \leq a+\varepsilon\}$, $Xf = -1$. Let φ_t be the flow generated by the vector field X. Then, obviously, within the interval above there arise diffeomorphisms $\varphi_t\colon M_s \to M_{s-t}$, in particular, $M_a \to M_b$.

Moreover, if $\mu\colon \mathbb{R} \to \mathbb{R}$ is a function with the graph shown on the left, that is such that $\mu(t) = t$ for $t > a + \varepsilon$ and for $t < b - \varepsilon$, $\mu(a) = b$, $\mu(t) \leq t$ and $\mu'(t) > 0$ for all t, then the maps $\varphi_{t-\mu(t)}\colon M_t \to M_{\mu(t)}$, together with the identity $M_{\leq b-\varepsilon} \to M_{\leq b-\varepsilon}$, compose a diffeomorphism $M_{\leq a} \to M_{\leq b}$.

Notice also that the formula $p \mapsto \varphi_{f(t)-b}(p)$ defines a diffeomorphism $\overline{M_a - M_b} \to M_b \times [b,a]$ which takes the function f into the projection $M_b \times [b,a] \to [b,a]$.

The theorem above shows that among the manifolds $M_{\leq c}$ (as well as among $M_c = \partial M_{\leq c}$) there are finitely many different. To complete the description of these manifolds (among them, M) we need to understand what happens to these manifold when c passes through a critical value.

5.4.3. Key operations: attaching handles and Morse surgery.

5.4.3.1. Attaching handles.

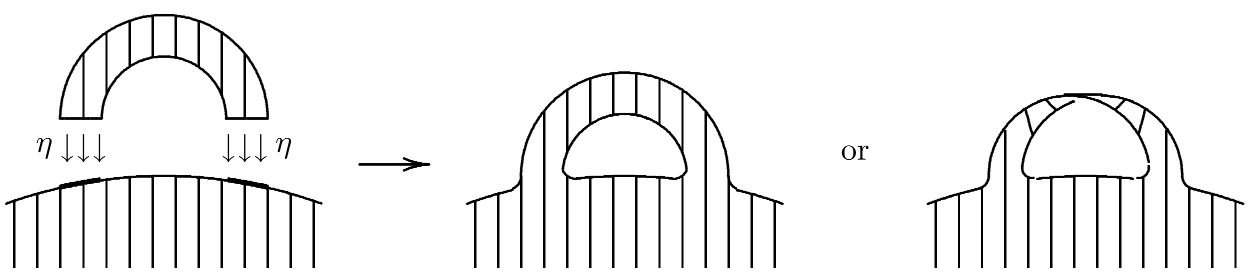
Let P be an n-dimensional manifold with boundary, and let $\eta\colon S^{i-1} \times D^{n-i} \to \partial P$ where $0 \leq i \leq n$ be an embedding. (Notice that if $i = 0$, then $S^{i-1} \times D^{n-i} = \emptyset$, so η means nothing; if $i = n$, then $S^{i-1} \times D^{n-i} = S^{n-1}$, so η has to be a diffeomorphism of S^{n-1} onto a component of ∂P.) On the other hand, $S^{i-1} \times D^{n-i}$ is a part of the boundary of the product $D^i \times D^{n-1}$ which is diffeomorphic to D^n (see 4.3.2 for a discussion of a product of two manifolds with boundaries). Now we use a construction of 4.4.4.2 to attach the product $D^i \times D^{n-1}$ to P. Denote the resulting manifold as $\widetilde{P}$. The transition from P to $\widetilde{P}$ is called *attaching a handle of index i by means of the embedding η.*

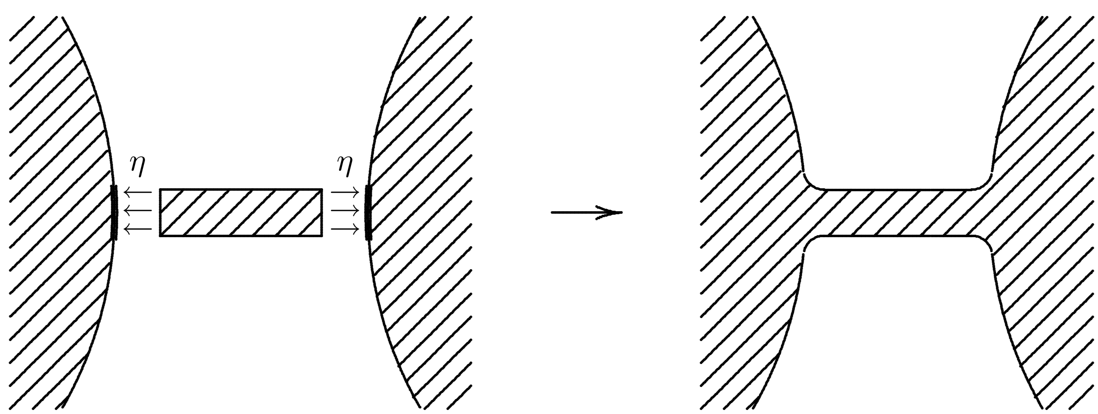
For a better understanding of this important operation, let us consider several examples.

Index 0. Since in this case η is an embedding of the empty set, the attaching a handle of index 0 means simply adding a new component: $\widetilde{P} = P \bigsqcup D^n$.

Index n. Since in this case η is a diffeomorphism of S^{n-1} onto a component of the boundary, attaching a handle of index n may be visualized as capping a hole.

Index 1. Since S^0 consists of two points, in this case η is an embedding of the union of two (disjoint) $(n-1)$-dimensional discs into ∂P, and attaching a handle of index 1 is attaching a cylinder $D^{n-1} \times [-1, 1]$ by the two bases. This cylinder may join two different components of P or to be attached to the same component. In the last case, attaching a handle may destroy the orientability of P. Several pictures (for $\dim P = 2$) are given below.





This geometric description of attaching handles on index 1 is the origin of the name of the operation.

Attaching handles of other indices also has a transparent geometric sense. For example, attaching a handle of index 2 means attaching a low cylinder along its side surface.

5.4.3.2. Morse surgery.

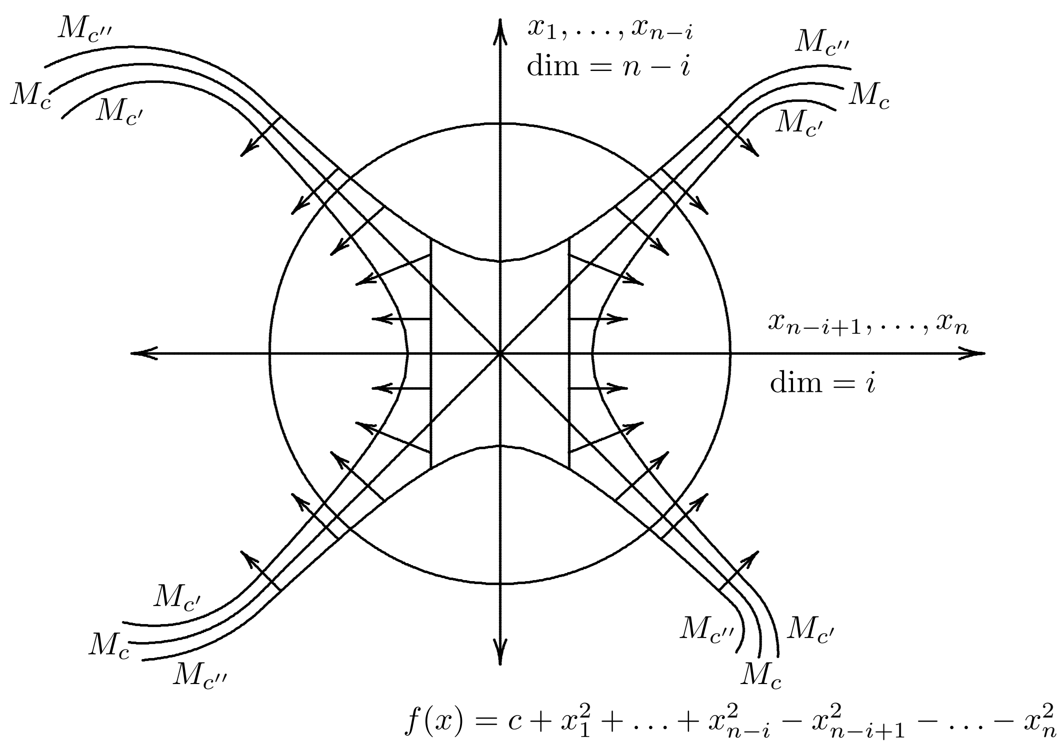
This is an operation which is applied to ∂P when P undergoes attaching a handle. We will not use it any seriously, but it is very important in differential topology, and by this reason it deserves a brief independent consideration.

Let Q be a manifold of dimension m, and let $\zeta\colon S^{i-1}\times D^{m-i+1}\to \operatorname{Int} Q$ $(0\le i\le m)$ be an embedding. Let us cut $\zeta(\operatorname{Int}(S^{i-1}\times D^{m-i+1}))$ from Q. We get a manifold Q' with the boundary $\partial Q'=\partial Q\coprod(S^{i-1}\times S^{m-i})$. Attach $D^i\times S^{m-i}$ to Q' along the diffeomorphism $\partial(D^i\times S^{m-i})=S^{i+1}\times S^{m-1}\subset\partial Q'$ of $\partial(D^i\times S^{m-i})$ onto a component of $\partial Q'$. We say that the resulting manifold $\widetilde{Q}$ is obtained from Q by a Morse surgery along the embedding ζ. For example, if $\widetilde{P}$ is obtaoned from P by an attaching a handle along an embedding $\eta\colon S^{i-1}\times D^{n-i}\to\partial P$, then $\partial\widetilde{P}$ is obtained from ∂P by a Morse surgery along the same embedding η.

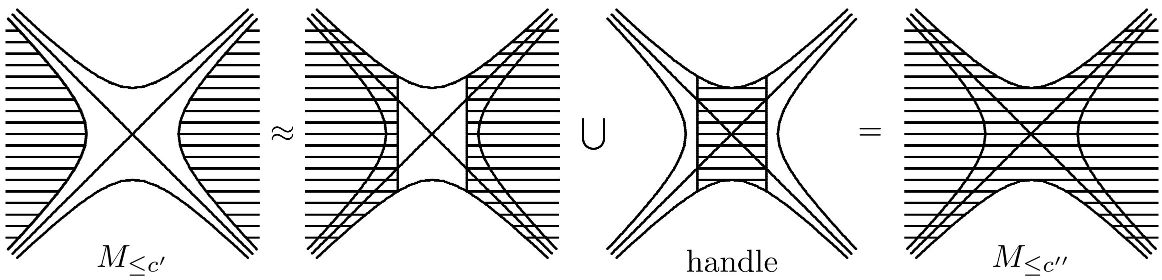
5.4.4. Passing a critical value.

Theorem. *Let c be a critical value of index i of the Morse function M, and let $[c',c'']$ be an interval such that $c'<c<c''$ and $[c',c']'$ contains no other critical values of f. Then $M_{c''}$ is obtained from $M_{c'}$ by attaching a handle of index i.*

Proof. It is all shown (at least when $0<i<n$) in the pictures below.



The first picture shows the structure of the sets $M_{\le c'}$, $M_{\le c}$, $M_{\le c''}$ in a neighborhood of the critical point with the critical value c. The three set closely follow each other everywhere outside this neighborhood. $M_{\le c''}$ consists of a part diffeomorphic to $M_{\le c'}$ (the diffeomorphism is shown by short arrows) and a handle $D^i\times D^{n-i}$ which is seen as a curvilinear rectangle in the middle of the picture. Further details are shown in the picture next page.



The cases $i = 0$ and $i = n$ are simple and do not deserve any drawing. When c passes through a critical value of index 0, that is, through a value at a local minimum, there arises first a point, and then a small disc centered at this point. This is attaching a handle of index 0. When c passes through a critical value of index n, that is through a value at a local maximum, a component of the boundary which is a small sphere collapses to a point. This is attaching a handle of index n.

5.4.5. The final statement.

The results of previous sections together give the following structure result.

Theorem. *Any smooth manifold M can be obtained from the empty set by successive attaching of finitely many handles. For a Morse function f on M, the number of handles of index i can be made equal to the number of critical points of f having index i.*

Corollary. *Let a compact manifold M (without boundary) possess a Morse function with precisely two critical points. Then M is homeomorphic to a sphere.*

Proof. Our manifold M is obtained from $\emptyset$ by attaching handles of indices 0 and $n = \dim M$. The first makes it D^n, the second consists in attaching D^n to D^n by some diffeomorphism $\varphi\colon S^{n-1} \to S^{n-1}$. The standard S^n is also made of two copies of D^n, but they attached to each other by id: $S^{n-1} \to S^{n-1}$. We compose a homeomorphism $M \to S^n$ from two homeomorphisms $D^n \to D^n$: the first is id, the second is defined by the formula $tx \mapsto t\varphi(x)$ for $x \in S^{n-1}$, $0 \le t \le 1$.

Remarks. (1) *Homeomorphic* in the last statement does not necessarily means *diffeomorphic* (J. Milnor, 1956.) (2) It is true, actually that if M posseses a smooth function with precisely two critical points *which are not assumed non-degenerate* then it also homeomorphic to a sphere (G. Reeb, 1949).

5.4.6. The boundary case.

This case is not different from the boundary free case. It was remarked in 5.2.3.3 that a ∂-manifold possesses a Morse function which is 0 on the boundary and positive inside. For our current purposes we prefer to attach a minus to this function. Then again we consider $M_{\le c}$ with c growing and observe the same process of attaching handles which stops when we reach $M_{\le 0} = M$.

5.4.7. Homotopy type.

Up to a homotopy equivalence, attaching a handle of index i to a ∂-manifold P is the same as attaching a disc D^i by means of some continuous map $\eta\colon S^{i-1} \to \partial P \subset P$. Moreover, up to a homotopy equivalence the result of attaching depends only of homotopy types of P and η. If we apply this observation to the whole process of building a manifold M (with or without boundary) with a Morse function f from the empty set by successive attaching handles, we will get a CW complex homotopy equivalent to M with the number of i-dimensional cells equal to the number of critical points of index i of the function f.

REMARK. It is known (S. Smale, 1960) that if a compact smooth manifold M *of dimension* ≥ 5 is homotopy equivalent to some CW complex X then it possesses a Morse function with the number of critical points of index i equal to the number of i-dimensional cells in X. With Corollary in 5.4.4.5, this implies that if a manifold (compact, without boundary) is homotopy equivalent to a sphere, then it is homeomorphic to this sphere ("generalized Poincaré conjecture").

5.5. Application of the Morse theory: classification of compact manifolds of dimension 2 (aka compact surfaces).

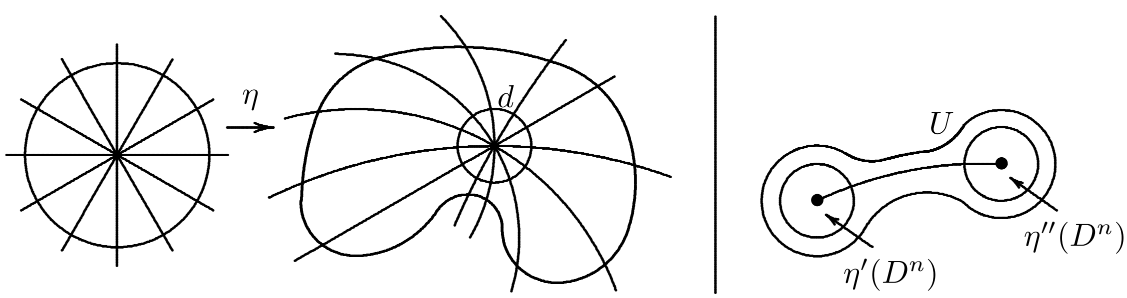
Terminology: a 2-dimensional compact manifold, possibly with boundary, is called a *compact surface*.

5.5.1. Basic operations.

5.5.1.1. Drilling a hole.

Let S be a connected compact surface, and let $\eta\colon D^2 \to \operatorname{Int} S$ be a smooth embedding (for clarity, we can assume that η can be extended to an embedding of an open disc of some radius > 1 centered at the origin. We consider the manifold $S' = S - \eta(\operatorname{Int} D^n)$. (Obviously, $\partial S' = \partial S \bigsqcup S^1$.) The transition from S to S' is called *drilling a hole*.

