UNIVERZA NA PRIMORSKEM
FAKULTETA ZA MATEMATIKO, NARAVOSLOVJE IN INFORMACIJSKE TEHNOLOGIJE

Zaključna naloga
(Final project paper)

Sardov izrek, Morsejeve funkcije in Whitneyjev vložitveni izrek za diferenciabilne mnogoterosti

(Sard's theorem, Morse functions, and Whitney embedding theorem for differentiable manifolds)

Ime in priimek: Anastasiya Tanana
Študijski program: Matematika
Mentor: prof. dr. Bojan Kuzma

Koper, junij 2015

Ključna dokumentacijska informacija

Ime in PRIIMEK: Anastasiya TANANA

Naslov zaključne naloge: Sardov izrek, Morsejeve funkcije in Whitneyjev vložitveni izrek za diferenciabilne mnogoterosti

Kraj: Koper

Leto: 2015

Število listov: 30

Število referenc: 5

Mentor: prof. dr. Bojan Kuzma

Ključne besede: diferenciabilna mnogoterost, imerzija, vložitev, kritične vrednosti.

Math. Subj. Class. (2010): 57R40, 57R42, 58A05.

Izvleček:
V prvih dveh poglavjih, Uvod in Preslikave med mnogoterostmi, so predstavljene nekatere od osnovnih pojmov diferencialne topologije. Glavni del zaključne naloge predstavlja dokaz Sardovega izreka v poglavju 3. S pomočjo Sardovega izreka nato pokažemo, da na vsaki diferenciabilni mnogoterosti obstajajo Morsejeve funkcije. V poglavju 5 dokazujemo verzijo Whitneyjevega vložitvenega izreka, in sicer, da ima vsaka n-dimenzionalna diferenciabilna mnogoterost vložitev v $\mathbb{R}^{2n+1}$.

Key words documentation

Name and SURNAME: Anastasiya TANANA

Title of final project paper: Sard's theorem, Morse functions, and Whitney embedding theorem for differentiable manifolds

Place: Koper

Year: 2015

Number of pages: 30

Number of references: 5

Mentor: Prof. Bojan Kuzma, PhD

Keywords: differentiable manifold, immersion, embedding, critical values.

Math. Subj. Class. (2010): 57R40, 57R42, 58A05.

Abstract: In the first two chapters, Introduction and Mappings between manifolds, we introduce some of the basic notions of differential topology. The main part of the project paper is the proof of the Sard's theorem in Chapter 3. With the help of Sard's theorem we show that on any differentiable manifold there exist Morse functions. In Chapter 5 we prove a version of Whitney embedding theorem, namely, that any n-dimensional differentiable manifold embeds in $\mathbb{R}^{2n+1}$.

Acknowledgement

Thanks to my mentor Prof. Bojan Kuzma.

Contents

1 Introduction **1**

2 Mappings between manifolds **6**

3 Sard's theorem **11**

4 Morse functions **18**

5 Whitney embedding theorem **22**

6 Conclusion **27**

7 Povzetek naloge v slovenskem jeziku **28**

8 Bibliography **30**

1 Introduction

The primary source for writing this project paper was the first chapter of the book *"Differential topology"* by V. Guillemin and A. Pollack, [1].

In the introductory chapter we will define the basic notions that will be considered later. We will also state without proof and generalize some results from Analysis that we will use later.

By $A \subset B$ we denote that A is a subset of B, that is, either A is a proper subset of B, or $A = B$.

Definition 1.1. A continuous function f from an open subset X of $\mathbb{R}^n$ to $\mathbb{R}^m$ is called **smooth** if all its partial derivatives of all orders are continuous. A function f from arbitrary $X \subset \mathbb{R}^n$ to $\mathbb{R}^m$ is called smooth if for all $x \in X$ there exist an open neighbourhood $U \in \mathbb{R}^n$ of x and a smooth function $F : U \to \mathbb{R}^m$ such that $f = F$ on $X \cap U$.

Definition 1.2. A map f from $X \subset \mathbb{R}^n$ to $Y \subset \mathbb{R}^m$ is called **diffeomorphism** if f is smooth, bijective, and its inverse is smooth. Subsets of Euclidean spaces X and Y are called **diffeomorphic** if such a map exists.

Now we define a differentiable manifold.

Definition 1.3. Let X be a subset of $\mathbb{R}^N$. Then X is called an n-**dimensional differentiable manifold** if it is locally diffeomorphic to $\mathbb{R}^n$, i.e., if for each $x \in X$ there exists an open[1] neighbourhood U of x in X which is diffeomorphic to an open set V in $\mathbb{R}^n$. The corresponding diffeomorphism from V to U is called a **parametrization** of the neighbourhood U. Its inverse is called a **coordinate system** on U.

In this definition we took the manifold initially to be a subset of sufficiently large Euclidean space. This is a simplification of the abstractly defined differentiable manifold X which is by definition a Hausdorff topological space such that each point $x \in X$ has a neighbourhood $U_x \subset X$ homeomorphic to an open subset of $\mathbb{R}^n$; and if $x, y \in X$, $\varphi_x : U_x \to \mathbb{R}^n, \varphi_y : U_y \to \mathbb{R}^n$ are homeomorphisms from open neighbourhoods U_x, U_y of x and y to some open subsets of $\mathbb{R}^n$ and $U_x \cap U_y \neq \emptyset$, then the transition map

$$\varphi = \varphi_y \circ \varphi_x^{-1}|_{\varphi_x(U_x \cap U_y)} : \varphi_x(U_x \cap U_y) \to \varphi_y(U_x \cap U_y)$$

[1] Topology on X is induced by the usual topology on $\mathbb{R}^N$.

is smooth. But it can be proved that any differentiable manifold (in its general definition) has a homeomorphic copy (can be embedded) in some big enough Euclidean space. We will not prove this fact and we will use Definition 1.3 throughout the paper. Since we will only deal with differentiable manifolds, the word "differentiable" will usually be omitted.

If X is a manifold and $Z \subset X$ is a manifold, then Z is called a *submanifold* of X. To denote that X is a manifold of dimension n we write dim $X = n$.

Example 1.4. Within Analysis IV we defined a surface in $\mathbb{R}^3$ as a subset $M \subset \mathbb{R}^3$ such that for each $p \in M$ there exist an open neighbourhood $U_p \subset \mathbb{R}^3$ of p, an open neighbourhood $V_p \subset \mathbb{R}^3$ of 0 and a smooth bijection $F_p : U_p \to V_p$ with smooth inverse, such that

$$F_p(U_p \cap M) = V_p \cap (\mathbb{R}^2 \times \{0\}).$$

This is actually equivalent to the definition of a 2-dimensional manifold embedded in $\mathbb{R}^3$, since the composition of $F_p|_{U_p \cap M}$ with the projection map $\pi : (x, y, 0) \mapsto (x, y)$ is precisely a local coordinate system on the open neighbourhood $U_p \cap M$ of p in M.

Proposition 1.5. *If $X \subset \mathbb{R}^N$ is a manifold of dimension n and $Y \subset \mathbb{R}^M$ is a manifold of dimension m, then $X \times Y \subset \mathbb{R}^{N+M}$ is a manifold of dimension $n + m$.*

We will define tangent space of a manifold at a given point, derivative of a smooth mapping between manifolds and state the chain rule for the mappings between manifolds. So, let us revise the definition of derivative and recall the chain rule for usual smooth mappings of open subsets of Euclidean spaces.

Definition 1.6. Let f be a smooth mapping from an open subset U of $\mathbb{R}^n$ to $\mathbb{R}^m$. Then its *derivative* at point $x \in U$ is a linear transformation $df_x : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{h \to 0} \frac{||f(x + h) - f(x) - df_x(h)||}{||h||} = 0.$$

We know from Analysis that the above defined linear transformation is unique (see [4] for the proof), and if we write $f = (f_1, ..., f_m),\ x = (x_1, ..., x_n)$, df_x is represented by its Jacobian matrix

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & ... & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & ... & \frac{\partial f_2}{\partial x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \frac{\partial f_m}{\partial x_2}(x) & ... & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}.$$

Theorem 1.7 (Chain rule)**.** *Let $f : U \to V,\ g : V \to \mathbb{R}^m$ be smooth mappings. Then*

$$d(g \circ f)_x = dg_{f(x)} \circ df_x.$$

The proof of the chain rule can be found in [4].

Now we return to the manifolds and define a *tangent space*. A tangent space of manifold X at point x is a linear subspace of $\mathbb{R}^N$ such that lifted by x it is the best flat approximation of a manifold at point x. And we identify it with the help of the derivative of a local parametrization, which is by the definition of derivative the best linear approximation of a local parametrization.

Definition 1.8. Let $X \subset \mathbb{R}^N$ be an n-dimensional manifold and let $x \in X$. Let $\phi : U \to X$, $0 \in U \subset \mathbb{R}^n$, be a local parametrization around x such that $\phi(0) = x$. A **tangent space** of manifold X at point x is the image of the mapping $d\phi_0 : \mathbb{R}^n \to \mathbb{R}^N$.

We will denote the tangent space of manifold X at point x by $T_x(X)$.

We have to verify that the definition does not depend on the choice of a local parametrization. Take another parametrization $\varphi : V \to X, 0 \in V \subset \mathbb{R}^n$ with $\varphi(0) = x$. Taking the restrictions of the maps, we can assume that $\phi(U) = \varphi(V)$. Then $h = \varphi^{-1} \circ \phi$ is a diffeomorphism from U to V. Since $\phi = \varphi \circ h$, by the chain rule $d\phi_0 = d\varphi_0 \circ dh_0$. Therefore, $\mathrm{Im}(d\phi_0) \subset \mathrm{Im}(d\varphi_0)$. The converse follows by interchanging the roles of ϕ and φ.

Let us show that dim $T_x(X) = n$. The inverse ϕ^{-1} is a smooth mapping from an open subset of X to U. Hence, it has a smooth extension Φ' to an open subset of $\mathbb{R}^N$ containing x. Then $\Phi' \circ \phi$ is the identical map of U onto itself, which means that $d\Phi'_x \circ d\phi_0$ is the identical map from $\mathbb{R}^n$ to $\mathbb{R}^n$. Recall that $d\phi_0$ is a linear map from $\mathbb{R}^n$ to $\mathbb{R}^N$ and $d\Phi'_x$ is a linear map from $\mathbb{R}^N$ to $\mathbb{R}^n$. This implies that dim $T_x(X) = \dim \mathrm{Im}(d\phi_0) = n$. We have shown that $d\phi_0$ is an isomorphism of $\mathbb{R}^n$ and $T_x(X)$.

Example 1.9. An open subset U of $\mathbb{R}^n$ is an n-dimensional manifold since the identity map on it is its local parametrization at each point $x \in U$. The tangent space $T_x(U)$ of U at x is an n-dimensional linear subspace of $\mathbb{R}^n$, hence $T_x(U) = \mathbb{R}^n$.

Now we will generalize the definition of derivative to a smooth mapping between arbitrary manifolds.

Definition 1.10. Let $f : X \to Y$ be a smooth mapping between manifolds and let $x \in X$ with $f(x) = y$. Let $\phi : U \to X$ and $\varphi : V \to Y$ be local parametrizations around x and y, where $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ are open and $\phi(0) = x$, $\varphi(0) = y$. Then $h = \varphi^{-1} \circ f \circ \phi$ is a smooth map from U to V. We know that $d\phi_0$, dh_0 and $d\varphi_0$ exist, and since $f = \varphi \circ h \circ \phi^{-1}$, we define $df_x : T_x(X) \to T_y(Y)$, the **derivative** of f at x, as follows:

$$df_x = d\varphi_0 \circ dh_0 \circ (d\phi_0)^{-1}.$$

Here, h is a smooth map from U to V, because ϕ is smooth, f can be extended to a smooth map on an open subset of $\mathbb{R}^N$ containing x and φ^{-1} can be extended to a smooth map on an open subset of $\mathbb{R}^M$ containing y. We have to check that the definition does not depend on the choice of local parametrizations ϕ and φ. Let $\phi_1 : U_1 \to X$ and $\varphi_1 : V_1 \to Y$ be other local parametrizations around x and y, where $U_1 \subset \mathbb{R}^n$, $V_1 \subset \mathbb{R}^m$ are open and $\phi_1(0) = x$, $\varphi_1(0) = y$. The map $h_1 = \varphi_1^{-1} \circ f \circ \phi_1$ is smooth from U_1 to V_1. Taking the restrictions of $\phi, \phi_1, \varphi, \varphi_1$ we can assume that $\phi(U) = \phi_1(U_1)$ and $\varphi(V) = \varphi_1(U_1)$. Since $f = \varphi \circ h \circ \phi^{-1} = \varphi_1 \circ h_1 \circ \phi_1^{-1}$, we have $h = (\varphi^{-1} \circ \varphi_1) \circ h_1 \circ (\phi_1^{-1} \circ \phi)$. By the usual chain rule, $dh_0 = d(\varphi^{-1} \circ \varphi_1)_0 \circ d(h_1)_0 \circ d(\phi_1^{-1} \circ \phi)_0$. Now, $\varphi^{-1} \circ \varphi_1$ is a mapping from U_1 to U and φ^{-1} has a smooth extension Φ' to an open subset of $\mathbb{R}^N$ containing x. Therefore, we have the equality $\varphi^{-1} \circ \varphi_1 = \Phi' \circ \varphi_1$, and by the usual chain rule $d(\varphi^{-1} \circ \varphi_1)_0 = d(\Phi' \circ \varphi_1)_0 = d\Phi'_x \circ d(\varphi_1)_0$. The derivative $d(\varphi_1)_0$ maps U_1 to $T_x(X)$, and we have seen already that $d\Phi'_x$ restricted to $T_x(X)$ is precisely $(d\varphi_0)^{-1}$. Therefore, $d(\varphi^{-1} \circ \varphi_1)_0 = d\Phi'_x \circ d(\varphi_1)_0 = (d\varphi_0)^{-1} \circ d(\varphi_1)_0$. Similarly, $d(\phi_1^{-1} \circ \phi)_0 = (d(\phi_1)_0)^{-1} \circ d\phi_0$. Hence,

$$dh_0 = d(\varphi^{-1} \circ \varphi_1)_0 \circ d(h_1)_0 \circ d(\phi_1^{-1} \circ \phi)_0 = (d\varphi_0)^{-1} \circ d(\varphi_1)_0 \circ d(h_1)_0 \circ (d(\phi_1)_0)^{-1} \circ d\phi_0$$

which implies that

$$d(\varphi_1)_0 \circ d(h_1)_0 \circ (d(\phi_1)_0)^{-1} = d\varphi_0 \circ dh_0 \circ (d\phi_0)^{-1}.$$

This proves the independence from the choice of local parametrizations.

The generalized definition matches with the usual one for mappings of open subsets of Euclidean spaces.