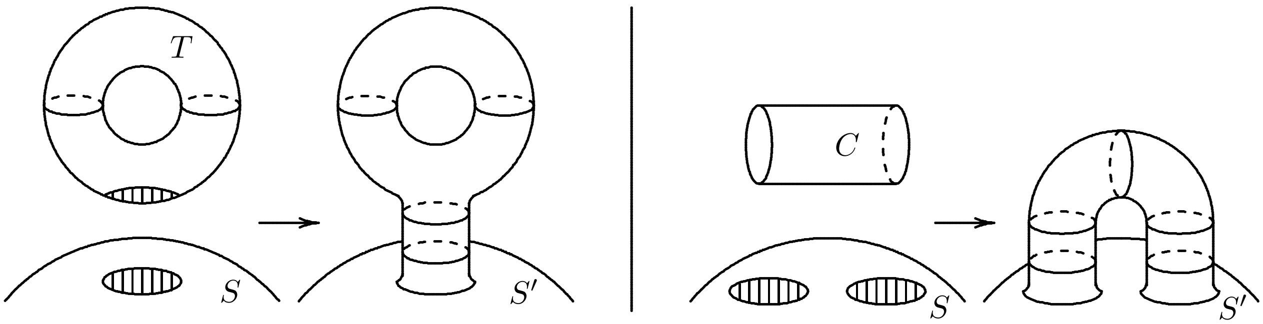
It is important to notice that, up to a diffeomorphism, drilling a hole is a well-defined operation, that is, if S', S'' are obtained from the same S by drilling holes by means of embeddings $\eta', \eta''\colon D^n \to S$, then S' and S'' are diffeomorphic. The proof is shown schematically on the pictures below. First, a self-diffeomorphism of S along the η-images of radial rays (identity on the rest of the surface and in the neighborhood of $\eta(0)$) takes $\eta(D^n)$ into a small disc centered at $\eta(0)$ in a local coordinate system around $\eta(0)$ (see the left picture); this shows that two embeddings, $\eta', \eta''\colon D^n \to S$ such that $\eta'(0) = \eta''(0)$ may be related by a diffeomorphism of S. Second, we can pull $\eta(0)$ along a path using a diffeomorphism within U, as shown in the right picture.



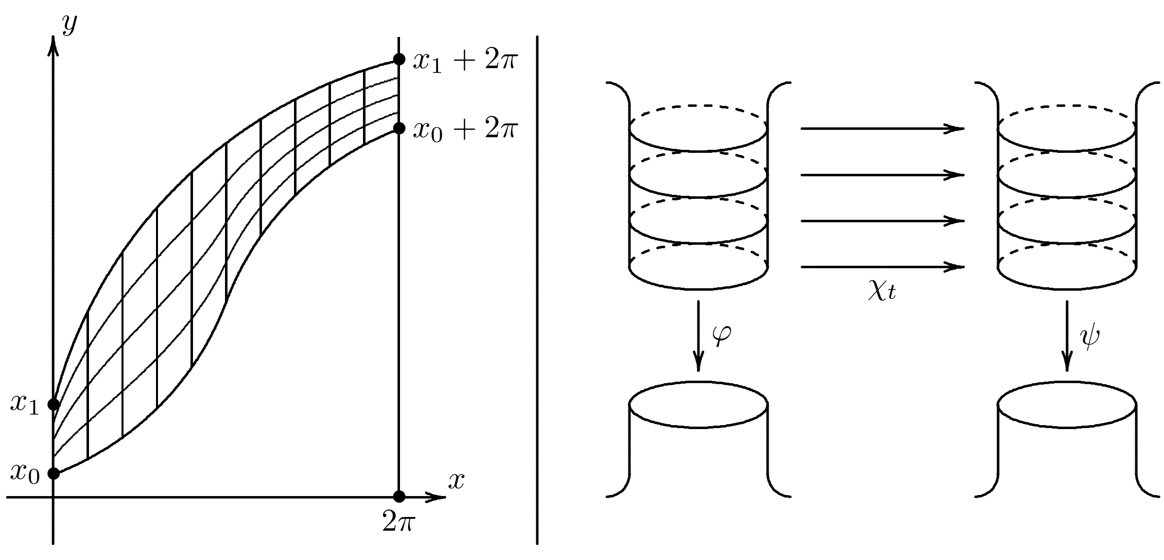
5.5.1.2. Attaching a handle.

To begin with, this is *not* the same operation which was considered above, in 5.4.3.1. To avoid a confusion, I could rename one of these operation (for instance, the operation of 5.4.3.1 could be called "attaching a handlebody.") But this confusion exists everywhere in literature, and it is not for me to fight this well-established tradition.

The operation we consider now is the following. Take a connected surface S surface and drill a hole in it. Then take a torus T and also drill a hole. Then attach T to S along the edges of the holes (see a picture next page, left). Another construction of the same is shown on the right picture: we attach a cylinder C to a surface S with two holes. The resulting surface is denoted as S'.



Again, we need to check that this operation is well defined. Regarding the construction on the left picture, we need only to check that if we attach the torus to S using two different diffeomorphisms, $\varphi, \psi\colon S^1 \to S^1$, we will get diffeomorphic surfaces S'. First we notice that we can assume that φ both preserve orientations (otherwise, we can, before the attachment, reflect the torus in any of two vertical planes of symmetry). Then we observe that φ and ψ must be *isotopic*, that is, there is a smooth family of diffeomorphisms $\{\varphi_t\colon S^1 \to S^1,\ 0 \le t \le 1$ of diffeomorphisms such that $\varphi_0 = \varphi, \varphi_1 = \psi$. To prove this, we present φ and ψ as smooth functions $\widetilde{\varphi}\colon [0, 2\pi] \to [x_0, x_0 + 2\pi]$, $\widetilde{\psi}\colon [0, 2\pi] \to [x_1, x_1 + 2\pi]$ with positive $\widetilde{\varphi}', \widetilde{\psi}'$ and put $\widetilde{\varphi}_t(x) = (1-t)\widetilde{\varphi}(x) + t\widetilde{\psi}(x)$ (see the picture below on the left). After this, we put $\chi_t = \psi \circ \varphi_t$ and, before attaching the cylinder pulled from the torus to the cylinder pulled from the surface S, we apply to the upper cylinder a twisting map defined by means the family χ_t, as shown on the right picture below.



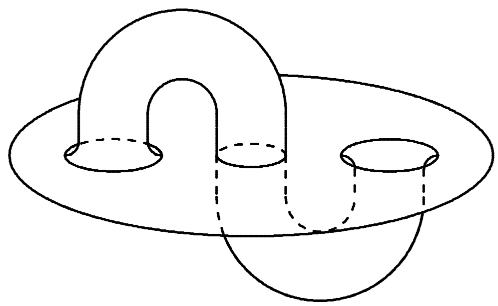
As to attaching a handle to two hole (as shown above on the right), there is a possibility to attach the cylinder to two oppositely oriented circles. We will return to this version of attaching a handle later in 5.5.1.4.

5.5.1.3 Attaching a Möbius band.

The boundary of a Möbius band is (diffeomorphic to) a circle. Hence, we can transform a surface by first drilling a hole and then attaching s Möbius band to the edge of this hole. There is another description of the same operation: we attach an arc to every pair of opposite points of the edge of the hole (again, it is not possible to do that without self-intersections in space). The arguments above show that the operation of attaching a Möbius band is well defined.

5.5.1.4. Attaching an inverted handle.

This is an attaching a cylinder to edges of two holes shown in a picture below. This is also a well-defined operation.



Sphere with attached g handles, m Möbius bands, and drilled h holes is denoted as $S(g, m, h)$.

5.5.1.5. Examples.

Sphere with one hole, $S(0, 0, 1)$, is a disk.

Sphere with one handle, $S(1, 0, 0)$, is a torus. Sphere with g handles, $S(g, 0, 0)$, is called (especially in algebraic geometry) a surface of genus g.

Sphere with one hole and one Möbius band, $S(0, 1, 1)$, is a Möbius band.

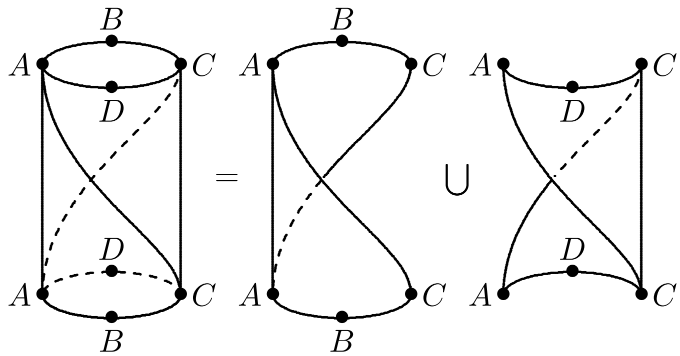
Sphere with one Möbius band, $S(0, 1, 0)$, is a projective plane. The surface $S(g, 1, 0)$ is usually referred to as a projective plane with g handles.

The Klein bottle has two descriptions. First, there is a popular picturebook description which presents the Klein bottle as a sphere with an inverted handle; second, it is obtained by attaching two copies of a Möbius band along their boundaries; the latter may be also identified as $S(0, 2, 0)$. The picture on the next page shows a way how to cut a Klein bottle into two Möbius bands. (The Klein bottle is shown there as a circular cylinder with the two bases attached to each other after reflection in a diameter.)

The equivalence of the two descriptions of the Klein bottle has an immediate consequence of the following proposition which will be also useful later.

Proposition. *If $m \geq 3$, then, for any g, h, there is a diffeomorphism $S(g, m, h) \approx S(g + 1, m - 2, h)$.*

Proof. The surface $S(g, m, h)$ may be regarded as a Klein bottle $S(0, 2, 0)$ with additional g handles, $m - 2$ Möbius bands, and h holes. Because of the equivalence of the two construction above, this is a sphere with inverted handle plus g handles, $m - 2$ Möbius bands, and h holes. Since $m - 2 > 0$, there is at least one Möbius band. We pull one of the holes of attachment of the inverted handle to this Möbius band, then drag it around this Möbius band, and then pull it back ti its in initial position. As a result, the inverted handle becomes an usual handle, and the whole surface becomes a sphere with a handle plus g more handles, $m-2$ Möbius bands, and h holes, that is, the surface becomes $S(g+1, m-2, h)$ as was stated.



Klein bottle Two Möbius bands

5.5.2. Main result.

5.5.2.1. The statement and a scheme of proof.

THEOREM. *Any connected compact surfaces is diffeomorphic to precisely one from the following surfaces:*

– *sphere with handles and holes;*
– *projective plane with handles and holes;*
– *Klein bottle with handles and holes.*

We will call these surfaces *standard*.

With the exception of the clause "precisely one" this theorem is proven below.

Scheme of a proof. First we prove the following

PROPOSITION. *Any connected compact surface is diffeomorphic to some surface of the form $S(g, m, h)$.*

This is proved in 5.5.2.2; the proof is based on results of the Morse theory (5.4.4 and 5.4.5). Combined with Proposition in 5.5.1.5, this shows that any connected compact surface is diffeomorphic to one of the surfaces listed in Theorem.

It remains to show that the surfaces of this list are not diffeomorphic to each other. We will discuss it in 5.5.2.3, but will not give a full proof.

5.5.2.2. Proof of diffeomorphism with a standard surface.

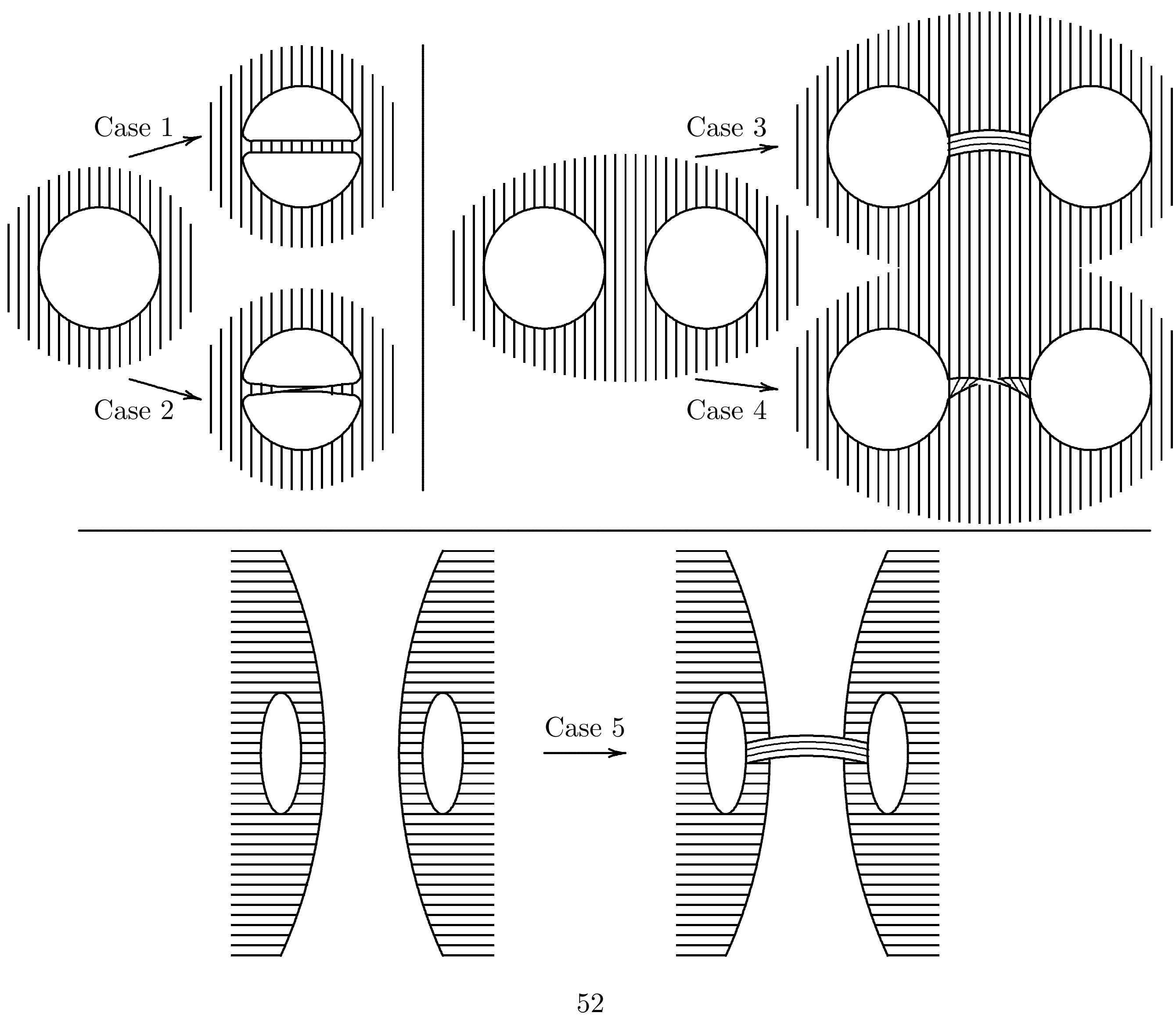
According to results of 5.4.4 and 5.4.5, an arbitrary surface can be obtained from an empty set (or from a disk) by a sequence of attaching handles. We assume that before an attaching a handle, every component of our surface was of the form $S(g, m, h)$; we want to prove that the same will be true after attaching a handle. A handle can have indices 0, 1, or 2. Let us begin with two trivial cases.

Attaching a handle of index 0 creates a new component, which is a disk, that is $S(0, 0, 1)$, and makes no other changes. Thus, we have no problems in this case.

Attaching a handle of index 2 consists in capping a hole. This affects only one component, some $S(g, m, h)$ with $h > 0$, and makes it $S(g, m, h-1)$. Again, everything is OK in this case.

Attaching a handle of index 1 is attaching a rectangle ("a piece of skotch") by two opposite sides to two intervals on the boundary of our surface. Here we have to consider 5 different cases. Cases 1 and 2: we attach the rectangle to the edge of one hole in such a way that the union of the rectangle and a thin neighborhood of the edge of the hole is orientable (Case 1) or not (Case 2). Cases 3 and 4: we attach the rectangle to two separate edges of two holes within one component of the surface; again, the union of a rectangle, thin neighborhoods of the edges and a strip joining the edges within the surface may be orientable (Case 3) or not (Case 4). (Actually, if the component to which we attach the piece of skotch was not orientable, then this two cases are the same; we do not need to pay any attention to this.) Case 5: we attach the rectangle to edges of holes on different component of the surface.

These 5 cases are shown on a picture below.



In Case 1 we just get an additional hole: $S(g,m,h)$ becomes $S(g,m,h+1)$. In Cases 2–5, we use 1 or 2 holes, but get a piece of a surface with a connected boundary. Let us fill this appearing hole with a cap, then look what we get, and then remove the cap, that is, drill a hole. In cases 2–4, this composition operation (attaching a piece of scotch and then a cap) is equivalent to attaching a Möbius band (we attach an arc to every pair of opposite points on the edge of the hole), attaching a handle, and attaching a perverted handle (which is the same as attaching a Möbius band to each of the holes, see 5.5.1.5). Attaching a piece of scotch leaves the number of holes unchanged in Case 2 and reduces it by 1 in Cases 3 and 4. Hence $S(g,m,h)$ in Cases 2–4, becomes, respectively, $S(g,m+1,h)$, $S(g+1,m,h-1)$, and $S(g,m+2,h-1)$. Finally, in Case 5 we are dealing with two surfaces, $S(g,m,h)$ and $S(g',m',h')$, the handles and Möbius bands on these two surfaces stay intact, the two holes participating in the operation become one.

Here is the summary of our results:

Case 1: $S(g,m,h) \Longrightarrow S(g,m,h+1)$.
Case 2: $S(g,m,h) \Longrightarrow S(g,m+1,h)$.
Case 3: $S(g,m,h) \Longrightarrow S(g+1,m,h-1)$.
Case 4: $S(g,m,h) \Longrightarrow S(g,m+2,h-1)$.
Case 5: $S(g,m,h) \sqcup S(g',m',h') \Longrightarrow S(g+g',m+m',h+h'-1)$.

These results complete the proof of Proposition in 5.5.2.1.

5.5.2.3. All standard surfaces are not diffeomorphic to each other.

This result is not especially difficult, but any proof requires some material not covered in these notes.

There are two things known to us. First, the number of holes is the number of components of the boundary. If two surfaces are diffeomorphic, then their boundaries are diffeomophic, hence, they have the same number of components. Thus, if two standard surfaces are diffeomorphic, then the number of holes is the same. We can fill the holes, and the surfaces will remain diffeomorphic. Thus, we need to prove that standard surfaces without boundaries, that is, spheres with handles, projective planes with handles, and Klein bottles with handles, are all different. Another observation is that some of the standard surfaces are orientable, and some are not, and an orientable manifold cannot be diffeomorphic to a non-orientable manifold. Thus, we need to distinguish separately between spheres with handles, and separately between projective planes and Klein bottles with handles.

To do this, we must use some tools from topology. For example, it is not hard to prove that all standard surfaces without boundary have non-isomorphic fundamental groups. It is easier to apply a simpler invariant: the Euler characteristic. The latter may be defined by the means of the Morse theory. Namely, let f be a Morse function on a compact manifold M, with or without boundary, and let $c_i(f)$ be the number of critical points of f of index i.

Proposition. *The alternating sum $\sum_i(-1)^i c_i(f)$ does not depend on f, that is, is determined by M.*

Idea of Proof. Any two Morse functions on M may be connected by a "generic deformation" f_t. In this deformation, only finitely many of functions f_t are not Morse, and when t passes a value for which f_t is not Morse, only two events may happen: two critical points of neighboring indices may appear or disappear. (For example, in a generic family of functions of one variable, a local maximum and a local minimum may simultaneously appear or disappear.) Obviously, neither of these events affects the alternating sum in the statement.

The alternating sum in Proposition is called the *Euler characteristic* of the manifold. Obviously, diffeomorphic manifolds have equal Euler characteristics.

A direct computation shows that the Euler characteristic of $S(g, m, h)$ is equal to $2-2g-m-h$. In particular, the Euler characteristic of a sphere with g handles is $2-2g$, so all spheres with handles are not diffeomorphic to each other. The Euler characteristic of a projective plane with g handles is $1-2g$, the Euler characteristic of a Klein bottle with g handles is $-2g$; thus all these non-orientable surfaces are not diffeomorphic to each other. (But Euler characteristics do not distinguish between spheres with handles and Klein bottles with handles!)

Final Remark. Both fundamental groups and Euler characteristics are *homotopy* invariants. Thus, standard surfaces without boundaries are not only not diffeomorphic, but also not homotopy equivalent to each other. For surfaces with boundary, a similar thing is not true.

5.6. An application to three-dimensional manifolds: Heegard splitting.

The following result is one of the most important tools of topology of three-dimensional manifold. We will deduce it from the Morse theory using Proposition 2 of 5.2.3.4 which was given without a proof. We use the following notations: $S_g = S(g, 0, 0)$ is a two-dimensional sphere with g handles, and B_g is a "handlebody, a solid bounded in $\mathbb{R}^3$ by S_g embedded in $\mathbb{R}^3$ in the standard way (a garland of solid tori connected by solid cylinders with a common axis). Thus, $\partial B_g = S_g$.

Theorem. *Let M be a connected orientable compact three-dimensional manifold without boundary. There exist a g and a diffeomophism $\varphi\colon S_g \to S_g$ such that M is diffeomorphic to the manifold $B_g \cup_\varphi B_g$ obtained by attaching B_g to B_g by means of φ.*

The splitting $M = B_g \cup_\varphi B_g$ is called a *Heegard splitting*.

Proof of Theoem. Let f be a Morse function on M satisfying the condition of Proposition 2 of 5.2.3.4: if p, q are critical points of f and $\operatorname{ind}(p) < \operatorname{ind}(q)$, then $f(p) < f(q)$. Let c be a real number which is greater than any critical value of f of indices 0 and 1, but less than any critical value of f of indices 1 and 2. Then $M = M_{\le c} \cup M_{\ge c}$ and $M_{\le c} \cap M_{\ge c} = \partial M_{\le c} = \partial M_{\ge c} = M_c$. The manifold M_c must be connected (since attaching handles of indices > 1 would not change the number of components) and orientable (as a codimension 0 submanifold of an orientable manifold). It is obtained by attaching solid cylinders to balls, thus it is a handlebody B_g with some g. Its boundary M_c is S_g. Finally, $M_{\ge c}$ is also a handlebody, since it is $M_{\le -c}$ for the function $-f$. It is B_g with the same g, since its boundary is S_g.

Remark. This theorem has an appearance of a full classification of compact orientable thee-dimensional manifold. However, it leaves unanswered two questions: a classification of (isotopy classes of) self-diffeomorphisms of S_g, and also the problem of possible diffeomorphisms $B_g \cup_\varphi B_g \approx B_{g'} \cup_{\varphi'} B_{g'}$. Luckily, the first problem has been settled by

Max Dehn in the thirties: there exists an explicit description of the isotopy classes of self-diffeomorphisms of S_g. The second problem, however, is algorithmically unsolvable.

6. Differential forms and the Stokes theorem.

6.1. A linear algebra introduction.

6.1.1. The space of exterior forms.

Let V be an n-dimensional (real) vector space. An exterior form α of degree q on V is defined as a real-valued function $[(v_1, \ldots, v_q) \in \underbrace{V \times \ldots \times V}_{q}] \mapsto \alpha(v_1, \ldots, v_q)$ which is q-linear,

$$\alpha(v_1, \ldots, a'v_i' + a''v_i'', \ldots, v_n) = a'\alpha(v_1, \ldots, v_i', \ldots, v_n) + a''\alpha(v_1, \ldots, v_i'', \ldots, v_n),$$

for $1 \le i \le q$, $v_1, \ldots, v_i', v_i'', \ldots, v_n \in V$, $a', a'' \in \mathbb{R}$, and skew-symmetric,

$$\alpha(v_{\tau(1)}, \ldots, v_{\tau(q)}) = \operatorname{sgn}(\tau)\alpha(v_1, \ldots, v_n)$$

for $\tau \in S_q$. The set $\Lambda^q V^*$ of all exterior forms of degree q in V has a natural structure of a vector space, and its dimension is $\binom{n}{q}$. Indeed, if $\{e_1, \ldots, e_n\}$ is a basis in V, then an exterior form $\alpha \in \Lambda^q V^*$ is fully determined by $\binom{n}{q}$ real numbers

$$\alpha_{i_1 \ldots i_q} = \alpha\left(e_{i_1}, \ldots, e_{i_q}\right), \ 1 \le i_1 < \ldots < i_q \le n,$$

and these numbers can be chosen arbitrarily. For example,

$$\begin{aligned}\alpha(e_1 + e_2, e_2 + e_3, e_3 + e_4) &= \alpha(e_1, e_2, e_3) + \alpha(e_1, e_2, e_4) + \alpha(e_1, e_3, e_3) + \alpha(e_1, e_3, e_4)\\ &\quad + \alpha(e_2, e_2, e_3) + \alpha(e_2, e_2, e_4) + \alpha(e_2, e_3, e_3) + \alpha(e_2, e_3, e_4)\\ &= \alpha_{123} + \alpha_{124} + \alpha_{134} + \alpha_{234}.\end{aligned}$$

General formula:

$$\alpha\left(\sum_{i=1}^{n} a_{i1}e_i, \ldots, \sum_{i=1}^{n} a_{iq}e_i\right) = \sum_{1 \le i_1 < \ldots < i_q \le n} \det \begin{bmatrix} a_{i_1 1} & \ldots & a_{i_1 q} \\ \ldots & \ldots & \ldots \\ a_{i_q 1} & \ldots & a_{i_q q} \end{bmatrix} \cdot \alpha_{i_1 \ldots i_q}.$$

Proof: exercise.

Thus, $\dim \Lambda^0 V^* = 1$ ($\Lambda^0 V^*$ is just $\mathbb{R}$), $\dim \Lambda^1 V^* = n$ ($\Lambda^1 V^*$ is just V^*), $\dim \Lambda^n V^* = 1$, and $\Lambda^q V^* = 0$ for $q > n$.

6.1.2. Reaction to linear maps.

If $f\colon V \to W$ is a linear map, then the formula

$$[f^*(\alpha)](v_1, \ldots, v_q) = \alpha(f(v_1), \ldots, f(v_q)), \text{ where } \alpha \in \Lambda^q W^*,\ v_1, \ldots, v_q \in V,$$

defines a linear map $f^*\Lambda^q W^* \to \Lambda^q V^*$. This operation possesses standard "functorial" properties: $\mathrm{id}^* = \mathrm{id}$, $(f \circ g)^* = g^* \circ f^*$.

6.1.3. Exterior products.

Let $\alpha \in \Lambda^q V^*, \beta \in \Lambda^r V^*$. The form $\alpha \wedge \beta \in \Lambda^{q+r} V^*$ is defined by the formula

$$(\alpha \wedge \beta)(v_1, \dots, v_{q+r}) = \sum_{1 \le i_1 < \dots < i_q \le q+r} (-1)^{i_1 + \dots + i_q - \frac{q(q+1)}{2}} \alpha(v_{i_1}, \dots, v_{i_q}) \beta(v_{j_1}, \dots, v_{j_r})$$

where $1 \le j_1 < \dots < j_r \le q + r$, $\{j_1, \dots, j_r\} = \{1, \dots, q + r\} - \{i_1, \dots, i_q\}$. The sign $(-1)^{i_1 + \dots + i_q - \frac{q(q+1)}{2}}$ is needed to make $\alpha \wedge \beta$ skew-symmetric.

Proposition 1. *For all* $\alpha \in \Lambda^q V^*, \beta \in \Lambda^r V^*, \gamma \in \Lambda^s V^*$,

(a) $\alpha \wedge (\beta \wedge \gamma) = (\alpha \wedge \beta) \wedge \gamma$.

(b) $\beta \wedge \alpha = (-1)^{qr} \alpha \wedge \beta$.

Proof: exercise.

Proposition 2. *Let* $\{\varepsilon_1, \dots, \varepsilon_n\}$ *be a basis of* $V^* = \Lambda^1 V^*$ *dual to* $\{e_1, \dots, e_n\}$ *(that is,* $\varepsilon_i(e_j) = \delta_{ij}$*). Then, for all* $\alpha \in \Lambda^q V^*$,

$$\alpha = \sum_{1 \le i_1 < \dots < i_q \le n} \alpha_{i_1 \dots i_q} \varepsilon_{i_1} \wedge \dots \wedge \varepsilon_{i_q}$$

where (as before) $\alpha_{i_1 \dots i_q} = \alpha(e_{i_1}, \dots, e_{i_q})$.

Proof: exercise.

6.2. Differential forms on a manifold.

6.2.1. Definition.

Let M be an n-dimensional manifold, and let $q \ge 0$. A differential form α of degree q on M is defined as a family of exterior forms $\alpha_p \in \Lambda^q (T_p M)^*$ (a more common notation for $(T_p M)^*$ is $T_p^* M$) which is $\mathcal{C}^\infty$ with respect to p. The latter can be defined in three equivalent ways (the equivalence will be obvious(.

(1). *Coordinate description.* Let $(x_1, \dots, x_n)$ be a local coordinate system defined in $U \subset M$. Then, for every $p \in U$, thee arises a basis $\left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right)$ of $T_p M$, and α is fully determined, within U, by the functions

$$\alpha_{i_1 \dots i_q}(p) = \alpha_p \left(\frac{\partial}{\partial x_{i_1}}, \dots, \frac{\partial}{\partial x_{i_q}} \right),\ 1 \le i_1 < \dots < i_q \le n.$$

We require that all these functions are $\mathcal{C}^\infty$. Certainly, one needs to check that this property does not depend on the choice of local coordinate, but this is done by a routine application of the transition formulas.

(2) *A coordinate-free version of the same.* We can require that for any vector fields $X_1, \dots, X_q \in \mathrm{Vect}(M)$, the function

$$\alpha(X_1, \dots, X_q)\colon M \to \mathbb{R},\ (\alpha(X_1, \dots, X_q))(p) = \alpha_p(X_{1,p}, \dots, X_{q,p})$$

belongs to $\mathcal{C}^\infty(M)$.

(3) *Using* TM. Let

$$(TM)^q_\Delta = \bigcup_{p\in M} \underbrace{T_pM \times \ldots \times T_pM}_{q} \subset \bigcup_{p_1\ldots p_q \in M} (T_{p_1}M \times \ldots \times T_{p_q}M) = \underbrace{TM \times \ldots \times TM}_{q}.$$

Obviously, $(TM)^q_\Delta$ is a submanifold of $TM \times \ldots \times TM$. The formula $\alpha(\xi_1, \ldots, \xi_q) = \alpha_p(\xi_1, \ldots, \xi_q)$ for $\xi_1, \ldots, \xi_q \in T_pM$ defines a function on $(TM)^q_\Delta$, and we require that this function is $\mathcal{C}^\infty$.

Notice that Condition (2) gives rise to a definition of a differential form, seemingly independent of Linear algebra. A differential form of degree q on M is a function

$$\alpha: \underbrace{\mathrm{Vect}(M) \times \ldots \times \mathrm{Vect}(M)}_{q} \to \mathcal{C}^\infty(M)$$

which is skew-symmetric and multilinear *over functions*, that is

$$\alpha(X_1, \ldots, f'X'_i + f''X''_i, \ldots, X_n) = f'\alpha(X_1, \ldots, X'_i, \ldots, X_n) + f''\alpha(X_1, \ldots, X''_i, \ldots, X_n),$$

for $1 \le i \le q$, $X_1, \ldots, X'_i, X''_i, \ldots, X_n \in \mathrm{Vect}(M)$, $f', f'' \in \mathcal{C}^\infty(M)$. We will use below all three versions of definition.

The space of differential forms of degree q on M is denoted as $\Omega^q M$. It is obvious that $\Omega^0 M = \mathcal{C}^\infty(M)$.

The operation of a wedge-product exists for differential forms: if $\alpha \in \Omega^q M, \beta \in \Omega^r M$, then there arises a differential form $\alpha \wedge \beta \in \Omega^{q+r} M$, and Proposition 1 of 6.1.3 holds without any changes. It is clear also that a smooth map $f: M \to N$ between two manifolds induces linear maps $f^*: \Omega^q N \to \Omega^q M$, and $\mathrm{id}^* = \mathrm{id}$, $(f \circ g)^* = g^* \circ f^*$, $f^*(\alpha \wedge \beta) = f^*\alpha \wedge f^*\beta$.