The chain rule also holds:

Theorem 1.11. *Let $f : X \to Y$, $g : Y \to Z$ be smooth mappings between manifolds. Then*

$$d(g \circ f)_x = dg_{f(x)} \circ df_x.$$

Proof. Let $\phi : U \to X$, $\varphi : V \to Y$ and $\psi : W \to Z$ be local parametrizations around x, $y = f(x)$ and $z = g(f(x))$ respectively, where $\phi(0) = x$, $\varphi(0) = y$, $\psi(0) = z$. Then $h = \varphi^{-1} \circ f \circ \phi$ is a smooth map from U to V and $k = \psi^{-1} \circ g \circ \varphi$ is a smooth map from V to W. By definition, $df_x = d\varphi_0 \circ dh_0 \circ d\phi_0^{-1}$ and $dg_{f(x)} = d\psi_0 \circ dk_0 \circ d\varphi_0^{-1}$. By the chain rule,

$$dg_{f(x)} \circ df_x = (d\psi_0 \circ dk_0 \circ d\varphi_0^{-1}) \circ (d\varphi_0 \circ dh_0 \circ d\phi_0^{-1}) =$$

$$= d\psi_0 \circ dk_0 \circ dh_0 \circ d\phi_0^{-1} = d\psi_0 \circ d(k \circ h)_0 \circ d\phi_0^{-1}.$$

But since $k \circ h = \psi^{-1} \circ (g \circ f) \circ \phi$, the latter expression is precisely $d(g \circ f)_x$. □

We will also need the inverse function theorem:

Theorem 1.12 (Inverse function theorem)**.** *Let $f : U \to V$ be a smooth mapping, where U and V are open subsets of $\mathbb{R}^n$. If the derivative df_x of f at point x is an isomorphism, then f is a local diffeomorphism at x, i.e. f has a smooth inverse in some open neighbourhood of $f(x)$, or, in other words, f is a diffeomorphism of some open neighbourhood of x in U and some open neighbourhood of $f(x)$ in V.*

The proof can be found in [4]. Using local parametrizations as above, the theorem can be generalized to mappings of arbitrary manifolds:

Theorem 1.13. *Let $f : X \to Y$ be a smooth mapping between manifolds whose derivative df_x at point x is an isomorphism. Then f is a local diffeomorphism at x.*

Proof. Let $\phi : U \to X$, $\varphi : V \to Y$ be local parametrizations around x and y, where $\phi(0) = x, \varphi(0) = y$. Let $h = \varphi^{-1} \circ f \circ \phi$. The mapping φ^{-1} extends to a smooth map Φ' on an open set containing y. By the chain rule, $dh_0 = d\Phi'_y \circ df_x \circ d\phi_0$. Since $dh_0 : \mathbb{R}^n \to \mathbb{R}^n$ is a composition of three isomorphisms, $d\phi_0 : \mathbb{R}^n \to T_x(X)$, $df_x : T_x(X) \to T_y(Y)$ and $d\Phi'_y|_{T_y(Y)} : T_y(Y) \to \mathbb{R}^n$ which, as we have seen above, equals $(d\varphi_0)^{-1}$, it is an isomorphism itself. Hence, by the usual inverse function theorem $h : U \to V$ is a diffeomorphism at 0. Since $f = \varphi \circ h \circ \phi^{-1}$ is the composition of three diffeomorphisms, it is a diffeomorphism at x. □

In the next chapter we will define and consider in more details mappings between manifolds called immersions and submersions.

2 Mappings between manifolds

The statement of the inverse function theorem can be applied only if $\dim X = \dim Y$. If $\dim X < \dim Y$, the best possible behaviour of df_x is that it is injective.

Definition 2.1. If $f : X \to Y$ is a smooth mapping between manifolds and $df_x : T_x(X) \to T_y(Y)$ is injective then f is called an **immersion** at x. Mapping f is called an immersion, if it is an immersion at each point of X. The mapping $\mathfrak{i} : (x_1, ..., x_n) \mapsto (x_1, ..., x_n, 0, ..., 0)$ from $\mathbb{R}^n$ to $\mathbb{R}^m$, $n \leq m$, is called a **canonical immersion**.

The following statement shows that locally all immersions are equivalent to the canonical immersion:

Theorem 2.2 (Local immersion theorem). *Let $f : X \to Y$ be an immersion at x and $f(x) = y$. Then there exist local coordinates around x and y such that*

$$f(x_1, ..., x_n) = (x_1, ..., x_n, 0, ..., 0).$$

Proof. Let $\phi : U \to X$ and $\varphi : V \to Y$ be arbitrary local parametrizations around x and y, $\phi(0) = x$ and $\varphi(0) = y$. Since f is continuous, $f^{-1}(\varphi(V))$ is open in X. Hence shrinking U to $\phi^{-1}(f^{-1}(\varphi(V)) \cap \phi(U))$, which is open as the preimage of the open set $f^{-1}(\varphi(V)) \cap \phi(U)$ under the smooth map ϕ, we can assume that $(f \circ \phi)(U) \subset \varphi(V)$. Then $h = \varphi^{-1} \circ f \circ \phi$ is a mapping from $U \subset \mathbb{R}^n$ to $V \subset \mathbb{R}^m$ whose derivative at 0 is injective, because $dh_0 = d(\varphi^{-1})_y \circ df_x \circ d\phi_0$, mappings $d(\varphi^{-1})_y : T_y(Y) \to \mathbb{R}^m$ and $d\phi_0 : \mathbb{R}^n \to T_x(X)$ are isomorphisms and $df_x : T_x(X) \to T_y(Y)$ is injective.

The linear map $dh_0 : \mathbb{R}^n \to \mathbb{R}^m$ is injective and therefore maps the standard basis $\{e_1, ..., e_n\}$ of $\mathbb{R}^n$ to n linearly independent vectors $\{v_1, ..., v_n\}$ of $\mathbb{R}^m$. These n vectors can be completed to a basis $\{v_1, ..., v_n, v_{n+1}, ..., v_m\}$ of $\mathbb{R}^m$. The matrix $P' = (v_1|v_2|...|v_n)$ maps the standard basis $\{e'_1, ..., e'_m\}$ of $\mathbb{R}^m$ to the basis $\{v_1, ..., v_m\}$. Hence, if we define $P = (P')^{-1} : \mathbb{R}^m \to \mathbb{R}^m$, the composition $P \circ dh_0$ maps $e_i \in \mathbb{R}^n$ to $e'_i \in \mathbb{R}^m$ for each $1 \leq i \leq n$, which means that

$$P \circ dh_0 = \begin{pmatrix} I_n \\ \hline 0 \end{pmatrix},$$

where I_n is $n \times n$ identity matrix and 0 is $(m - n) \times n$ zero matrix.

Now, define a mapping $H : U \times \mathbb{R}^{m-n} \to \mathbb{R}^m$ by

$$H(x, z) = (P \circ h)(x) + (0, z).$$

It holds that $dH_0 = I_m$, and by the inverse function theorem H is a local diffeomorphism around $0 \in \mathbb{R}^m$. Therefore, the composition $\varphi \circ H$ is a local parametrization of Y around y. By the definition of H, $P \circ h = H \circ \mathfrak{i}$, where $\mathfrak{i}$ is the canonical immersion. Hence, at a neighbourhood of x we have:

$$f = \varphi \circ h \circ \phi^{-1} = (\varphi \circ P^{-1} \circ H) \circ \mathfrak{i} \circ \phi^{-1},$$

which means that ϕ and $\varphi \circ P^{-1} \circ H$ define the sought-for local coordinates around x and y. □

A corollary of this theorem is that if f is an immersion at x, then f is an immersion in a neighbourhood of x.

Although an immersion $f : X \to Y$ maps a neighbourhood of $x \in X$ diffeomorphically to a submanifold of Y, globally $f(X)$ does not have to be a submanifold of Y. For instance, it can be shown that Klein bottle can be immersed in $\mathbb{R}^3$ but only with self-intersections, so it is not a submanifold of $\mathbb{R}^3$.

On the other hand, even if the image $f(X)$ of immersion f is a manifold, it doesn't have to be diffeomorphic to X: an example is the mapping $f : \mathbb{R} \to S^1$, $f : t \mapsto (\cos t, \sin t)$. It is an immersion because $df_t = (-\sin t, \cos t) \neq (0,0)$ for $t \in \mathbb{R}$, but $\mathbb{R}$ and S^1 are not diffeomorphic because they are not even homeomorphic.

Notice that when $\dim X = \dim Y$, immersions are precisely local diffeomorphisms.

Therefore, being an immersion is a local property. But it suffices to require immersion to be injective and *proper* to make it a topological embedding of X into Y.

Definition 2.3. A mapping $f : X \to Y$ is called proper if the preimage of every compact set in Y is compact in X.