6.2.2. Differentials.

6.2.2.1. Differentials of functions.

Recall that we know an example of a differential form, and this is the differential of a function. If $f \in \mathcal{C}^\infty(M) = \Omega^0 M$ then any of the formulas $d_pf(\xi) = \xi(f)$, $df(X) = Xf$ where $p \in M$, $\xi \in T_pM$, $X \in \mathrm{Vect}\, M$ defines a differential form $df \in \Omega^1(f)$. Thus, we have a homomorphism $d: \Omega^0 M \to \Omega^1 M$. Our goal is to generalize it to a homomorphism $d: \Omega^q M \to \Omega^{q+1} M$.

First let us notice that if $(x_1, \ldots, x_n)$ is a local coordinate system in a $U \subset M$, then $dx_i\left(\dfrac{\partial}{\partial x_j}\right) = \delta_{ij}$; in particular, $\{d_px_1, \ldots, d_px_n\}$ is a basis of T^*_pM dual to the standard basis $\left\{\dfrac{\partial}{\partial x_1}, \ldots \dfrac{\partial}{\partial x_n}\right\}$ of T_pM. Applying Proposition 2 of 6.1.3 to every T_pM, $p \in U$, we get the following result.

Proposition. *In* U*, an arbitrary differential form* $\alpha \in \Omega^q M$ *has the form*

$$\alpha = \sum_{1 \le i_1, < \ldots < i_q \le n} \alpha_{i_1 \ldots i_q}(x_1, \ldots, x_n)\, dx_{i_1} \wedge \ldots \wedge dx_{i_q}$$

where $\alpha_{i_1\dots i_q}(x_1,\dots,x_n)$ *is the function* $\alpha_{i_1\dots i_q}(p) = \alpha_p\left(\frac{\partial}{\partial x_{i_1}},\dots,\frac{\partial}{\partial x_{i_q}}\right)$ *expressed in the local coordinates* $(x_1,\dots,x_n)$*.*

6.2.2.2. An axiomatic description of the differential.

THEOREM. *For any manifold* M*, there exists a unique sequence of linear maps* $d\colon \Omega^q M \to \Omega^{q+1}M,\ q = 0,1,2,\dots,$ *with the following properties:*

(1) $d^2 = d\circ d\colon \Omega^q M \to \Omega^{q+2}M$ *is always zero;*

(2) *for* $\alpha \in \Omega^q M,\ \beta\in\Omega^r M,$

$$d(\alpha\wedge\beta) = (d\alpha)\wedge\beta + (-1)^q\alpha\wedge d\beta;$$

(3) *for a function* $f\in\Omega^0 M,\ df\in\Omega^1 M$ *is the differential of* f *as defined above.*

Proof is contained in Sections 6.2.2.3 and 6.2.2.4 below.

The differential d arising from this Theorem is called sometimes the *exterior differential* and sometimes *de Rham differential*.

6.2.2.3. Proof of uniqueness.

This is the easiest part of a proof. Since in the domin of local coordinates an arbitrary differential form α has a form

$$\alpha = \sum_{1\le i_1<\dots<i_q\le n} \alpha_{i_1\dots i_q}\, dx_{i_1}\wedge\dots\wedge dx_{i_q}$$

where $\alpha_{i_1\dots i_q}$ is a smooth function of variables $x_1,\dots,x_n$, and $d(dx_i) = 0$, we must have

$$\begin{aligned} d\alpha &= \sum_{1\le i_1<\dots<i_q\le n} d\alpha_{i_1\dots i_q}\wedge dx_{i_1}\wedge\dots\wedge dx_{i_q} \\ &= \sum_{1\le i_1<\dots<i_q\le n}\left(\sum_{j=1}^n \frac{\partial\alpha_{i_1\dots i_q}}{\partial x_j}\, dx_j\right)\wedge dx_{i_1}\wedge\dots\wedge dx_{i_q} \\ &= \sum_{1\le i_1<\dots<i_{q+1}}\sum_{r=1}^{q+1}(-1)^{r-1}\frac{\partial\alpha_{i_1\dots\widehat{i_s}\dots i_{q+1}}}{\partial x_{i_s}} dx_{i_1}\wedge\dots\wedge dx_{i_{q+1}}. \end{aligned}$$

(we use the formula $df = \sum\frac{\partial f}{\partial x_i}dx_i$). The formulas above (actually, even the first of them) show that $d\alpha$ is uniquely defined by the conditions of Theorem.

Notice that the last formula implies some formulas well known from classical multivariable calculus. For example,

$$d(Pdx+Qdy) = \left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)dx\wedge dy,$$

$$\begin{aligned} d(Pdx+Qdy+Rdz) = \left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)dx\wedge dy + \left(\frac{\partial R}{\partial x}-\frac{\partial P}{\partial z}\right)dx\wedge dz \\ + \left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)dy\wedge dz, \end{aligned}$$

$$d(Pdy\wedge dz + Qdz\wedge dx + Rdx\wedge dy) = \left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}\right)dx\wedge dy\wedge dz.$$

6.2.2.4. Proof of existence.

We need to provide a construction of $d\alpha$ for a given α and then check the axioms (1)–(3) from Theorem. We will do it in two different ways, but will give the details of the proof only for second constructions. (Traditionally, this work is left to the reader. I prefer to give all the details, partially, to convince myself that the statement is right. Those who do not want to participate in this game, can skip the most part of the next three pages.)

The first construction stems from the formula from 6.2.2.3: using a local coordinate system $(x_1, \dots, x_n)$, we assign to a form

$$\alpha = \sum_{1 \le i_1 < \dots < i_q \le n} \alpha_{i_1 \dots i_q} \, dx_{i_1} \wedge \dots \wedge dx_{i_q}$$

its differenhtial as

$$d\alpha = \sum_{1 \le i_1 < \dots < i_{q+1}} \sum_{r=1}^{q+1} (-1)^{r-1} \frac{\partial \alpha_{i_1 \dots \widehat{i_s} \dots i_{q+1}}}{\partial x_{i_s}} dx_{i_1} \wedge \dots \wedge dx_{i_{q+1}}.$$

Besides checking axioms (1)–(3) (this is not too difficult), we need to verify that $d\alpha$ does not depend of the choice of a coordinate system. We leave this work to a reader.

The second construction uses the definition of a differential form of degree q as a function of q vector fields with values in functions (see 6.2.1). For an $\alpha \in \Omega^q M$ and $X_1, \dots, X_{q+1}$, we need to define $(d\alpha)(X_1, \dots, X_{q+1}) \in \mathcal{C}^\infty(M)$. This is done by means of the so called *Cartan formula*:

$$\begin{aligned}(d\alpha)(X_1, \dots, X_{q+1}) = & \sum_{s=1}^{q+1} (-1)^{s-1} X_s(\alpha(X_1, \dots \widehat{X}_s \dots, X_{q+1})) \\ & + \sum_{1 \le t < u \le q+1} (-1)^{t+u} \alpha([X_t, X_u], X_1, \dots \widehat{X}_t \dots \widehat{X}_u \dots, X_{q+1}).\end{aligned}$$

Example ($q = 1$). $d\alpha(X_1, X_2) = X_1(\alpha(X_2)) - X_2(\alpha(X_1)) - \alpha([X_1, X_2])$.

First, we need to check that this $(d\alpha)(X_1, \dots, X_{q+1})$ has necessary properties. Obviously, is is skew-symmetric and multilinear (over $\mathbb{R}$). It remains to check that

$$(d\alpha)(X_1, \dots, X_q, fX_{q+1}) = f(d\alpha)(X_1, \dots, X_q, X_{q+1}).$$

Here it is done:

$$\begin{aligned}(d\alpha)(X_1, \dots, X_q, fX_{q+1}) = & \sum_{s=1}^{q} (-1)^{s-1} X_s(\alpha(X_1, \dots \widehat{X}_s \dots, X_q, fX_{q+1})) \\ & + (-1)^q f X_{q+1} \alpha(X_1, \dots, X_q) \\ & + \sum_{1 \le t < u \le q} (-1)^{t+u} \alpha([X_t, X_u], X_1, \dots \widehat{X}_t \dots \widehat{X}_u \dots, X_q, fX_{q+1}) \\ & + \sum_{t=1}^{q} (-1)^{t+q+1} \alpha([X_t, fX_{q+1}], X_1, \dots \widehat{X}_t \dots, X_q)\end{aligned}$$

$$= \sum_{s=1}^{q}(-1)^{s-1}X_s(f\alpha(X_1,\dots\widehat{X}_s\dots,X_{q+1})) + (-1)^q fX_{q+1}(\alpha(X_1,\dots,X_q))$$
$$+ \sum_{1\le t<u\le q}(-1)^{t+u}f\alpha([X_t,X_u],X_1,\dots\widehat{X}_t\dots\widehat{X}_u\dots,X_{q+1})$$
$$+ \sum_{t=1}^{q}(-1)^{t+q-1}\alpha(X_t(f)X_{q+1}+f[X_t,X_{q+1}],X_1,\dots\widehat{X}_t\dots,X_q)$$
$$= \sum_{s=1}^{q}(-1)^{s-1}X_s(f)\alpha(X_1,\dots\widehat{X}_s\dots,X_{q+1}) + \sum_{s=1}^{q+1}(-1)^{s-1}fX_s(\alpha(X_1,\dots\widehat{X}_s\dots,X_{q+1}))$$
$$+ \sum_{1\le t<u\le q+1}(-1)^{t+u}f\alpha([X_t,X_u],X_1,\dots\widehat{X}_t\dots\widehat{X}_u\dots,X_{q+1})$$
$$+ \sum_{t=1}^{q}(-1)^{t}X_t(f)\alpha(X_1,\dots\widehat{X}_t\dots,X_{q+1}) = fd\alpha(X_1,\dots,X_{q+1}).$$

Next thing to do is to check the $d^2=0$ property. We have

$$(d(d\alpha))(X_1,\dots,X_{q+2}) = \sum_{s=1}^{q+2}(-1)^{s-1}X_s(d\alpha(X_1,\dots\widehat{X}_s\dots,X_{q+2}))$$
$$+ \sum_{1\le t<u\le q+2}(-1)^{t+u}d\alpha([X_t,X_u],X_1,\dots\widehat{X}_t\dots\widehat{X}_u\dots,X_{q+2})$$

and we need to plug the Cartan formula for $d\alpha$ into the right hand side of the last formula. We will obtain terms of the following five types:

(1) $X_sX_t\alpha(X_1,\dots\widehat{X}_s\dots\widehat{X}_t\dots,X_{q+2})$;
(2) $X_s\alpha([X_t,X_u],X_1,\dots\widehat{X}_s\dots\widehat{X}_t\dots\widehat{X}_u\dots,X_{q+2})$ $(t<u)$;
(3) $[X_s,X_t](\alpha(X_1,\dots\widehat{X}_s\dots\widehat{X}_t\dots,X_{q+2})$, $(s<t)$;
(4) $\alpha([\,[X_t,X_u],X_v],X_1,\dots\widehat{X}_t\dots\widehat{X}_u\dots\widehat{X}_v\dots,X_{q+2})$ $(t<u)$;
(5) $\alpha([X_t,X_u],[X_v,X_w],X_1,\dots\widehat{X}_t\dots\widehat{X}_u\dots\widehat{X}_v\dots\widehat{X}_w\dots,X_{q+2}),(t<u,v<w)$

where all $s,t,\dots$ must be different.

Term (1) appears from the first sum of the formula above. It appears with the coefficient $(-1)^{s+t-1}$, if $s<t$, and with the coefficient $(-1)^{s+t}$, if $s>t$. We combine this with term (3) which appears from the second sum with the coefficient $(-1)^{s+t}$. Together, they give for every $s<t$, $(-1)^{s+t-1}(X_sX_t-X_tX_s-[X_s,X_t])\alpha(X_1,\dots\widehat{X}_s\dots\widehat{X}_t\dots,X_{q+2})=0$. Thus, terms (1) and (3) cancel.

Term (2) appears both from the first and the second sum. In the first sum, the coefficient is $(-1)^{s+t+u-1}$, if $t<u<s$, is $(-1)^{s+t+u}$, if $t<s<u$, and is $(-1)^{s+t+u-1}$, if $s<t<u$. In the second sum, the coefficient is $(-1)^{t+u+s}$, if $s<t<u$, is $(-1)^{t+u+s-1}$, if $t<s<u$, and is $(-1)^{t+u+s}$, if $t<u<s$. Obviously, everything here cancels.

Term (4) appears from the second sum. The coefficient is $(-1)^{t+u+v+1}$, if $t<u<v$, is $(-1)^{t+u+v}$, if $t<v<u$, and is $(-1)^{t+u+v-1}$, if $t<u<v$. In other

words, for every $t < u < v$ there arises $(-1)^{t+u+v-1}\alpha([\,[X_t, X_u], X_v] - [\,[X_v, X_u], X_t] + [\,[X_v, X_t], X_u], X_1, \dots \widehat{X}_t \dots \widehat{X}_u \dots \widehat{X}_v \dots, X_{q+2}) = 0$ by the Jacobi identity.

At last, term (5) appears from the second sum. The coefficient is $(-1)^{t+u+v+w}$, if $v < w < t < u$, or $v < t < u < w$, or $t < v < w < u$, or $t < u < v < w$, and is $(-1)^{t+u+v+w-1}$, if $v < t < w < u$ or $t < v < u < w$. In other words, for every $t < u < v < w$, there arises $(-1)^{t+u+v+w}$ times

$$\begin{aligned}
&\alpha([X_t, X_u], [X_v, X_w], X_1, \dots \widehat{X}_t \dots \widehat{X}_u \dots \widehat{X}_v \dots \widehat{X}_w \dots, X_{q+2})\\
-&\alpha([X_t, X_v], [X_u, X_w], X_1, \dots \widehat{X}_t \dots \widehat{X}_u \dots \widehat{X}_v \dots \widehat{X}_w \dots, X_{q+2})\\
+&\alpha([X_t, X_w], [X_u, X_v], X_1, \dots \widehat{X}_t \dots \widehat{X}_u \dots \widehat{X}_v \dots \widehat{X}_w \dots, X_{q+2})\\
+&\alpha([X_u, X_v], [X_t, X_w], X_1, \dots \widehat{X}_t \dots \widehat{X}_u \dots \widehat{X}_v \dots \widehat{X}_w \dots, X_{q+2})\\
-&\alpha([X_u, X_w], [X_t, X_v], X_1, \dots \widehat{X}_t \dots \widehat{X}_u \dots \widehat{X}_v \dots \widehat{X}_w \dots, X_{q+2})\\
+&\alpha([X_v, X_w], [X_t, X_u], X_1, \dots \widehat{X}_t \dots \widehat{X}_u \dots \widehat{X}_v \dots \widehat{X}_w \dots, X_{q+2}).
\end{aligned}$$

Obviously, everything cancels. Thus, $d^2 = 0$.

Next, let us turn to the multiplicative property. Let $\alpha \in \Omega^q M, \beta \in \Omega^r M$, and let $X_1, \dots, X_{q+r+1} \in \mathrm{Vect}(M)$. Consider

$$(d(\alpha \wedge \beta))(X_1, \dots, X_{q+r+1}), (d(\alpha) \wedge \beta)(X_1, \dots, X_{q+r+1}), (\alpha \wedge d\beta)(X_1, \dots, X_{q+r+1}).$$

The first of these three functions contains the terms of the following four types:

(1) $(X_s(\alpha(X_I))\beta(X_J)$ where $I = \{i_1 < \dots < i_q\}$, $J = \{j_1 < \dots < j_r\}$, $\{1, \dots, q+r+1\} = I \cup J \cup \{s\}$;
(2) $\alpha(X_I)(X_s(\beta(X_J))$ where I, J, s have the same meaning;
(3) $\alpha([X_t, X_u], X_I)\beta(X_J)$ where $I = \{i_1 < \dots < i_{q-1}\}$, $J = \{j_1 < \dots < j_r\}$, $\{1, \dots, q + r + 1\} = I \cup J \cup \{t, u\}$;
(4) $\alpha(X_I)\beta([X_t, X_u], X_J)$ where $I = \{i_1 < \dots < i_q\}$, $J = \{j_1 < \dots < j_{r-1}\}$, $\{1, \dots, q + r + 1\} = I \cup J \cup \{t, u\}$

(we use abbreviated notations like $X_I = \{X_{i_1}, \dots, X_{i_q}$ for $I = \{i_1, \dots, i_q\}$, etc.). The second of the three functions above contains term (1) and (3), the third one contains the terms (2) and (4), and it remains to compare the coefficients.