Definition 2.4. An immersion which is injective and proper is called an **embedding**.

The following theorem shows that the above definition indeed gives a topological embedding:

Theorem 2.5. *An embedding $f : X \to Y$ maps X diffeomorphically onto a submanifold of Y.*

Proof. First, let us show that the image of an open set $W \subset X$ is open in $f(X)$. Assume that $f(W)$ is not open. Then $f(X) \setminus f(W)$ is not closed and there exists a sequence $y_1, y_2, \ldots \in f(X) \setminus f(W)$ converging to $y \in f(W)$. Since the set $\{y_i;\ i \in \mathbb{N}\} \cup \{y\}$ is compact, its preimage $\{x_i;\ i \in \mathbb{N}\} \cup \{x\}$ $(f(x_i) = y_i, f(x) = y)$ is compact as well. Hence, the sequence $\{x_i\}_{i\in\mathbb{N}}$ has a limit point $z \in X$. Then $f(z)$ is a limit point of $\{f(x_i)\}_{i\in\mathbb{N}}$, which implies that $f(z) = y = f(x)$ and by injectivity $z = x$. But $x \in W^{open}$, while $x_i \notin W$, a contradiction. Therefore, $f(W)$ is open in $f(X)$.

By the local immersion theorem, there exist local parametrizations $\phi : U \to X$ and $\varphi : V \to Y$ around x and $f(x)$, where $\phi(0) = x$ and $\varphi(0) = f(x)$, such that the map $h = \varphi^{-1} \circ f \circ \phi : U \to V$ is the canonical immersion, i.e.

$$h(x_1, ..., x_n) = (x_1, ..., x_n, 0, ..., 0)$$

for $(x_1, ..., x_n) \in U$. Since f maps open sets to open sets, by shrinking U and V we can achieve that $f(\phi(U)) = \varphi(V)$: namely, take instead of U the set $\phi^{-1}(\phi(U) \cap f^{-1}(\varphi(V)))$ and instead of V the set $\varphi^{-1}(f(\phi(U) \cap f^{-1}(\varphi(V))))$. Since $h(U) = V \cap (\mathbb{R}^n \times \{0\})$ and $h = \varphi^{-1} \circ f \circ \phi$, we have $f(\phi(U)) = \varphi(V \cap (\mathbb{R}^n \times \{0\}))$, i.e. the image of the open neighbourhood $\phi(U)$ of x in X under f is the subset $\varphi(V \cap (\mathbb{R}^n \times \{0\})) = \varphi(V) \cap f(X)$ of $f(X)$ open and containing $f(x)$. But $\varphi(V \cap (\mathbb{R}^n \times \{0\}))$ is diffeomorphic to $V \cap (\mathbb{R}^n \times \{0\})$ which is diffeomorphic to an open subset of $\mathbb{R}^n$ because V is open. Hence, for each $x \in X$ there exists an open parametrizable neighbourhood of $f(x)$ in $f(X)$, which means that $f(X)$ is a manifold.

Since mapping $f : X \to f(X)$ is bijective, the inverse f^{-1} is well-defined. Since f is a diffeomorphism locally, f^{-1} is smooth at each point of $f(X)$. Since f is smooth and its inverse is well-defined and smooth, f is a diffeomorphism of X and $f(X)$. □

When X is a compact manifold, all smooth mappings $f : X \to Y$ are proper: every compact subset of Y is closed, and its preimage is closed in X because f is continuous, but since X is compact, closed subset of X are also compact. Therefore, whenever X is compact, embeddings are precisely injective immersions.

We defined n-dimensional differentiable manifold as a subset of sufficiently large Euclidean space $\mathbb{R}^N$. Later we will prove that every n-dimensional manifold can be embedded in $\mathbb{R}^{2n+1}$.

The notion dual to immersion for the case dim $X >$ dim Y is a submersion:

Definition 2.6. If $f : X \to Y$ is a smooth mapping of manifolds and $df_x : T_x(X) \to T_y(Y)$ is surjective then f is called a **submersion** at x. Mapping f is called a submersion, if it is a submersion at each point of X. The mapping $\mathfrak{s} : (x_1, ..., x_m, x_{m+1}, ..., x_n) \mapsto (x_1, ..., x_m)$ from $\mathbb{R}^n$ to $\mathbb{R}^m$, $n \geq m$, is called a **canonical submersion**.

By definition, if dim $X =$ dim Y then submersions are precisely local diffeomorphisms.

The statement analogous to the local immersion theorem also holds:

Theorem 2.7 (Local submersion theorem)**.** *Let $f : X \to Y$ be a submersion at x and $f(x) = y$. Then there exist local coordinates around x and y such that*

$$f(x_1, ..., x_n) = (x_1, ..., x_m).$$

Proof. The proof is similar to the proof of the local immersion theorem. Let $\phi : U \to X$ and $\varphi : V \to Y$ be arbitrary local parametrizations around x and y, $\phi(0) = x$ and $\varphi(0) = y$. The mapping $h = \varphi^{-1} \circ f \circ \phi$ from $U \subset \mathbb{R}^n$ to $V \subset \mathbb{R}^m$ has surjective derivative at 0. Hence, there exists a linear isomorphism $P : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$dh_0 \circ P = \left(I_m \middle| \; 0 \right),$$

where I_m is $m \times m$ identity matrix and 0 is $m \times (n-m)$ zero matrix. Define a mapping $H : P^{-1}(U) \to \mathbb{R}^n$ by

$$H(x) = ((h \circ P)(x), x_{m+1}, ..., x_n).$$

Since $dH_0 = I_n$, H is a local diffeomorphism around $0 \in \mathbb{R}^n$. So, H^{-1} is a local diffeomorphism around $0 \in \mathbb{R}^n$. By the definition of H, $h \circ P = \mathfrak{s} \circ H$, where $\mathfrak{s}$ is the canonical submersion. We obtain:

$$f = \varphi \circ h \circ \phi^{-1} = \varphi \circ \mathfrak{s} \circ (H \circ P^{-1} \circ \phi^{-1}),$$

therefore, $\phi \circ P \circ H^{-1}$ and φ define the required local coordinates around x and y. □

Similarly as for immersions, we see that if f is a submersion at x, then it is a submersion in a neighbourhood of x.

Now we define regular and critical points and values of f.

Definition 2.8. Let $f : X \to Y$ be a smooth mapping of manifolds. A point $x \in X$ is called a **regular point** of f if $df_x : T_x(X) \to T_y(Y)$ is surjective. If df_x is not surjective, x is called a **critical point** of f. A point $y \in Y$ is called a **regular value** of f if all points in $f^{-1}(y)$ are regular points of f. If y is not a regular value of f, it is called a **critical value**.

Observe that all points in $Y \setminus f(X)$ are regular values by definition.

If dim $X <$ dim Y then all points in X are critical points and all points in $f(X)$ are critical values.

If dim $X =$ dim Y then $x \in X$ is a regular point if and only if f is a local diffeomorphism at x. Point $y \in Y$ is a regular value if and only if f is a local diffeomorphism at each point of the preimage $f^{-1}(y)$.

If dim $X >$ dim Y then $x \in X$ is a regular point if and only if f is a submersion at x. Point $y \in Y$ is a regular value if and only if f is a submersion at each point of the preimage $f^{-1}(y)$.

In the next chapter we will prove the Sard's theorem which states that the set of critical values of a smooth mapping of manifolds has Lebesgue measure zero.