We will use the following notations: for $I = \{i_1 < \dots < i_q\}$ and $s \notin I$, we put $i(s) = \#(i_u > s)$ (and similarly $j(s)$ for J instead of I). It is clear that if we insert s into $I:\ \{i_1 < \dots < s < \dots < i_q\}$, then it acquires the number $q - i(s) + 1$; it is clear also that if $\{1, \dots, q + r + 1\} = I \cup J \cup \{s\}$, then $i(s) + j(s) = q + r + s - 1$.

The term (1) appears in $(d(\alpha \wedge \beta))(X_1, \dots, X_{q+r+1})$ as a summand in $(-1)^{s-1}(\alpha \wedge \beta)(X_1, \dots \widehat{X}_s \dots, X_{q+r+1})$, and this summand corresponds to the summand $\alpha(X_{I'})\beta(X_{J'})$ of $(\alpha \wedge \beta)(X_1, \dots, X_{q+r+1})$ where I', J' are obtained from I, J in (1) by subtracting 1 from all i_u, j_v exceeding s. The last summand bears the sign $(-1)^{i_1+\dots+i_q-i(s)-\frac{q(q+1)}{2}}$ Thus, the coefficient at the term (1) in $(d(\alpha \wedge \beta))(X_1, \dots, X_{q+r+1})$ is

$$(-1)^{i_1+\dots+i_q+s-1-i(s)-\frac{q(q+1)}{2}}.$$

The term (2) appears in $(d(\alpha\wedge\beta))(X_1,\dots,X_{q+r+1})$ with the same coefficient.

The term (3) appears in $(d(\alpha\wedge\beta))(X_1,\dots,X_{q+r+1})$ as a summand in $(-1)^{t+u}(\alpha\wedge\beta)([X_s,X_t],X_1,\dots\widehat{X}_t\dots\widehat{X}_u\dots,X_{q+r+1})$, and this summand corresponds to the summand $\alpha(X_1,X_{I'})\beta(X_{J'})$ of $(\alpha\wedge\beta)(X_1,\dots,X_{q+r+1})$ where I',J' are obtained from I,J in (3) by adding 1 from all i_u,j_v less than t and subtracting 1 from all i_u,j_v exceeding u. Thus, the coefficient at the term (3) in $(d(\alpha\wedge\beta))(X_1,\dots,X_{q+r+1})$ is

$$(-1)^{t+u+1+i_1+\dots+i_{q-1}+(q-1-i(t))-i(u)-\frac{q(q+1)}{2}}.$$

For the term (4) in $(d(\alpha\wedge\beta))(X_1,\dots,X_{q+r+1})$ we repeat the same as for the term (3) with only one change: instead $\alpha(X_1,X_{I'})\beta(X_{J'})$ we take $\alpha(X_{I'})\beta(X_1,X_{J'})$ (and, certainly, take I and J from (4), not (3)). The coefficient becomes

$$(-1)^{t+u+i_1+\dots+i_q+(q-i(t))-i(u)-\frac{q(q+1)}{2}}.$$

The term (1) in $((d\alpha)\wedge\beta)(X_1,\dots,X_{q+r+1})$ appears from $(d\alpha)(X_{i_1},\dots,X_s,\dots,X_{i_q})\cdot\beta(X_J)$. Since s in $\{i_1,\dots,s,\dots,i_q\}$ has the number $q-i(s)-1$, the coefficient at this term is

$$(-1)^{q-i(s)+i_1+\dots+i_q+s-\frac{(q+1)(q+2)}{2}}.$$

The term (2) in $(\alpha\wedge d\beta)(X_1,\dots,X_{q+r+1})$ appears from $\alpha(X_I)\cdot d\beta(X_{j_1},\dots,s,\dots,X_{j_r})$. Since s in $\{j_1,\dots,s,\dots,j_r\}$ has the number $r-j(s)-1$, the coefficient at this term is

$$(-1)^{r-j(s)+i_1+\dots+i_q-\frac{q(q+1)}{2}}.$$

The term (3) in $((d\alpha)\wedge\beta)(X_1,\dots,X_{q+r+1})$ appears from $d\alpha(X_{i_1},\dots,X_t,\dots,X_u,\dots,X_{i_{q-1}})\cdot\beta(X_J)$. Since s and t in $\{i_1,\dots,t,\dots,u\dots,i_{q-1}\}$ have the numbers $q-i(t)$ and $q-i(u)+1$, the coefficient at this term is

$$(-1)^{(q-i(t))+(q-i(u)+1)+i_1+\dots+i_{q-1}+t+u-\frac{(q+1)(q+2)}{2}}.$$

Finally, the term (4) in $(d(\alpha\wedge\beta))(X_1,\dots,X_{q+r+1})$ appears from $\alpha(X_I)d\beta(X_{j_1},\dots,X_t,\dots,X_u,\dots,X_{j_{r-1}})$. Similarly to the last computation, we have the following value for the coefficient:

$$(-1)^{(r-j(t))+(r-j(u)+1)+i_1+\dots+i_q-\frac{q(q+1)}{2}}.$$

A direct comparing the coefficients gives the desired relation $d(\alpha\wedge\beta)=(d\alpha)\wedge\beta+(-1)^q\alpha\wedge d\beta$.

And, at the very last, axiom (3) holds automatically.

6.2.2.5. Differentials and smooth maps.

PROPOSITION. *Let $\varphi: M\to N$ be a smooth map, and let $\alpha\in\Omega^qN$. Then $d(\varphi^*\alpha)=\varphi^*(d\alpha)$. More algebraically: the diagram*

$$\begin{array}{ccc}\Omega^qN & \xrightarrow{\varphi^*} & \Omega^qM\\ \Big\downarrow d & & \Big\downarrow d\\ \Omega^{q+1}N & \xrightarrow{\varphi^*} & \Omega^{q+1}M\end{array}$$

is commutative.

Proof. First, it follows directly from Definition in 2.2.1 that $\varphi^*(df) = d(\varphi^* f)$ for every $f \in \mathcal{C}^\infty(N) = \Omega^0 N$. Indeed, for a $\xi \in T_p M$, $\varphi^*(df))(\xi) = df(d_p\varphi(\xi)) = (d_p\varphi(\xi))(f) = \xi(f \circ \varphi) = \xi(\varphi^*(f)) = d(\varphi^*(f))(\xi)$. Now, an arbitrary differential form $\alpha \in \Omega^q N$ is a linear combination of the forms $f\, dg_1 \wedge \ldots \wedge dg_q$; let α be this product. Then

$$\begin{aligned}\varphi^*(d\alpha) &= \varphi^*(df \wedge dg_1 \wedge \ldots \wedge dg_q) = \varphi^*(df) \wedge \varphi^*(dg_1) \wedge \ldots \wedge \varphi^*(dg_q) \\ &= d(\varphi^*(f)) \wedge d(\varphi^*(g_1)) \wedge \ldots \wedge d(\varphi^*(g_q)) = d(\varphi^*(f) \wedge d(\varphi^*(g_1)) \wedge \ldots \wedge d(\varphi^*(g_q)) \\ &= d(\varphi^*(f) \wedge \varphi^*(dg_1) \wedge \ldots \wedge \varphi^*(dg_q) = d(\varphi^*(\alpha)).\end{aligned}$$

6.2.2.6. Closed forms, exact forms, de Rham cohomology.

A differential form $\alpha \in \Omega^q M$ is called *closed*, if $d\alpha = 0$; it is called *exact*, if $\alpha = d\beta$ for some $\beta \in \Omega^{q-1}$. The equality $d^2 = 0$ means precisely that *every exact form is closed*. The inverse is not true: a closed form is not necessarily exact. The simplest example is any non-zero (locally) constant function $f \in \mathcal{C}^\infty(M) = \Omega^0 M$: it is closed, since $\Omega^{-1} M = 0$. Another example: if $f \in \mathcal{C}^\infty(S^1)$ is a function on a circle (aka a periodic function of one variable) and x is the angular coordinate on the circle, then the form $f\, dx \in \Omega^1(S^1)$ is closed (since $\Omega^2(S^1) = 0$), but it is exact if and only if $\int_{S^1} f(x)\, dx = 0$ (proof: exercise).

Let $\Omega^q_{\mathrm{cl}} M, \Omega^q_{\mathrm{ex}} M$ be the spaces of closed and exact differential forms of degree q on M. Since $\Omega^q_{\mathrm{ex}} M \subset \Omega^q_{\mathrm{cl}} M$, we can form the quotient $\Omega^q_{\mathrm{cl}} M / \Omega^q_{\mathrm{ex}} M$; this quotient is called the q-th de Rham cohomology of M and is denoted as $H^q_{\mathrm{DR}}(M)$.

The computation of the de Rham cohomology is a problem of algebraic Topology (see a discussion in Section 6.5.3). The most important fact is that for a compact manifold all the cohomology spaces are finite-dimensional. Here we single out two obvious facts: if M is connected, then $H^0_{\mathrm{DR}}(M) = \mathbb{R}$ (closed forms of degree 0 are precisely constant functions: no non-zero function is exact); $H^q_{\mathrm{DR}}(M) = 0$ for $q > \dim M$.

6.2.2.7. Other versions of the theory.

If a manifold M is not compact, we can consider differential forms with *compact support*, that is forms $\alpha \in \Omega^q M$ such that for some (not fixed) compact set $K \subset M$, $\alpha_p = 0$ for all $p \in M - K$. Forms with compact support form a subspace $\Omega^q_{\mathrm{c}} M$ of $\Omega^q M$. Obviously, $d(\Omega^q_{\mathrm{c}} M) \subset \Omega^{q+1}_{\mathrm{c}} M$; in particular, there arise "de Rham cohomology with compact support", $H^q_{\mathrm{DR,c}}(M)$ (in the definition of the latter, we use the space $\Omega^q_{\mathrm{ex,c}} M$ of forms which are differentials of forms *with compact support*, which is not the same as $\Omega^q_{\mathrm{ex}} M \cap \Omega^q_{\mathrm{c}} M$).

Also, it is possible to consider differential forms whose restrictions to a submanifold $N \subset M$ (the restriction of $\alpha|_N$ is $\iota^*\alpha$ where $\iota\colon N \to M$ is the inclusion map) or to the boundary ∂M is zero. There arise de Rham cohomologies $H^q_{\mathrm{DR}}(M, N)$ and $H^q_{\mathrm{DR}}(M, \partial M)$.

6.3. Integration. The Stokes formula.

6.3.1. Definition of integrals.

Our goal is to define, for an oriented n-dimensional manifold M and a form $\alpha \in \Omega^n M$, an integral $\int_M \alpha \in \mathbb{R}$. We will have to assume also that M is compact, or, at least. the form α is compactly supported.

6.3.1.1. The case of a domain in $\mathbb{R}^n$.

If $U \subset \mathbb{R}^n$ is a domain with compact closure $\overline{U}$ and f is a smooth function in a neighborhood of $\overline{U}$, then the integral $\int_U \alpha$ where $\alpha = f\, dx_1 \wedge \ldots \wedge dx_n$ is define as the usual integral from calculus: $\int_U f(x_1, \ldots, x_n)\, dx_1 \ldots dx_n$.

6.3.1.2. Coordinate change.

Let $x'_1, \ldots, x'_n$ be another coordinate system (in general, non-linear) around $\overline{U}$. Then

$$\begin{aligned}\alpha &= f(x'_1, \ldots, x'_n) \left(\sum\nolimits_j \frac{\partial x_1}{\partial x'_j}\, dx'_j\right) \wedge \ldots \wedge \left(\sum\nolimits_j \frac{\partial x_1}{\partial x'_j}\, dx'_j\right)\\ &= f(x'_1, \ldots, x'_n) \sum_{\tau \in S_n} \frac{\partial x_1}{\partial x'_{\tau(1)}} \cdots \frac{\partial x_n}{\partial x'_{\tau(n)}}\, dx'_{\tau(1)} \wedge \ldots \wedge dx'_{\tau(n)}\\ &= f(x'_1, \ldots, x'_n) \sum_{\tau \in S_n} \operatorname{sgn}(\tau) \cdot \frac{\partial x_{\tau(1)}}{\partial x'_1} \cdots \frac{\partial x_{\tau(n)}}{\partial x'_n}\, dx'_1 \wedge \ldots \wedge dx'_n\\ &= f(x'_1, \ldots, x'_n) \cdot \det \left\| \frac{\partial x_i}{\partial x'_j} \right\|\, dx'_1 \wedge \ldots \wedge dx'_n.\end{aligned}$$

On the other hand, a formula from calculus says that

$$\int_U f(x_1, \ldots, x_n)\, dx_1 \ldots dx_n = \int_U f(x'_1, \ldots, x'_n)\, \operatorname{abs} \left(\det \left\| \frac{\partial x_i}{\partial x'_j} \right\|\right)\, dx'_1 \ldots dx'_n.$$

Comparing these two formulas we conclude that $\int_U \alpha$ is preserved by coordinate change with positive Jacobian; in other words, $\int_U \alpha$ is dertermined by U, α, and orientation of U.

6.3.1.3. General case.

Let M be an oriented n-dimensional manifold, and let $\alpha \in \Omega^n M$ be a compactly supported differential form. Let $\{(U_i, \varphi_i)\}$ be an oriented atlas of M belonging to the chosen orientation and such that all $\overline{U}_i$ are compact and $\varphi_i(U_i)$ form a compact, locally finite covering of M. Let $\{f_i \colon M \to [0,1]\}$ be a partition of unity subordinated to this covering. First, we define $\int_M (f_i \alpha)$ as in 6.3.1.1, using the local coordinates corresponding to the chart (U_i, φ_i) (in other words, $\int_M f_i \alpha = \int_{U_i} \varphi_i^*(f_i \alpha)$ as defined in 6.3.1.1). Then we put $\int_M \alpha = \sum_i \int_M f_i \alpha$.

Proposition. *The integral $\int_M \alpha$ is well defined.*

Proof. First, since the form α is compactly supported, he sum in this definition is finite. Let $\{(V_j, \psi_j)\}$ and $\{g_j\}$ satisfy the same conditions as $\{(U_i, \varphi_i)\}$ and $\{f_i\}$. We denote the integrals defined using local coordinates from these two atlases (and thses two partitions of unity) as $\int^U$ and $\int^V$. We have:

$$\int_M^U \alpha = \sum_i \int_M^U f_i \alpha = \sum_{i,j} \int_M^U f_i g_j \alpha = \sum_{i,j} \int_M^V f_i g_j \alpha = \sum_j \int_M^V g_j \alpha = \int_M^V \alpha$$

(the middle equality follows from the result of 6.3.1.2: the integrals $\int_M^U f_i g_j \alpha, \int_M^V f_i g_j \alpha$ have the same integrand and are taken over the same domain with compact closure covered by two orientably compatible charts, (U_i, φ_i) and (V_j, ψ_j)).

6.3.1.4. Further generalizations.

First, it is OK, if M has a boundary.

Second, if $\alpha \in \Omega^q M$, and $q < n$, then $\int_M \alpha$ does not exist, but there exist integrals $\int_N \alpha = \int_N \iota^* \alpha|_N$ for all oriented q-dimensional submanifolds N of M where $\alpha|_N$ is the restriction of α to N (see 6.2.2.7). (Notice that M does not need to be oriented or even orientable.) Certainly, we require that N is compact, or, at least, $\iota^*\alpha$ has a compact support. It is more common to call $\iota^*\alpha$ the *restriction* of α

We do not exclude the case when $q = 0$. In this case, N is a discrete subset of M; the orientation is a function $\varepsilon \colon N \to \{\pm 1\}$ (see 4.2.2); a form α of degree q is a function; the compactness condition means that α can be non-zero only at finitely many points of N (which holds, if N is finite). The integral $\int_N \alpha$ is defined as $\sum_{p \in N} \varepsilon(p) \cdot \alpha(p)$.

6.4. The Stokes Theorem.

6.4.1. The statement.

THEOREM (the Stokes Theorem). *Let M be an n-dimensional manifold, let $\alpha \in \Omega^{q-1} M$ $(q \geq 1)$, and let N be an oriented submanifold of M such that the restrictions $\alpha|_{\partial N}$ and $(d\alpha)|_N$ are compactly supported. (We assume that ∂N is oriented accordingly to the orientation of N.) Then*

$$\int_N d\alpha = \int_{\partial N} \alpha.$$

A poof will be given in 6.4.3.

6.4.2. Particular cases.

Some particular cases of this theorem are well known in Analysis and have canonical names. We always assume that $M = \mathbb{R}^n$ and $1 \leq q \leq n$.

CASE $n = 1, q = 1$. Let $N = [a, b]$, $\alpha = f$, and N is oriented in such a way that $\partial N = b - a$. Since $d\alpha = f'(x)\, dx$, the Stokes theorem assumes the form

$$\int_a^b f'(x)\, dx = f(b) - f(a).$$

This is the fundamental theorem of calculus.

CASE $n = 2, q = 2$. Let $N = U \subset \mathbb{R}^2$ be a domain bounded by a closed smooth non-self-intersecting curve $\gamma = \partial N$ (which may be disconnected). If N is oriented accordingly to the standard orientation of the plane, then the outer component(s) of γ are oriented counterclockwise and the inner components of γ are oriented clockwise. If $\alpha = f(x, y)\, dx + g(x, y)\, dy$, then $d\alpha = \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dx \wedge dy$. The Stokes theorem assumes the form

$$\int_\gamma (f\, dx + g\, dy) = \int_U \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dx \wedge dy.$$

This is the Green Theorem (in classical Analysis the symbol "$\wedge$" is usually not used).

CASE $n = 3, q = 2$. In this case N is a domain S on an oriented surface in $\mathbb{R}^3$ bounded by a closed curve γ; the orientations of S and γ should be compatible. If $\alpha = f\,dx + g\,dy + h\,dz$, then $d\alpha = \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dx \wedge dy + \left(\frac{\partial h}{\partial x} - \frac{\partial f}{\partial z}\right) dx \wedge dz + \left(\frac{\partial h}{\partial y} - \frac{\partial g}{\partial z}\right) dy \wedge dz$, and the Stokes theorem assumes the form

$$\int_\gamma (f\,dx + g\,dy + h\,dz) \\ = \int_S \left[\left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dx \wedge dy + \left(\frac{\partial h}{\partial x} - \frac{\partial f}{\partial z}\right) dx \wedge dz + \left(\frac{\partial h}{\partial y} - \frac{\partial g}{\partial z}\right) dy \wedge dz\right].$$

In classical Analysis, this theorem (and not Theorem of 6.4.1) is attributed to Stokes.

CASE $n = 3, q = 3$. In this case $N = U \subset \mathbf{R}^3$ is a domain bounded by a closed surface S (U and S are oriented coherently), $\alpha = f\,dy \wedge dz + g\,dz \wedge dx + h\,dx \wedge dy$, $d\alpha = \left(\frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} + \frac{\partial h}{\partial z}\right) dx \wedge dy \wedge dz$, and the Stokes theorem assumes the form

$$\int_S (f\,dy \wedge dz + g\,dz \wedge dx + h\,dx \wedge dy) = \int_U \left(\frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} + \frac{\partial h}{\partial z}\right) dx \wedge dy \wedge dz.$$

A generic name for this result is the Gauss-Ostrogradski Theorem.

COMMENTS REGARDING THE NAMES. The fundamental theorem of calculus is usually attributed to Isaac Newton and Gottfried Leibnitz (who were involved in a hot argument concerning the authorship of this result). Actually, it was known before these illustrious figures; the first published proofs belong to James Gregory and Isaac Barrow (1670). Green's theorem is named after the English physicist George Green who used this formula in his works (1828); as a mathematical theorem it was first singled out by Cauchy (1846). The Gauss-Ostrogradski formula was published by Mikhail Ostrogradski (in Paris) in 1826; as early as in 1813, it was used by Carl-Friedrich Gauss in his work on gravitation of ellipsoids (Gauss had so many results that it was impossible for him to publish them all in separate papers). (It is amusing that the Gauss-Ostogradski formula appeared before the Green formula which may be regarded as a particular case of the Gauss-Ostrogradski formula.) The Stokes formula was first published by George Stokes (1854) in his account of the annual student's contest for the Smith fellowship in Cambridge. Stokes was responsible for this contest for many years, and not all the problems of this contest belonged to him. In particular, what we call today the Stokes formula, was communicated to Stokes by William Thompson (1850) (who claimed later that he deduced this formula from an earlier work of Stokes). The language of differential forms and their differentials which we use here was developed in 1899 by Elie Cartan, and the theorem which we call the Stokes theorem here was first proved by Edouard Goursat in 1917 (although, in some form, it was apparently known much earlier to Elie Cartan).

6.4.3. Proof of the Stokes theorem.

This is sort of unfair that such a deep theorem has such a simple proof. I will try to give all the details, just to make the proof look slightly longer.

We can assume that α is not zero only within one chart (U, φ), and this $U \subset \mathbb{R}^n_-$ is contained in the cube $I^n = \{(x_1, \dots, x_n) \mid -1 \le x_i \le 0 \forall i$ and is disjoint from all the planes $x_i = -1$ and all the planes $x_i = 0$, except, possibly $x_n = 0$. We need to prove that for $\alpha = \sum_{i=1}^n f_i(x_1, \dots, x_n)\, dx_1 \wedge \dots \widehat{dx_i} \dots \wedge dx_n$, $\int_{\mathbb{R}^n_-} d\alpha = \int_{\mathbb{R}^{n-1}} \alpha|_{\mathbb{R}^{n-1}}$. (We assume that α is defined in $\mathbb{R}^n_-$ but is zero in the complement of I^n.) First we notice that

$$\alpha|_{\mathbb{R}^{n-1}} = f_n(x_1, \dots, x_{n-1}, 0)\, dx_1 \wedge \dots \wedge dx_{n-1}$$
$$d\alpha = \sum_{i=1}^n \frac{\partial f_i}{\partial x_i} dx_i \wedge dx_1 \wedge \dots \widehat{dx_i} \dots \wedge dx^n.$$

From this,

$$\begin{aligned}\int_{\mathbb{R}^n_-} d\alpha &= \sum_{i=1}^n \int_{-1}^0 \dots \int_{-1}^0 \left(\int_{-1}^0 \frac{\partial f_i}{\partial x_i}(x_1, \dots, x_n)\, dx_i \right) dx_1 \dots \widehat{dx_i} \dots dx_n \\ &= \sum_{i=1}^n \int_{-1}^0 \dots \int_{-1}^0 (f_i \mid_{x_i=0} - f_i \mid_{x_i=-1}) dx_1 \dots \widehat{dx_i} \dots dx_n \\ &= \int_{-1}^0 \dots \int_{-1}^0 f_n(x_1, \dots, x_{n-1}, 0)\, dx_1 \dots dx_{n-1} = \int_{\mathbb{R}^{n-1}} \alpha|_{\mathbb{R}^{n-1}}.\end{aligned}$$

6.4.4. Integrals of closed and exact forms.

Proposition 1. *Let M be an oriented compact m-dimensional submanifold without boundary of an n-dimensional manifold N, and let $\alpha \in \Omega^m N$.*

(1) *If α is exact, then $\int_M \alpha = 0$.*

(2) *If α is closed and $M = \partial P$ where P is an oriented compact $(m+1)$-dimensional submanifold of N, then $\int_M \alpha = 0$.*

Proof. (1) If $\alpha = d\beta$, then $\int_M \alpha = \int_M d\beta = \int_{\partial M} \beta = \int_\emptyset \beta = 0$.

(2) If $M = \partial P$, then $\int_M \alpha = \int_P d\alpha = \int_P 0 = 0$.

Proposition 1 creates a possibility to distinguish between exact and closed forms: if a differential form has a non-zero integral over a closed (= compact, without boundary) manifold, then it cannot be exact (although can be closed).

Examples. (1) Let $N = \mathbb{R}^2 - (0,0)$ and let $\alpha = \dfrac{xdy - ydx}{x^2 + y^2}$. This form is closed but not exact, since if S is a counterclockwise oriented unit circle centered at the origin, then $\int_S \alpha = 2\pi \ne 0$. (Actually, $\alpha = d\theta$ where θ is the polar angle; this implies closedness of α, but not exactness, since θ is not a (univalent) function on N; the same fact gives a computation of the above integral.

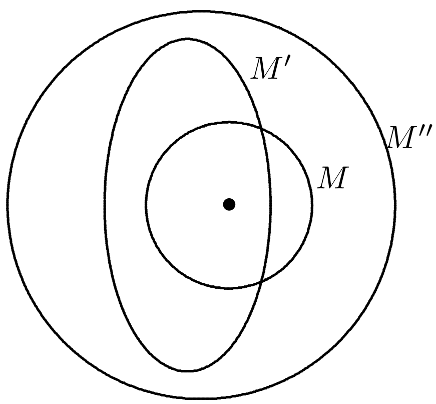
(2). Let M be an oriented closed m-dimensional ($m > 0$) submanifold of $\mathbb{R}^n$. For $v_1, \dots, v_m \in T_pM$, define $vol(v_1, \dots, v_m)$ as a signed volume of the m-dimensional parallelepiped spanned by $v_1, \dots, v_m$ ("signed" means that we take it with the sign $+$, if it induces in T_pM orientation compatible with the orientation of M, and with the sign $-$ otherwise). It is easy to understand that $vol \in \Omega^m M$, and that $\int_M vol$ is the volume of

M (a positive number). The form vol is closed (simply because $\Omega^{m+1}M = 0$) and is not exact by Proposition 1.

Finally, let M, M' be closed oriented m-dimensional submanifolds of N and α, α' be closed degree m differential forms of N. We say that M, M' *cobound*, if there exists a compact oriented $(m+1)$-dimensional submanifold P of N such that $\partial P = M' - M$ (the minus sign means that the orientations of ∂P and M are opposite). We say that the forms α, α' are cohomological, if the difference between them is exact, $\alpha' - \alpha = d\beta$; in other words, α and α' are cohomological, if they represent the same element of $H^m_{\mathrm{DR}}(N)$.

PROPOSITION 2. If M, M' cobound and α, α' are cohomological, then

$$\int_M \alpha = \int_{M'} \alpha'.$$



Proof. $\int_M \alpha' - \int_M \alpha = \int_M d\beta = 0$ by (1) in Proposition 1. $\int_{M'} \alpha' - \int_M \alpha' = \int_{\partial P} \alpha' = 0$ by (2) in Proposition 1. Hence, $\int_M \alpha = \int_{M'} \alpha'$.

Thus, we can speak of integrals of de Rham cohomology classes over closed manifolds given up to a cobounding. Notice that it may happen that M, M' do not cobound, but each of them cobounds with some M'' (see the picture on the left); then the statement of Proposition 2 still holds).

6.5. More on de Rham cohomology.

6.5.1. Homotopy invariance.

6.5.1.1. Induced maps.

A smooth map $f\colon M \to N$ between two smooth manifolds induces homomorphisms $f^*\Omega^q N \to \Omega^q M$ (see 6.2.1) commuting with differentials (see 6.2.2.5). Hence, $f^*(\Omega^q_{\mathrm{cl}} N) \subset \Omega^q_{\mathrm{cl}} M$ (if $d\alpha = 0$, then $df^*\alpha = f^*d\alpha = 0$) and $f^*(\Omega^q_{\mathrm{ex}} N) \subset \Omega^q_{\mathrm{ex}} M$ (if $\alpha = d\beta$, then $f^*\alpha = f^*d\beta = df^*\beta$). Hence, f^* gives rise to a linear map

$$H^q_{\mathrm{DR}}(N) = \Omega^q_{\mathrm{al}} N / \Omega^q_{\mathrm{ex}} N \to \Omega^q_{\mathrm{al}} M / \Omega^q_{\mathrm{ex}} M = H^q_{\mathrm{DR}}(M)$$

which is also denoted as f^*. It is obvious that $(f \circ g)^* = g^* \circ f^*$ and $\mathrm{id}^* = \mathrm{id}$.

EXERCISE. If M and N are connected, then, for every $f\colon M \to N$, $f^*\colon H^0_{\mathrm{DR}}(N) \to H^0_{\mathrm{DR}}(M)$ is an isomorphism.

6.5.1.2. Induced maps and homotopies.

Two smooth maps $f, g\colon M \to N$ are called *homotopic* (notation: $f \sim g$), if there exists a (smooth) *homotopy* $F\colon M \times \mathbb{R} \to N$ (or $F\colon M \times [0,1] \to N$) *joining* f and g, that is such that $F|_{M\times 0} = f$, $F|_{M\times 1} = g$.

THEOREM. *If $f \sim g$, then $f^* = g^*$.*

Proof. We need to show that for any closed form $\alpha \in \Omega^q N$, the forms $f^*\alpha, g^*\alpha$ are cohomologous, that is, the difference $g^*\alpha - f^*\alpha$ is exact (this means that for a representative α of arbitrary $a \in H^q_{\mathrm{DR}}(N)$, the forms $f^*\alpha, g^*\alpha$ represent the same element of $H^q_{\mathrm{DR}}(M)$, that is, $f^*(a) = g^*(a)$). Let $F: M \times \mathbb{R} \to N$ be a homotopy joining f and g, and let $\iota_t: M \to M \times \mathbb{R}$ be the embedding $p \mapsto (p, t)$. Then $f = F \circ \iota_0, g = F \circ \iota_1$. We need to establish that $\iota^*(F^*\alpha)$ cohomologous to $\iota_1^*(F^*\alpha)$. We will do more: we will construct ($\mathcal{G}^\infty$-continuous) linear maps

$$h = h_q: \Omega^q(M \times \mathbb{R}) \to \Omega^{q-1}(M)$$

such that for any $\beta \in \Omega^q N$,

$$dh(\beta) + h(d\beta) = \iota_1^*(\beta) - \iota_0^*(\beta);$$

applying this to $\beta = F^*\alpha$ and using the fact that $d\beta = F^* d\alpha = 0$, we get the necessary result.

We will need one more notation. Let P be a manifold, $\gamma \in \Omega^r P$, and $X \in \operatorname{Vect} P$; then the form $\eta_X \gamma \in \Omega^{r-1} P$ is defined by the formula

$$\eta_X \gamma(X_1, \ldots, X_{r-1}) = \gamma(X, X_1, \ldots X_{r-1}).$$

Obviously, $\eta_X(df) = Xf$ and $\eta_X(\gamma \wedge \gamma') = \big(\eta_X \gamma\big) \wedge \gamma' + (-1)^{\deg \gamma} \gamma \wedge \eta_X \gamma'$.

We put

$$h(\beta) = \int_0^1 \left(\eta_{\frac{\partial}{\partial t}} \beta|_{M \times t}\right) dt = \int_0^1 \left[\iota_t^* \eta_{\frac{\partial}{\partial t}} \beta\right] dt$$

(here $\dfrac{\partial}{\partial t}$ is the "vertical vector field" on $M \times \mathbb{R}$, that is the derivative with respect to the coordinate t on $\mathbb{R}$). Let us prove the promised equality. For β, we take a monomial form, and we distinguish two cases:

(1) $\beta = f(x,t)\, dt \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_{q-1}}$,
(2) $\beta = f(x,t)\, dx_{i_1} \wedge \ldots \wedge dx_{i_q}$

(notice, that the operator $\eta_{\frac{\partial}{\partial t}}$ applied to a monomial form, eliminates dt, if there is a dt, and yields 0, if there is no dt).

Case (1). In this case,

$$\begin{aligned}
\iota_t^* \beta &= 0 \qquad (\text{since } \iota_t^*(dt) = 0),\\
\eta_{\frac{\partial}{\partial t}} \beta &= f(x,t)\, dx_{i_1} \wedge \ldots \wedge dx_{i_{q-1}},\\
d\beta &= \sum_j \frac{\partial f}{\partial x_j}(x,t)\, dx_j \wedge dt \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_{q-1}},\\
\iota_t^* \eta_{\frac{\partial}{\partial t}} d\beta &= -d\iota_t^* \eta_{\frac{\partial}{\partial t}} \beta \Rightarrow hd\beta = -dh\beta \Rightarrow dh\beta + hd\beta = 0 = \iota_1^*\beta - \iota_0^*\beta.
\end{aligned}$$

Case (2). In this case,

$$\begin{aligned}
\eta_{\frac{\partial}{\partial t}}\beta &= 0 \Rightarrow h\beta = 0,\\
d\beta &= \frac{\partial}{\partial t} f(x,t)\, dt \wedge dx_{i_1} \wedge \ldots \wedge dx_{i_q} + \text{terms without } dt\\
\eta_{\frac{\partial}{\partial t}} d\beta &= \frac{\partial}{\partial t} f(x,t)\, dx_{i_1} \wedge \ldots \wedge dx_{i_q},\\
hd\beta &= \left(\int_0^1 \frac{\partial}{\partial t} f(x,t) dt\right) dx_{i_1} \wedge \ldots \wedge dx_{i_q}\\
&= [f(x,1) - f(x,0)] dx_{i_1} \wedge \ldots \wedge dx_{i_q} = \iota_1^*\beta - \iota_0^*\beta.
\end{aligned}$$

6.5.1.3. De Rham cohomology and homotopy equivalences.

A smooth map $f: M \to N$ is called a *homotopy equivalence*, if there exists a smooth map $g: N \to M$ such that $g \circ f \sim \mathrm{id}_M$, $f \circ g \sim \mathrm{id}_N$. Manifolds M, N are called *homotopy equivalent*, if there exists a homotopy equivalence $M \to N$. (We mentioned homotopy equivalences before, in 5.4.6, but at this moment we need more details.)