Let us prove now the following result:

Theorem 2.9 (Preimage theorem)**.** *If y is a regular value of $f : X \to Y$, then the preimage $f^{-1}(y)$ is a submanifold of X of dimension dim X − dim Y.*

Proof. Let $x \in f^{-1}(y)$. By the local submersion theorem, there exist local coordinate systems around x and y such that $x = (0, ..., 0) \in \mathbb{R}^n$, $y = (0, ..., 0) \in \mathbb{R}^m$ and

$$f(x_1, ..., x_n) = (x_1, ..., x_m).$$

Let W be an open neighbourhood of x where the above coordinates are defined. Then $W \cap f^{-1}(y)$ is precisely the set of points that have first m coordinates equal to 0: $(0, ..., 0, x_{m+1}, ..., x_n)$. Therefore, $(x_{m+1}, ..., x_n)$ define a coordinate system on the open neighbourhood $W \cap f^{-1}(y)$ of x in $f^{-1}(y)$, and hence $f^{-1}(y)$ is an $(n-m)$-dimensional manifold. □

Example 2.10. We can now easily prove that the unit sphere S^{n-1} is an $(n-1)$-dimensional manifold. Consider the mapping $f : \mathbb{R}^n \to \mathbb{R}$ defined by

$$f(x_1, x_2, ..., x_n) = x_1^2 + x_2^2 + ... + x_n^2.$$

Its derivative $df_x = (2x_1, 2x_2, ..., 2x_n)$ is surjective everywhere except at 0. Therefore, 1 is a regular value, and its preimage, the unit sphere S^{n-1}, is an $(n-1)$-dimensional manifold.

Proposition 2.11. *If Z is a submanifold of X then $T_x(Z) \subset T_x(X)$ for $x \in Z$.*

Proof. Let $\phi : U \to X$ and $\varphi : V \to Z$ be local parametrizations around $x \in Z \subset X$. Shrinking U and V we can assume that $\phi(U) \cap Z = \varphi(V)$. Define $h = \phi^{-1} \circ \varphi : V \to U$. This is a smooth map and $h(0) = 0$. We have $\varphi = \phi \circ h$ and by the chain rule: $T_x(Z) = \mathrm{Im}(d\varphi_0) = \mathrm{Im}(d\phi_0 \circ dh_0) = d\phi_0(\mathrm{Im}(dh_0)) \subset d\phi_0(\mathbb{R}^n) = T_x(X)$. □

When Z is a preimage of a regular value, $T_x(Z)$ is precisely the kernel of df_x:

Proposition 2.12. *Let $f : X \to Y$ be a smooth mapping of manifolds, y its regular value and $Z = f^{-1}(y)$. Then the kernel of the derivative $df_x : T_x(X) \to T_y(Y)$ for $x \in Z$ is precisely $T_x(Z)$.*

Proof. Map f is constant on Z, which implies that $df_x(T_x(Z)) = 0$, i.e. $T_x(Z) \subset \ker(df_x)$. By Theorem 2.9, dim $Z = n - m$, and hence dim $T_x(Z) = n - m$. On the other hand, since $df_x : T_x(X) \to T_y(Y)$ is surjective, dim (ker df_x) = dim $T_x(X)$ − dim $T_y(Y) = n - m$. Since the dimensions are the same, $T_x(Z)$ coincides with the kernel. □

3 Sard's theorem

To formulate the Sard's theorem we will need the definition of a set of measure zero.

Definition 3.1. Let $(a_1, a_2, ..., a_n) \in \mathbb{R}^n$ and $h > 0$. We call a subset

$$(a_1, a_1 + h) \times (a_2, a_2 + h) \times ...(a_n, a_n + h)$$

of $\mathbb{R}^n$ an *open cube* and define its *volume* as h^n. A subset

$$[a_1, a_1 + h] \times [a_2, a_2 + h] \times ...[a_n, a_n + h]$$

of $\mathbb{R}^n$ is called a *closed cube* and its *volume* is defined as h^n.

Definition 3.2. A set $A \subset \mathbb{R}^n$ is said to be **of measure zero** if for every $\varepsilon > 0$ it can be covered by a countable number of cubes whose total sum of volumes is at most ε.

A cube itself is never of measure zero (for the proof see, for instance, [1], page 203). Consequently, no open subset of $\mathbb{R}^n$ is of measure zero.

The following statement is true:

Lemma 3.3. *The union of countably many sets of measure zero is of measure zero.*

Proof. Let $A_1, A_2, ...$ be sets of measure zero and let $\varepsilon > 0$. Since we can cover each A_i by a countable number of cubes with sum of volumes at most $\frac{\varepsilon}{2^i}$, the union $\cup_{i\in\mathbb{N}} A_i$ has a countable cover with total volume at most $\sum_{i=1}^{\infty} \frac{\varepsilon}{2^i} = \varepsilon$. □

Example 3.4. The set $\mathbb{R}^{n-1} \times \{0\}$ has measure zero in $\mathbb{R}^n$, because $\mathbb{R}^{n-1}$ is a (countable) union of cubes with vertices in integer coordinates, and each cube $S \subset \mathbb{R}^{n-1}$ has measure zero in $\mathbb{R}^n$ since it can be covered by $S \times [-\frac{\varepsilon}{2\cdot \mathrm{vol}(S)}, \frac{\varepsilon}{2\cdot \mathrm{vol}(S)}]$ for each $\varepsilon > 0$.

The definition of a subset of measure zero can be generalized for arbitrary manifolds:

Definition 3.5. A subset A of an n-dimensional manifold X is said to be of measure zero if for every $x \in X$ and for every local parametrization $\phi : U \to X$ around x the preimage $\phi^{-1}(A)$ has measure zero in $\mathbb{R}^n$.

The next lemma is not related to measure zero but we will need its special case to prove Proposition 3.7.

Lemma 3.6. *Let $U \subset \mathbb{R}^n$ be an open set and let $f : U \to \mathbb{R}^m$ be a continuously differentiable map. If $A \subset U$ is compact then there exists a constant $M \geq 0$ such that*

$$||f(x) - f(y)|| \leq M||x - y|| \quad \textit{for all } x, y \in A.$$

Proof. Since $A \subset \mathbb{R}^n$ is compact, the set $A \times A \subset \mathbb{R}^{2n}$ is also compact.

Let $D = \{(a, a);\ a \in A\}$ be the diagonal of $A \times A$.

Denote by $B_{x,r}$ the open ball centered at x with radius r, and by $\bar{B}_{x,r}$ its closure. For each $x \in U$ there exists $\bar{B}_{x,\delta_x} \subset U$. The open cover $\{B_{a,\delta_a/2};\ a \in A\}$ of A has a finite subcover, say

$$A \subset \bigcup_{j=1}^{r} B_{a_j,\delta_{a_j}/2} \subset \bigcup_{j=1}^{r} \bar{B}_{a_j,\delta_{a_j}} \subset U.$$

Then $W := \bigcup_{j=1}^{r} B_{(a_j,a_j),\delta_{a_j}}$ is a finite union of open balls in $\mathbb{R}^{2n}$, and it is an open cover of the diagonal D: if $(a, a) \in D$ then, since $a \in A$, $||a - a_j|| \leq \frac{\delta_{a_j}}{2}$ for some j, and therefore $||(a, a) - (a_j, a_j)|| = \sqrt{2}||a - a_j|| < \delta_{a_j}$, which means that $(a, a) \in B_{(a_j,a_j),\delta_{a_j}}$.

Write f as $f = (f_1, ..., f_m)$. Since f is continuously differentiable, f_i' is continuous for $i = 1, ..., m$, and therefore bounded on each ball $\bar{B}_{a_j,\delta_{a_j}}$. Hence, for each $j = 1, ..., r$ there exists a constant M_j such that

$$||f_i'(x)|| \leq M_j \quad \text{for } x \in \bar{B}_{a_j,\delta_{a_j}} \text{ and } i = 1, ..., m.$$

Then for $x, y \in \bar{B}_{a_j,\delta_{a_j}}$ the Taylor's formula around x of order $p = 1$ gives

$$f_i(y) - f_i(x) = f_i'(x + \theta_i(y - x)) \cdot (y - x), \qquad 0 < \theta_i < 1,$$

and therefore

$$\begin{aligned} ||f(y) - f(x)||^2 &= \sum_{i=1}^{m} |f_i(y) - f_i(x)|^2 \leq \sum_{i=1}^{m} ||f_i'\big(x + \theta_i(y - x)\big)||^2 \cdot ||y - x||^2 \leq \\ &\leq mM_j^2 \cdot ||y - x||^2 \leq T_1^2 \cdot ||y - x||^2, \end{aligned}$$

where $T_1 := m \cdot \max\{M_1, ..., M_r\}$.

For each $(x, y) \in W$ we have: $(x, y) \in B_{(a_j,a_j),\delta_{a_j}}$ for some j, hence $||x - a_j|| < \delta_{a_j}$ and $||y - a_j|| < \delta_{a_j}$, i.e. $x, y \in B_{\delta_{a_j}}(a_j)$, which means that $||f(x) - f(y)|| \leq T_1 \cdot ||x - y||$. Since the diagonal D lies in W, the mapping

$$\psi : (x, y) \mapsto \frac{||f(x) - f(y)||}{||x - y||}$$

is well-defined and continuous on the compact set $(A \times A) \setminus W$ and therefore is bounded above by some constant T_2. Hence, for arbitrary $x, y \in A$, either $(x, y) \in W$ and in this case $||f(x) - f(y)|| \leq T_1 \cdot ||x - y||$, or $(x, y) \in A \times A \setminus W$ and then $||f(x) - f(y)|| \leq T_2 \cdot ||x - y||$. Therefore, $M := \max\{T_1, T_2\}$ satisfies the condition from the statement of the lemma. □

Proposition 3.7. *Let U be an open subset of $\mathbb{R}^n$ and let $f : U \to \mathbb{R}^n$ be a smooth mapping. If $A \subset U$ is of measure zero, then $f(A)$ is of measure zero.*

Proof. Since $\mathbb{R}^n$ is second countable, any open set $U \subset \mathbb{R}^n$ is a countable union of open cubes each having a compact neighbourhood contained in U. Hence it suffices to prove the statement when U is a cube such that f is defined on its compact neighbourhood. In this case by Lemma 3.6 there exists a constant M such that $||f(x)-f(y)|| \leq M||x-y||$ for all $x, y \in U$. Therefore, the image of a cube $S \subset U$ with sides of length δ is contained in a cube with sides of length $\sqrt{n}M\delta$. Hence, if A is covered by cubes $S_1, S_2, ... \subset U$ with total volume less than ε, $f(A)$ has a countable cover by cubes $S'_1, S'_2, ...$ $(f(S_i) \subset S'_i)$ with total volume $< (\sqrt{n}M)^n\varepsilon$. □

By the second axiom of countability, an open cover $\{U_x;\ x \in X\}$ of X, where U_x is a neighbourhood of x in X diffeomorphic to an open subset of $\mathbb{R}^n$, has a countable subcover. This and Proposition 3.7 imply that it is enough to check that for one local parametrization around each point the preimage has measure zero. Hence, in particular, for open subsets of Euclidean spaces the two definitions of sets of measure zero match.

Proposition 3.8. *Let U be an open subset of $\mathbb{R}^n$ and let $f : U \to \mathbb{R}^m$ be a smooth map. If $m > n$ then $f(U)$ has measure zero in $\mathbb{R}^m$.*

Proof. Let $g : U \times \mathbb{R}^{m-n} \to \mathbb{R}^m$ be a mapping defined by $g(x, y) = f(x)$. It is defined on an open subset of $\mathbb{R}^m$ and smooth. Therefore, by the Proposition 3.7, the image $g(U \times \{0\}) = f(U)$ of the set $U \times \{0\}$ having measure zero in $\mathbb{R}^m$ is of measure zero. □

For the proof of Sard's theorem we will need one more proposition.

Let $n = k+l$ and write $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^l$. Let us call a "vertical slice" V_c the set $\{c\} \times \mathbb{R}^l$ for $c \in \mathbb{R}^k$. A set $\{c\} \times U$ is called of measure zero in V_c if U has measure zero in $\mathbb{R}^l$.

Proposition 3.9 (Fubini)**.** *Let A be a closed subset of $\mathbb{R}^n$. If $A \cap V_c$ has measure zero in V_c for each $c \in \mathbb{R}^k$, then A has measure zero in $\mathbb{R}^n$.*

Proof. Since A can be written as a countable union of compact sets, for example $A = \bigcup_{n\in\mathbb{N}}(A \cap \bar{B}_{0,n})$, we can assume that A is compact. It is sufficient to prove the statement for $k = 1,\ l = n-1$, since the rest will follow by induction on k.

The set A is compact, and therefore it can be covered by $[a, b] \times \mathbb{R}^{n-1}$ for some interval $[a, b] \subset \mathbb{R}$. Let $\varepsilon > 0$. Take arbitrary $t \in [a, b]$. The set $A \cap V_t$ has measure zero, hence there exists a countable collection of open cubes $S^t_1, S^t_2, ...$ in $\mathbb{R}^{n-1}$ such that $\sum_i \mathrm{vol}(S^t_i) < \varepsilon$ and $A \cap V_t$ is a subset of $\{t\} \times W_t$, where $W_t := \bigcup_i S^t_i$. Let us show that there exists an interval I_t containing t such that the vertical slice $A \cap (I_t \times \mathbb{R}^{n-1})$ is contained in $I_t \times W_t$. Assume that it is not true. Then there exists a sequence of points (t_i, w_i) in A such that $t_i \to t$ and $w_i \notin W_t$. Since A is closed and bounded, this sequence has a convergent subsequence, say, with limit $(t, w) \in A$. This would mean that $w \in W_i$, which is a contradiction.

The cover $\{I_t;\ t \in [a, b]\}$ of interval $[a; b]$ has a finite subcover. Now we will need the following lemma:

Lemma 3.10. *If an interval $I = [a, b]$ is covered by a finite number of intervals $I_i = [a_i, b_i] \subset I$, then there exists a subcover with total length $\leq 2(b-a)$.*

Proof of lemma. Take an arbitrary minimal subcover $I_1, ..., I_k$, i.e. such a subcover that after removing any interval it does not cover I any more. If $a_1 \leq a_2 \leq ... \leq a_k$, then by minimality $b_1 \leq b_2 \leq ... \leq b_k$. Moreover, $b_i \leq a_{i+2}$ for all $1 \leq i \leq k-2$, because otherwise we could remove $[a_{i+1}, b_{i+1}]$. Hence, $a_1 < b_1 \leq a_3 < b_3 \leq ...$, so the total length of $I_1, I_3, ...$ is at most $b - a$. Analogously the total length of $I_2, I_4, ...$ is at most $b - a$. □

By lemma, our finite cover will have a subcover $I_{t_1}, ..., I_{t_k}$ with total length $\leq 2(b-a)$. The corresponding cubes $I_{t_i} \times S_j^{t_i}$ cover A and their total volume is $\leq 2(b-a)\varepsilon$. □

Now we state and prove the Sard's theorem.

Theorem 3.11 (Sard's theorem)**.** *Let $U \subset \mathbb{R}^n$ be open and let $f : U \to \mathbb{R}^m$ be a smooth mapping. Then the set of critical values of f has measure zero in $\mathbb{R}^m$.*

Proof. Denote by C the set of critical points of f in U. Then the set of critical values is $f(C)$. Denote by C_i the set of points x such that all partial derivatives of order $\leq i$ vanish at x. Clearly, $C \supset C_1 \supset C_2 \supset ...$.

First, we prove that

$$f(C_i) \text{ has measure zero for } i > \frac{n}{m} - 1. \tag{3.1}$$

Proof of (3.1). We will take an arbitrary cube $S \subset U$ such that f is defined in some open neighbourhood of its closure $\bar{S}$ and show that $f(C_i \cap S)$ has measure zero in $\mathbb{R}^m$ for $i > \frac{n}{m} - 1$. Since U can be covered by countably many of such cubes, this proves that $f(C_i)$ has measure zero.