Examples. (1) $\mathbb{R}^n$ is homotopy equivalent to a point (as they say, *contractible*). Indeed, let $f: \mathrm{pt} \to \mathbb{R}^n$, $f(\mathrm{pt}) = 0$, and let g be the map $\mathbb{R}^n \to \mathrm{pt}$. Then $g \circ f = \mathrm{id}$ and $(f \circ g)(\mathbb{R}^n) = 0$; the latter is homotopic to id by the homotopy $(x,t) \mapsto tx$.

(2) $\mathbb{R}^n - 0$ is homotopy equivalent to S^{n-1}. Indeed, let $f: S^{n-1} \to \mathbb{R}^n - 0$ be the inclusion map, and $g: \mathbb{R}^n - 0 \to S^{n-1}$, $g(x) = \dfrac{x}{\|x\|}$. Then $g \circ f = \mathrm{id}$, and $f \circ g$ is homotopic to id by the homotopy $(x,t) = x \cdot \|x\|^{-t}$.

Theorem. *If $f: M \to N$ is a homotopy equivalence, then $f^* H^q_{\mathrm{DR}}(N) \to H^q_{\mathrm{DR}}(M)$ is an isomorphism. In partiucular, de Rham cohomologies of homotopy equivalent manifolds are isomorphic.*

Proof. Let $g \circ f \sim \mathrm{id}_M$ and $f \circ g \sim \mathrm{id}_N$. Then $f^* \circ g^* = (g \circ f)^* = \mathrm{id}^a\, st = \mathrm{id}$ and, similarly, $g^* \circ f^* = \mathrm{id}$. Thus, f^*, g^* are inverse to each other.

6.5.1.4. Application: the Poincaré lemma.

Theorem (the Poincaré lemma). *In $\Omega^q\mathbb{R}^n$, $q > 0$, every closed form is exact. Equivalently: $H^q_{\mathrm{DR}}(\mathbb{R}^n) = 0$.*

Proof. Theorem and Example (1) from 6.5.1.4 show that $H^q(\mathbb{R}^n) = H^q(\mathrm{pt})$. On the other hand, $H^q(\mathrm{pt}) = 0$ for $q > \dim \mathrm{pt} = 0$.

Corollary. *On any manifold, closed forms of positive degree are locally exact.*

Proof. Every point of any n-dimensional manifold M is covered by a chart (U, φ) with U diffeomorphic to $\mathbb{R}^n$. Hence, for any closed form $\alpha \in \Omega^q M$, $q > 0$, the restriction $\alpha|_{\varphi(U)}$ is exact.

6.5.2. Examples of computation of de Rham cohomology.

The Poincaré lemma, together with computations of $H^q_{\mathrm{DR}}(M)$ for $q = 0$ and $q > \dim M$, (see 6.2.2.6) provide examples of such computations. We will consider more examples here and will briefly discuss the general case in 6.5.3.

6.5.2.1. De Rham cohomology of spheres.

Theorem. *Any closed form in $\Omega^q S^n$ with $q \neq 0, n$ is exact; a closed form $\alpha \in \Omega^n S^n$ is exact if and only if $\int_{S^n} \alpha = 0$.*

Corollary. *If $n > 0$, then*

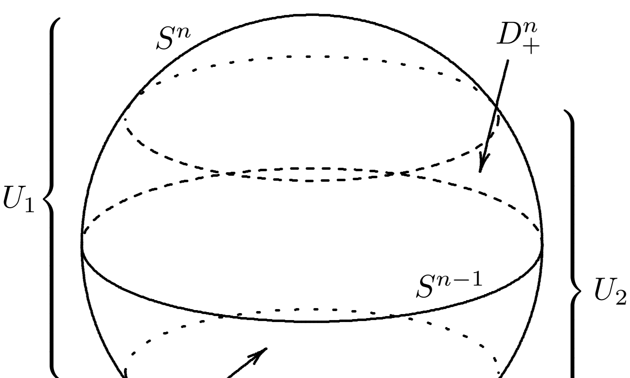
$$H^q_{\mathrm{DR}}(S^n) \begin{cases} \cong \mathbb{R}, & \text{if } q = 0, n, \\ = 0, & \text{if } q \neq 0, n. \end{cases}$$

Proof of Corollary. The only thing we need to add to Theorem is that there exists a form $\alpha \in \Omega^n S^n$ with a non-zero integral. A construction of such form ("the volume form") is given in Example (2) of 6.4.4.

Proof of Theorem. The *only if* part of the last statement of Theorem is known to us: if α is exact, then $\int_{S^n} \alpha = 0$ (see Proposition 1 of 6.4.4). All the rest is proved by induction with respect to n. It is possible to take for the base case $n = 0$, but this would require a slight modification of the statement in this case, so we begin with the case $n = 1$.

If $n = 1$, then a 1-form has the form $\alpha = f(\theta)\, d\theta$, or $f(x)\, dx$ where f is a 2π-periodic function of 1 variable. The last form is the differential of the function $g(x) = \int_0^x f(t)\, dt$, and the latter is 2π-periodic if and only if $\int_0^{2\pi} f(t)\, dt = \int_{S^1} \alpha = 0$.

Assume now that Theorem has been proved for S^{n-1} and take a closed form $\alpha \in \Omega^q S^n, q > 0$ such that if $q = n$, then $\int_{S^n} \alpha = 0$. Cover S^n by two open sets, $U_1 = \left\{ x_{n+1} > -\frac{1}{2} \right\}$ and $U_2 = \left\{ x_{n+1} < \frac{1}{2} \right\}$ (see the picture on the left). Then the forms $\alpha|_{U_i} \in \Omega^q U_i$ are exact by the Poincaré lemma, $\alpha|_{U_i} = d\beta_i,\ \beta_i \in \Omega^{q-1} U_i$. The form $\beta_2 - \beta_1 \in \Omega^{q-1}(U_1 \cap U_2)$ is closed (since $d\beta_1 = d\beta_2$). The intersection $U_1 \cap U_2$ is homotopy equivalent to S^{n-1} (and the inclusion map $S^{n-1} \to U_1 \cap U_2$ is a homotopy equivalence). Hence, if $q-1 \neq 0, n-1$, then the form $\beta_2 - \beta_1$ is exact (by the induction hypothesis). The same is true, however, for $q - 1 = 0, n - 1$: if $q - 1 = 0$, then $\beta_2 - \beta_1$ is a constant, and we can make it zero by adding an appropriate constant to β_1. If $q - 1 = n - 1$, then $\int_{S^{n-1}} \beta_1 = \int_{D^n_+} d\beta_1 = \int_{D^n_+} \alpha$ and, similarly, $\int_{S^{n-1}} \beta_2 = \int_{D^n_-} \alpha$ where $D^n_\pm$ are oriented by the orthogonal projections $D^n_\pm \to D^n$. But $0 = \int_{S^n} \alpha = \int_{D^n_+} \alpha - \int_{D^n_-} \alpha$ (since S^n is oriented coherently with D^n_+ and non-coherently with D^n_-). Hence $\int_{D^n_+} \alpha = \int_{D^n_-} \alpha,\ \int_{S^{n-1}} \beta_1 = \int_{S^{n-1}} \beta_2 \Rightarrow \int_{S^{n-1}} (\beta_2 - \beta_1) = 0$. Thus, the form $\beta_2 - \beta_1$ is exact in all cases; let $\beta_2 - \beta_1 = d\gamma,\ \gamma \in \Omega^{q-2}(U_1 \cap U_2)$.



Let $f_1, f_2 \colon S^n \to [0,1]$ be smooth functions such that $\operatorname{supp} f_i \subset U_i$ and $f_1 + f_2 = 1$

(see 1.3.3.4). Define $\varepsilon_i \in \Omega^{q-2}U_i$ by the formulas

$$\varepsilon_1 = \begin{cases} f_2\gamma & \text{on } U_1 \cap U_2, \\ 0 & \text{on } U_1 - U_2, \end{cases} \qquad \varepsilon_2 = \begin{cases} f_1\gamma & \text{on } U_1 \cap U_2, \\ 0 & \text{on } U_2 - U_1. \end{cases}$$

and put $\beta_1' = \beta_1 - d\varepsilon_1$, $\beta_2' = \beta_2 + d\varepsilon_2$. Then, on $U_1 \cap U_2$,

$$\beta_1' - \beta_2' = \beta_1 - d\varepsilon_1 - \beta_2 - d\varepsilon_2 = \beta_1 - \beta_2 - d(f_1 + f_2)\gamma - \beta_1 - \beta_2 - d\gamma = 0,$$

thus $\beta'1$ and β_2 agree on $U_1 \cap U_2$, thus they merge into one form $\beta' \in \Omega^{q-1}S^n$ and

$$d\beta' = \begin{cases} d\beta_1' = d\beta_1 = \alpha & \text{on } U_1, \\ d\beta_2' = d\beta_2 = \alpha & \text{on } U_2, \end{cases}$$

that is, $d\beta' = \alpha$ and the form α is exact.

6.5.2.2. Compactly supported cohomology of $\mathbb{R}^n$.

Theorem. *Let $\alpha \in \Omega^q_{\text{cl,c}}\mathbb{R}^n$, that is, α is a closed compactly supported form; if $q = n$, assume that $\int_{\mathbb{R}^n} \alpha = 0$. Then $\alpha \in \Omega^q_{\text{ex,c}}{}^n$, that is, α is a differential of a compactly supported form.*

Proof. If $q = 0$, then a closed form is a constant function, and if it is compactly supported, then it is zero. If $n = q = 1$, then $\alpha = f(x)\,dx$ and $\alpha = dg$ where $g(x) = \int_{-\infty} f(t)\,dt$, and if $\int_{-\infty}^{\infty} f(t)\,dt = 0$, then g is a compactly supported function. So we need to consider the cases when $q > 0$ and $n > 1$. Let $\operatorname{supp}\alpha \subset B$ where B is an open ball in $\mathbb{R}^n$ centered at 0. By the Poincaré lemma, $\alpha = d\beta$ where $\beta \in \Omega^{q-1}\mathbb{R}^n$. Since $\operatorname{supp}\alpha \subset B$, β is closed in $\mathbb{R}^n - B$. But $\mathbb{R}^n - B$ is homotopy equivalent to $\partial\overline{B} \approx S^{n-1}$. Hence, if $0 < q - 1 < n - 1$, then $\beta \in \Omega^{q-1}(\mathbb{R}^n - B)$ is exact, $\beta = d\gamma$, $\gamma \in \Omega^{q-2}(\mathbb{R}^n - B)$ (see 6.5.2.1). The same is true, if $q - 1 = 0$ or $n - 1$. Indeed, if $q - 1 = 0$, then β is a constant function in $\mathbb{R}^n - B$, and we can make this constant zero by subtracting a constant from $\beta \in \Omega^{q-1}\mathbb{R}^n$. If $q - 1 = n - 1$, then $\int_{\partial\overline{B}} \beta = \int_{\overline{B}} d\beta = \int_{\overline{B}} \alpha = \int_{\mathbb{R}^n} \alpha = 0$, thus again β is exact in $\mathbb{R}^n - B$ (see 6.5.2.1).

Choose a smooth function $f\colon \mathbb{R}^n \to [0, 1]$ equal to 0 in a neighborhood of $\overline{B}$ and equal to 1 in the complement of an open ball $B' \supset \overline{B}$. Let $\beta' = \beta - d(f\gamma)$ (we can regard $f\gamma$ as a form defined in the whole space $\mathbb{R}^n$). Then $d\beta' = d\beta = \alpha$ and $\beta' = \beta - d\gamma = 0$ in $\mathbb{R}^n - B'$, thus β' is compactly supported.

Corollary.

$$H^q_{\text{DR,c}}(\mathbb{R}^n) \begin{cases} \cong \mathbb{R}, & \text{if } q = n, \\ = 0, & \text{if } q \neq n. \end{cases}$$

Proof. In addition to Theorem, we need only to remark that any form from $\Omega^n_{\text{ex,c}}\mathbb{R}^n$ has zero integral over $\mathbb{R}^n$ and that there are forms in $\Omega^n_{\text{c}}\mathbb{R}^n$ with a non-zero integral over $\mathbb{R}^n$. The first follows from the Stokes theorem (if $\alpha = d\beta, \operatorname{supp}\beta \subset B$, then $\int_{\mathbb{R}}^n \alpha = \int_{\overline{B}} \alpha = \int_{\overline{B}} d\beta = \int_{\partial\overline{B}} \beta = 0$), for the second, take $\alpha = f(x)\,dx_1 \wedge \ldots \wedge dx_n$ where f is a compactly supported non-negative, non-zero function.

6.5.2.3. Highest degree cohomology.

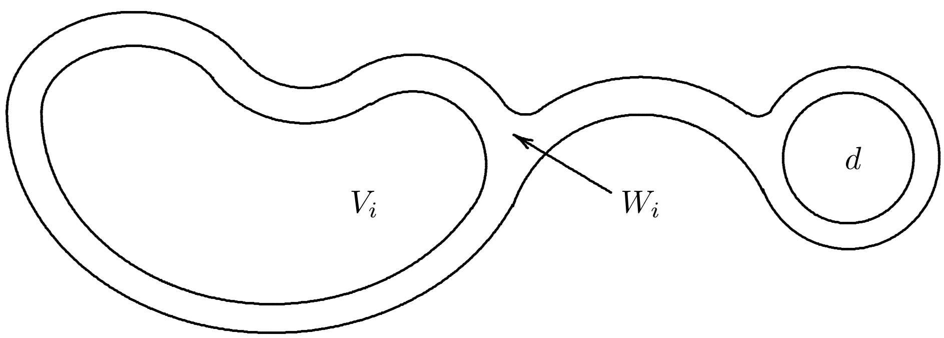
6.5.2.3.1. The case of compact orientable manifolds.

THEOREM. *Let M be an n-dimensional connected compact oriented manifold. A differential form α of degree n on M is exact if and only if $\int_M \alpha = 0$.*

Proof. The *only if* part is contained in 6.4.4: if α is exact, then $\int_M \alpha = 0$. Let now $\int_M \alpha = 0$. Our goal is to construct a decomposition $\alpha = \sum_{i=1}^N \alpha_i$ such that for every i, (1) $\operatorname{supp} \alpha_i$ is contained in an open set $W_i \subset M$ diffeomorphic to $\mathbb{R}^n$, and (2) $\int_{W_i} \alpha_i = \int_M \alpha_i = 0$. As soon as we do it, we can complete the proof of Theorem: Theorem of 6.5.2.2 implies that there is a $\beta_i \in \Omega^{n-1}M$ with $\operatorname{supp} \beta_i \subset W_i$ and $d\beta_i = \alpha_i$. It remains to put $\beta = \sum \beta_i$ and to observe that $d\beta = \alpha$.

We begin with an oriented atlas $\{(U_i, \varphi_i)\}$ of M with all U_i diffeomorphic to $\mathbb{R}^n$. Then we shrink the cover $\{\varphi_i(U_i)\}$ of M to $\{V_i\}$ with compact $\overline{V}_i \subset U_i$ and take the partition of unity $\{f_i\}$ subordinated to $\{V_i\}$. Then we put $\alpha'_i = f_i\alpha$; the only reason why we cannot take $\alpha_i = \alpha'_i$ is that the integrals $c_i = \int_{V_i} \alpha = \int_M \alpha'_i$ have no reason to be zero. All we have is $\sum_i c_i = \int_M (\sum_i \alpha'_i) = \int_M \alpha = 0$.