Let S have sides of length δ. By Taylor's formula of order $p = i$ and compactness of S, for $x \in C_i \cap S$ and $x + h \in S$ we have:

$$f(x + h) = f(x) + R(x, h),$$

where $||R(x, h)|| < a||h||^{i+1}$ and constant a depends only on f and S (all partial derivatives of order $\leq i$ vanish at x, hence they will not appear in Taylor's expansion).

Now divide S into r^n cubes with sides of length δ/r and consider those of them which contain points of C_i. Let S_1 be such a cube and let $x \in C_i \cap S_1$. Each point of

S_1 can be written as $x + h$, where $||h|| < \sqrt{n} \cdot \delta/r$. Therefore, $f(S_1)$ is contained in a cube with center at $f(x)$ and sides of length

$$2a(\sqrt{n}\delta/r)^{i+1}.$$

Hence $f(C_i \cap S)$ is contained in the union of at most r^n cubes with total volume

$$V \leq r^n \left(2a(\sqrt{n}\delta/r)^{i+1}\right)^m = (2a(\sqrt{n}\delta)^{i+1})^n \cdot r^{n-(i+1)m}.$$

If $i > \frac{n}{m} - 1$, then V tends to 0 as $r \to \infty$. Therefore, $f(C_i \cap S)$ has measure zero. □

Notice now that we have already proved Sard's theorem for $n = 1$: if $m \geq 2$ then $f(U)$ has measure zero by Proposition 3.8, if $m = 1$ then $C = C_1$, and we apply (3.1).

The proof for $n \geq 2$ will be done by induction. Assuming that the statement of Sard's theorem is true for $n - 1$, we show that

$$f(C \setminus C_1) \text{ has measure zero, and} \tag{3.2}$$

$$f(C_i \setminus C_{i+1}) \text{ has measure zero for } i \geq 1. \tag{3.3}$$

Since (3.1), (3.2), and (3.3) imply that $f(C)$ is a union of countably many sets of measure zero, this proves the statement for n.

Proof of (3.2). For each $x \in C \setminus C_1$ we will prove that there exists an open neighbourhood V_x of x such that $f(C \cap V_x)$ has measure zero. Since $f((C \setminus C_1) \cap V_x) \subset f(C \cap V_x)$, and by the second axiom of countability we can choose countably many x's such that the sets V_x cover $C \setminus C_1$, this will mean that $f(C \setminus C_1)$ has measure zero.

Since $x \notin C_1$, for some i and j the partial derivative $\frac{\partial f_i}{\partial x_j}$ is nonzero at x. Without loss of generality, $i = j = 1$. Then the mapping $h : U \to \mathbb{R}^n$ defined by

$$h(x) = (f_1(x), x_2, ..., x_n)$$

has a nonsingular derivative at x. Thus, it maps some open neighbourhood V_x of x diffeomorphically onto an open set $V' \subset \mathbb{R}^n$. The mapping $g = f \circ h^{-1}$ from V' to $\mathbb{R}^m$ has the same critical values as f restricted to V, and preserves the first coordinate:

$$g(t, x_2, ..., x_n) = (t, y_2, ..., y_m).$$

Let g_t be a mapping of the set $(\{t\} \times \mathbb{R}^{n-1}) \cap V'$ to $\{t\} \times \mathbb{R}^{m-1}$ induced by g:

$$g_t(x_2, ..., x_n) = (y_2, ..., y_m).$$

Since

$$\left(\frac{\partial g_i}{\partial x_j}\right) = \begin{pmatrix} 1 & 0 \\ * & \left(\frac{\partial (g_t)_i}{\partial x_j}\right) \end{pmatrix},$$

a point of $(\{t\}\times\mathbb{R}^{n-1})\cap V'$ is a critical point for g_t if and only if it is a critical point for g. By induction hypothesis, the set of critical values of g_t has measure zero in $\{t\}\times\mathbb{R}^{m-1}$. Now we would like to use Fubini's theorem to show that the set of critical values of g has measure zero in $\mathbb{R}^m$. But we can't do it directly because the set of critical values of g need not be closed. However, the set of critical points of g is closed whenever $m \leq n$ (if $m > n$ then the set of critical points is the whole image of g and by Proposition 3.8 has measure zero in $\mathbb{R}^m$): critical points of g are the points where the Jacobian matrix J_g is not surjective, i.e. the set of critical points is the intersection of the zero sets of the $\binom{n}{m}$ determinants of $(m \times m)$-submatrices of J_g. And since V' is a countable union of compact sets, so is the set of critical points. Therefore its image, the set of critical values, is a countable union of compact sets and applying Fubini's theorem to each of these compact sets separately gives the desired result. □

Proof of (3.3). Similarly, here we will prove that for each $x \in C_i \setminus C_{i+1}$ there exists an open neighbourhood V_x of x such that $f(C_i \cap V_x)$ has measure zero.

Since $x \in C_i \setminus C_{i+1}$, there exists some i-th partial derivative $\rho = \frac{\partial^i f_k}{\partial x_{k_1}\partial x_{k_2}...\partial x_{k_i}}$ of f which is zero but has some first derivative, say $\frac{\partial \rho}{\partial x_1}$, which is not zero at x. The mapping $h : U \to \mathbb{R}^n$ defined by

$$h(x_1, x_2, ..., x_n) = (\rho(x), x_2, ..., x_n)$$

has a nonsingular derivative at x, and therefore maps some open neighbourhood $V_x \subset \mathbb{R}^n$ of x diffeomorphically onto an open set V'. By construction, the set $h(C_i \cap V_x)$ lies in the hyperplane $\{0\} \times \mathbb{R}^{n-1}$. Consider the mapping $g = f \circ h^{-1} : V' \to \mathbb{R}^m$ and its restriction

$$\bar{g} : (\{0\} \times \mathbb{R}^{n-1}) \cap V' \to \mathbb{R}^m.$$

By induction, the set of critical values of $\bar{g}$ has measure zero in $\mathbb{R}^m$. By the chain rule $J_g(z) = J_f(h^{-1}(z)) \circ J_{h^{-1}}(z)$, hence all the points in $h(C_i \cap V_x)$ are critical points of g of type at least C_1. Since the Jacobian $J_{\bar{g}}$ is precisely J_g without the first column, points in $h(C_i \cap V_x)$ are also critical points of $\bar{g}$. Therefore, the set $\bar{g}(h(C_i \cap V_x)) = f(C_i \cap V_x)$ has measure zero in $\mathbb{R}^m$. □

The proof of Sard's theorem is finished. □

Now we generalize Sard's theorem to arbitrary manifolds:

Theorem 3.12. *Let $f : X \to Y$ be a smooth mapping of manifolds. Then the set of critical values of f has measure zero in Y.*

Proof. By the second axiom of countability, an open cover $\{U_x;\ x \in X\}$ of X, where each U_x is the image of some local parametrization $\phi_x : U'_x \to U_x$ around x, has a countable subcover. Hence, it suffices to check that $f(U_x \cap C)$ has measure zero in Y. Similarly, an open cover $\{V_y;\ y \in Y\}$ of Y, where each V_y is the image of some local parametrization $\varphi_y : V'_y \to V_y$ around y, has a countable subcover.