Choose a small open ball $d \subset M$ and a form $\gamma \in \Omega^n M$ with $\operatorname{supp} \gamma \subset d$ such that $\int_d \gamma = \int_M \gamma = 1$ (obviously, exists). Then we take for W_i a thin neighborhood of V_i, d and a path between these two domains (see the picture below).



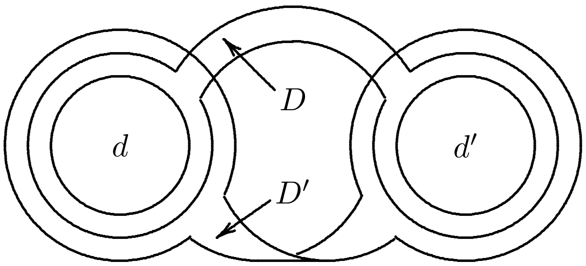
The form α_i is defined as α'_i on V_i, as $-c_i\gamma$ on d. The conditions above are all satisfied ($\sum_i \alpha_i = \alpha$, because $\sum_i c_i = 0$).

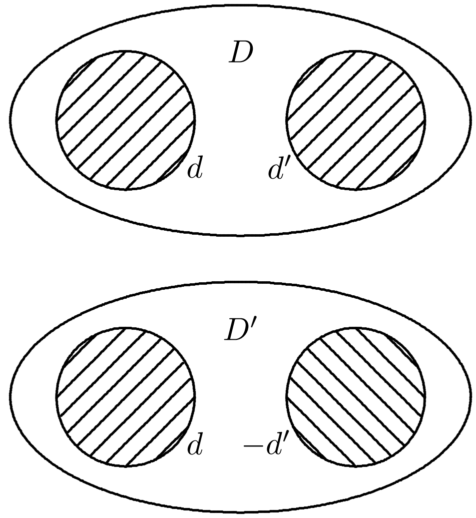
COROLLARY. *If M is a connected compact orientable manifold (without boundary), then the function $\int_M : \Omega^n M \to \mathbb{R}$ establishes an isomorphism $H^n_{\mathrm{DR}}(M) \cong \mathbb{R}$.*

6.5.2.3.2. The non-orientable case.

THEOREM. *Let M be a connected compact non-orientable manifold. Then $H^n_{\mathrm{DR}}(M) =$ 0 (equivalently: every closed form of degree n on M is exact).*

Proof. We define U_i, φ_i, V_i, and f_i precisely as in the proof in 6.5.2.3.1, only we, certainly, drop the condition that the atlas is oriented. Then we put $\alpha'_i = f_i\alpha$ and $c_i = \int_{V_i} \alpha'_i$ (no integrals over M!). Notice that the sum $\sum_i c_i$ is not necessarily zero. Next, we choose an *oriented* small ball d and a form γ with $\int_d \gamma = 1$ (it is important that the choice of d is arbitrary). W_i is defined as in 6.5.2.3.1 (see the picture above), then we put $\alpha_i = \alpha' i \pm c_i\gamma$ (the sign is plus, if the orientation of W_i compatible with the orientation of V_i is also compatible with the orientation of d and minus otherwise) and observe as above that α_i is a differential of a form with a compact support in W_i. Hence α_i is exact in M, as well as $\sum_i \alpha_i$. However, the last sum is not α: it is α plus a constant times γ. So, we need to prove that γ is also exact.





According to Theorem in 5.4.4.5, the manifold M can be obtained from a ball by attaching handles. Since the ball is orientable and M is not, there should be a handle attaching which turns an orientable manifold into a non-orientable manifold. This may happen only if a handle is attached to some N along a disconnected subset of a connected component of N. First, this shows that the handle must have index 1. Second, N must contain two small balls, d and d' which are contained in a bigger ball $D \subset N$ such that all D, d, and d' have coherent orientations, and also in a ball D' in N plus the handle such that D' has an orientation coherent with that of d and not coherent with that of d' (see the picture below). We do need N anymore, all we need is the balls $d, d', D, D' \subset M$ with the orientations as described above. We consider $\gamma, \gamma' \in \Omega^n M$ with supports in d, d' and $\int_d \gamma = \int_{d'} \gamma' = 1$ (with respect to the orientations of d and d'). Then we also have $\int_D (\gamma - \gamma') = 0$, $\int_{D'} (\gamma + \gamma') = 0$ (with respect to the orientations of D and D'). This implies that $\gamma - \gamma'$ is the differential of a form compactly supported in D, $\gamma + \gamma'$ is the differential of a form compactly supported in D', so both $\gamma - \gamma'$ and $\gamma + \gamma'$ are exact in M; hence γ (as well as γ') is exact in M.

6.5.2.3.3. The non-compact case.

THEOREM. *Let M be a connected non-compact n-dimensional manifold (orientable or not). Then $H^n_{\mathrm{DR}}(M) = 0$.*

I will not give a full proof of this result, but it is not hard, if one uses an elementary Lemma from the graph theory.

LEMMA. *Let Γ be an infinitely countable connected graph without multiple edges and loops and with all vertices having finite order. Orient all the edges. Then for every function $f\colon \mathrm{Vertices} \to \mathbb{R}$ there exists a function $g\colon \mathrm{Edges} \to \mathbb{R}$ such that for any vertex v,*

$$f(v) = \sum_{e, \mathrm{end}(e) = v} g(e) - \sum_{e, \mathrm{start}(e) = v} g(e)$$

(where e is the edge beginning at start(e) *and ending at* end(e).

Proof of Lemma: exercise.

Proof of Theorem. Suppose that M is oriented (it is not really important). Let $\alpha \in \Omega^n M$. Construct a compact locally finite cover $\{U_i \mid i = 1, 2, \ldots\}$ such that $U_i \approx \mathbb{R}^n$ and a partition of unity $\{f_i\}$ subordinated to this cover. Let $\alpha'_i = f_i \alpha$ and let $c_i = \int_{U_i} \alpha'_i = \int_M \alpha'_i$. Consider the graph with vertices v_i with v_i and $v_j, j \neq i$ joined by an edge, e_{ij}, if $U_i \cap U_j \neq \emptyset$. We orient e_{ij} from v_i to v_j, if $i < j$. Let $f(v_i) = c_i$, and let g be a function delivered by Lemma. Put $g(e_{ij}) = b_{ij}$. For every i, j with $j \neq i, U_i \cap U_j \neq \emptyset$, pick a small disc d_{ij}, $\overline{d}_{ij} \subset U_i \cap U_j$ and a form γ_{ij} with $\operatorname{supp} \gamma_{ij} \subset U_i \cap U_j$ and $\int_{d_{ij}} \gamma_{ij} = 1$. Then put $\alpha_i = \alpha'_i + \sum_{j<i} b_{ij}\gamma_{ij} - \sum_{j>i} b_{ij}\gamma_{ij}$. Obvious observations: (1) $\operatorname{supp} \alpha_i \subset U_i$; (2) $\int_M \alpha_i = \int_{U_i} \alpha_i = c_i + \sum_{j<i} b_{ij} - \sum_{j>i} b_{ij} = 0$; (3) $\sum_i \alpha_i = \sum_i \alpha'_i = \alpha$. The last equality is true, because every η_{ij} appears in $\sum_i \alpha_i$ twice: once with the coefficient b_{ij} and once with the coefficient $-b_{ij}$. Thus, the form α_i is the differential of some form β_i supported in U_i, and $\alpha = d\beta$, $\beta = \sum_i \beta_i$.

6.5.3. Further discussion of the de Rham cohomology. The de Rham theorem.

The idea of the computation of the de Rham cohomology of spheres in 6.5.2.1 can be applied to many other manifolds, like products of spheres, or $\mathbb{C}P^n$. (For the complex projective space $\mathbb{C}P^n$ one can use the standard covering by $n+1$ affine subspaces and prove that a closed form α of odd degree on $\mathbb{C}P^n$ is always exact, and the closed form α of even degree $2m$ is exact if and only if $\int_{\mathbb{C}P^m} \alpha = 0$.) But topology possesses some way more efficient methods for that. First of all, thee are many different constructions of cohomology in topology (singular cohomology, Čech cohomology, etc., etc.), and there are strong uniqueness theorems which assert that for sufficiently good spaces (certainly, including manifolds) all constructions yield the same results. The theorem stating that the de Rham cohomology of (arbitrary) manifold M is isomorphic to the singular cohomology (with coefficient in $\mathbb{R}$) is called the *de Rham theorem*. Its standard proof is based on the same geometric ideas as the computations of 6.5.2.1, but technically belongs to the sheaf theory.

Within the theory of manifolds, the expectable result would be a proposition that a closed form on a manifold is exact if and only if its integrals over compact orientable submanifolds without boundary are all zero. Unfortunately, in this nice form the theorem is false. It is true for forms of small degree (up to 6) or large degree (1 or 2 less than the dimension of the manifold); to make it true for all degrees, we have to consider integrals over manifolds with some kind of singulaities. Here we must restrict ourselves to the standard hypocritical excuse that all these fact are not within our reach is this course.

APPENDIX

Proof of Sard's Theorem

From Milnor's "Topology from the differential viewpoint."

First, let us recall the statement:

Theorem of Sard. *Let $f\colon U \to \mathbb{R}^n$ be a smooth map, with U open in $\mathbb{R}^m$, and let C be the set of critical points, that is, the set of all $x \in U$ with* $\operatorname{rank} d_x f < n$. *Then $f(C) \subset \mathbb{R}^n$ has measure zero.*

Proof: induction by m[1]. Note that the statement makes sense for $m \geq 0, n \geq 1$. To start the induction, the theorem is certainly true for $n = 0$.

Let $C_1 \subset C$ denote the set of all $x \in U$ such that $d_x f = 0$, that is, all first partial derivatives of all coordinate functions of f are zeroes. More generally, let C_i denote the set of all x such that all partial derivatives of all coordinate functions of f of order $\leq i$ vanish at x. Thus, we have a descending sequence of closed sets,

$$C \supset C_1 \supset C_2 \supset C_3 \supset \dots .$$

The proof will be divided into three steps as follows.

STEP 1. The image $f(C - C_1)$ has measure zero.

STEP 2. The image $f(C_i - C_{i+1})$ has measure zero for $i \geq 1$.

STEP 3. The image $f(C_k)$ has measure zero for k sufficiently large.

(If f is real analytic, then $\bigcap_i C_i = \emptyset$ unless f is constant on an entire component of U; hence, in this case, it is sufficient to carry out Steps 1 and 2.)

Proof of Step 1. This step is the hardest. We may assume $n \geq 2$, since if $n = 1$, then $C = C_1$. We will use the following well known Fubini's theorem: *a measurable set $A \subset \mathbb{R}^n = \mathbb{R} \times \mathbb{R}^{n-1}$ has measure zero, if $A \cap (t \times \mathbb{R}^{n-1})$ has measure zero for every $t \in \mathbb{R}$.*

For each $\overline{x} \in C - C_1$ we will find an open neighborhood $V \subset \mathbb{R}^m$ such that $f(V \cap C)$ has measure zero. Since $C - C_1$ is covered by countably many of these neighborhoods, this will prove that $f(C - C_1)$ has measure zero.

Since $\overline{x} \notin C_1$, there is some partial derivative, say $\dfrac{\partial f_1}{\partial x_1}$, which is not zero at $\overline{x}$. Consider the map $h: U \to \mathbb{R}^m$ defined by $h(x) = (f_1(x), x_2, \dots, x_m)$. Since $d_{\overline{x}} h$ is non-singular, h maps some neighborhood V of $\overline{x}$ diffeomorphically onto an open set V'. The composition $g = f \circ h^{-1}$ will then map V' into $\mathbb{R}^n$. Notice that (since h is a diffeomorphism), the set C' of critical points of g is precisely $h(V \cap C)$; hence the set $g(C')$ of critical values of g is equal to $f(V \cap C)$.

For each $(t, x_2, \dots, x_m) \in V'$, note that $g(t, x_2, \dots, x_m)$ belongs to the hyperplane $t \times \mathbb{R}^{n-1} \subset \mathbb{R}^n$ (if $h(x) = (t, x_2, \dots, x_m)$, then $f_1(x) = t$ and $g(t, x_2, \dots, x_m) = f(x) = (f_1(x), f_2(x), \dots, f_n(x)) = (t, f_2(x), \dots, f_p(x)))$; thus g carries hyperplanes into hyperplanes. Let

$$g^t : (t \times \mathbb{R}^{m-1}) \cap V' \to t \times \mathbb{R}^{m-1}$$

denote the restriction of g. Note that a point of $t \times \mathbb{R}^{m-1}$ is critical for g^t if and only if it is critical for g; for the Jacobian matrix of g has the form

$$\left[\frac{\partial g_i}{\partial x_j}\right] = \begin{bmatrix} 1 & 0 \\ * & \left[\dfrac{\partial g_i^t}{\partial x_j}\right] \end{bmatrix}.$$

---

[1]) Actually, the cases where $m \leq n$ are comparatively easy (see Section 3.3 of these notes); our proof, however, does not make much difference between these cases and the general case.

By the induction hypothesis, the set of critical values of g^t has measure zero in $t \times \mathbb{R}^{n-1}$. Therefore the set of critical values of g intersects each hyperplane $t \times \mathbb{R}^{n-1}$ in a set of measure zero. This set $g(C')$ is measurable, since it can be expressed as a countable union of compact subsets. Hence, by Fubini's theorem, the set $g(C') = f(V \cap C)$ has measure zero, and Step 1 is complete.

Proof of Step 2. For each $\overline{x} \in C_k - C_{k+1}$, there is some $(k+1)$-st deriative, $\dfrac{\partial^{k+1} f_r}{\partial x_{s_1} \ldots \partial x_{s_{k+1}}}$, which is not zero. Thus the function $w(x) = \dfrac{\partial^k f_r}{\partial x_{s_2} \ldots \partial x_{s_{k+1}}}$ vanishes at $\overline{x}$, but $\dfrac{\partial w}{\partial x_{s_1}}$ does not. Suppose, for definiteness, that $s_1 = 1$. Then the map $h\colon U \to \mathbb{R}^m$ defined by $h(x) = (w(x), x_2, \ldots, x_m)$ carries some neighborhood V of $\overline{x}$ diffeomorphically onto an open set V'. Note that h carries $C_k \cap V$ into the hyperplane $0 \times \mathbb{R}^{m-1}$. Again we consider $g = f \circ h^{-1}\colon V' \to \mathbb{R}^n$. Let $\overline{g}\colon (0 \times \mathbb{R}^{m-1}) \cap V' \to \mathbb{R}^n$ denote the restriction of g. By induction, the set of critical values of $\overline{g}$ has measure zero in $\mathbb{R}^n$. But each point in $h(C_k \cap V)$ is certainly a critical point of $\overline{g}$ (since all derivatives of order $\leq k$ vanish). Therefore $\overline{g}h(C_k \cap V) = f(C_k \cap V)$ has measure zero. Since $C_k - C_{k+1}$ is covered by finitely many such sets V, it follows that $f(C_k - C_{k+1})$ has measure zero.

Proof of Step 3. Let $I^m \subset U$ be a cube with edge δ. If k is sufficiently large ($k > \dfrac{n}{p} - 1$, to be precise) we will prove that $f(C_k \cap I^m)$ has measure zero.

From Taylor's theorem, the compactness of I^n, and the definition of C_k, we see that $f(x+h) = f(x) + R(x,h)$ where

$$\|R(x,h)\| \leq \|h\|^{k+1} \tag{1}$$

for $x \in C_k \cap I^n, x + h \in I^n$. Here c is a constant which depends only on f and I^n. Now subdivide I^n into r^n cubes of edge $\dfrac{\delta}{r}$ Let I_1 be a cube of the subdivision that contains a point x of C_k. Then any point of I_1 can be written as $x + h$ with

$$\|h\| \leq \sqrt{n} \cdot \frac{\delta}{r}. \tag{2}$$

From (1) it follows that $f(I_1)$ lies in a cube of edge $\dfrac{a}{r^{k+1}}$ centered about $f(x)$ where $a = 2c(\sqrt{n}\delta)^{k+1}$ is constant. Hence, $f(C_k \cap I^m)$ is contained in a union of at most r^m cubes having total volume

$$\mathrm{Vol} \leq r^m \left(\frac{a}{r^{k+1}}\right)^p = a^p r^{n-(k+1)p}.$$

If $k + 1 > \dfrac{n}{p}$, then evidently Vol tends to 0 as $r \to \infty$; so $f(C_k \cap I^n)$ must have measure zero. This completes proof of Sard's theorem.