Therefore, it suffices to check that the set of critical values of f restricted to $W = U_x \cap f^{-1}(V_y)$ has measure zero in V_y for $x \in X$ and $y \in Y$. Since $\phi_x : \phi_x^{-1}(W) \to W$ is a diffeomorphism, the set of critical values of $f|_W$ is precisely the set of critical values of $f \circ \phi_x|_{\phi_x^{-1}(W)}$, which is a mapping from the open subset $\phi_x^{-1}(W)$ of $\mathbb{R}^n$ to the open subset V_y of Y. But a subset A of V_y has measure zero in V_y if and only if the set $\varphi^{-1}(A)$ has measure zero in $\mathbb{R}^m$. Since φ is a diffeomorphism, $\varphi^{-1}(f|_W(C))$ is precisely the set of critical values of a smooth map $h = \varphi^{-1} \circ f \circ \phi$ from the open subset $\phi_x^{-1}(W)$ of $\mathbb{R}^n$ to $\mathbb{R}^m$, which by Sard's theorem has measure zero. Hence, $f|_W(C)$ is indeed of measure zero in V_y. □

Since no open subset of $\mathbb{R}^m$ is of measure zero, we have the following corollary:

Corollary 3.13. *The set of regular values of a smooth map $f : X \to Y$ is dense in Y.*

4 Morse functions

In this chapter we consider smooth *functions* on a differentiable manifold X, i.e., smooth mappings $f : X \to \mathbb{R}$. If point x is regular then by the local submersion theorem there exists a coordinate system around x such that f is just the first coordinate function. Hence, the behaviour in regular points is clear. So we will focus on critical points. From Analysis we know that the information about the behaviour of a function $f : U^{open} \to \mathbb{R}$ at critical point is provided by the matrix of its second partial derivatives, the Hessian matrix

$$H_{i,j} = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right),$$

in case when it is nonsingular: if H is positive-definite at x then f has a local minimum at x, if H is negative-definite at x then f has a local maximum at x, otherwise, if H is nonsingular and has both positive and negative eigenvalues, point x is a saddle point.

Definition 4.1. If $f : U^{open} \to \mathbb{R}$ is a smooth function and the Hessian matrix H is nonsingular at a critical point x of f, then x is called a *nondegenerate critical point* of function f.

Let us show that this definition can be generalized via local parametrizations to a function f on an arbitrary manifold X. For a critical point x of function $f : X \to \mathbb{R}$ we will say that it is a nondegenerate critical point if for any local parametrization $\phi : U \to X$ around x with $\phi(0) = x$, point 0 is a nondegenerate critical point of $f \circ \phi$. First of all, if $\phi : U \to X$ is a local parametrization around x then point 0 is a critical point of $f \circ \phi$ since $d(f \circ \phi)_0 = df_x \circ d\phi_0 = 0$. Thus, we only have to check that if 0 is a nondegenerate critical point for $f \circ \phi$, then 0 is a nondegenerate critical point for $f \circ \varphi$, where $\varphi : V \to X$ is another local parametrization around x with $\varphi(0) = x$. Since $\varphi^{-1} \circ \phi : U \to V$ is a diffeomorphism, it suffices to prove the following lemma:

Lemma 4.2. *Let f be a function on an open subset of $\mathbb{R}^n$ with a nondegenerate critical point 0, and let ψ be a diffeomorphism with $\psi(0) = 0$. Then 0 is a nondegenenerate critical point for $f \circ \psi$.*

Proof. The proof is done by computing the Hessian matrix of $f \circ \psi$ at 0 explicitly. By the chain rule,

$$\frac{\partial (f \circ \psi)}{\partial x_i}(x) = \sum_a \frac{\partial f}{\partial x_a}[\psi(x)] \cdot \frac{\partial \psi_a}{\partial x_i}(x).$$

Hence,

$$\frac{\partial^2(f\circ\psi)}{\partial x_i\partial x_j}(x)=\sum_a\sum_b\frac{\partial^2 f}{\partial x_a\partial x_b}[\psi(x)]\cdot\frac{\partial\psi_a}{\partial x_i}(x)\frac{\partial\psi_b}{\partial x_j}(x)+\sum_a\frac{\partial f}{\partial x_a}[\psi(x)]\frac{\partial^2\psi_a}{\partial x_i\partial x_j}(x).$$

The second summand vanishes at 0, since 0 is a critical point of f. Therefore, we obtain:

$$\left(\frac{\partial^2(f\circ\psi)}{\partial x_i\partial x_j}(0)\right)=J_\psi(0)^T\left(\frac{\partial^2 f}{\partial x_i\partial x_j}(0)\right)J_\psi(0),\tag{4.1}$$

where $J_\psi(0)$ is a Jacobian matrix of ψ at 0. Since $J_\psi(0)$ and $\left(\frac{\partial^2 f}{\partial x_i\partial x_j}(0)\right)$ are nonsingular, so is $\left(\frac{\partial^2(f\circ\psi)}{\partial x_i\partial x_j}(0)\right)$, which means that 0 is a nondegenerate critical point of $f\circ\psi$. □

Definition 4.3. Let x be a critical point of a smooth function $f : X \to \mathbb{R}$. If for any local parametrization $\phi : U \to X$ around x with $\phi(0) = x$, point 0 is a nondegenerate critical point of $f \circ \phi$, then x is called a **nondegenerate critical point** of f.

The next statement completely describes the behaviour of a function at its nondegenerate critical point. The proof can be found in [2].

Lemma 4.4 (Morse lemma)**.** *Let p be a nondegenerate critical point of $f : X \to \mathbb{R}$. Then there exists a local coordinate system $(x_1, ..., x_n)$ in some neighbourhood U of p such that $x_i(p) = 0$ and*

$$f(x)=f(p)-x_1^2-...-x_\alpha^2+x_{\alpha+1}^2+...+x_n^2$$

for all $x \in U$.

A straightforward corollary is the following:

Corollary 4.5. *Nondegenerate critical points are isolated from other critical points.*

The number α from Morse lemma is called the *index of f at p*. With a local coordinate system around p it can also be defined as the dimension of the largest linear subspace of $\mathbb{R}^n$ where the Hessian matrix is negative-definite. Since with a change of a local coordinate system the Hessian matrix transforms as in (4.1), this will be well-defined. Intuitively, the index is the number of directions in which f decreases.

Definition 4.6. A function $f : X \to \mathbb{R}$ whose all critical points are nondegenerate is called a **Morse function**.

With the next theorem we will prove that there exist Morse functions on each manifold. The motivation to consider Morse functions is that they give the information about the topology of their domain manifolds. More precisely, consideration of a Morse function f on X and its critical points with their indices enables to construct a CW compex homotopically equivalent to X. The reader is referred to [2] for further studying.

For a function f on a manifold $X \subset \mathbb{R}^N$ and for an arbitrary vector $a = (a_1, ..., a_N) \in \mathbb{R}^N$ let us denote

$$f_a = f + a_1 x_1 + ... + a_N x_N.$$

As usually in measure theory, the statement holds for "almost all" values if all the values for which it does not hold constitute a set of measure zero.

Lemma 4.7. *Let $U \subset \mathbb{R}^n$ and let $f : U \to \mathbb{R}$ be a smooth function. Then for almost all $a \in \mathbb{R}^n$, f_a is a Morse function on U.*

Proof. Define the mapping $g : U \to \mathbb{R}^n$ by

$$g(x) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, ..., \frac{\partial f}{\partial x_n}\right).$$

Since $(df_a)_x = g(x) + a$, a point x is critical for f_a if and only if $g(x) = -a$. The Hessian of f_a at x is dg_x. Now, if the point $-a$ is a regular value of g, then at each point $x \in g^{-1}(-a)$ the derivative dg_x is nonsingular, which means that x is a nondegenerate critical point of f_a. By Sard's theorem almost all values of g are regular, hence almost all f_a's are Morse functions. □

Now we prove the statement for arbitrary manifolds.

Theorem 4.8. *Let $X \subset \mathbb{R}^N$ be a manifold and let $f : X \to \mathbb{R}$ be a smooth function. Then for almost all $a \in \mathbb{R}^N$, f_a is a Morse function on X.*

Proof. First, we will show that at each point $x \in X$ the restriction of some $n = \dim X$ coordinate functions $x_{i_1}, x_{i_2}, ..., x_{i_n}$ to X forms a local coordinate system around x. Indeed, the tangent space $T_x(X)$ is an n-dimensional linear subspace of $\mathbb{R}^N$, hence there exist some n vectors $e_{i_1}, ..., e_{i_n}$ from the standard basis of $\mathbb{R}^N$ such that the projections of the basis vectors of $T_x(X)$ to the subspace spanned by $e_{i_1}, ..., e_{i_n}$ are linearly independent. Hence, the projection π of $T_x(X)$ to the subspace $< e_{i_1}, ..., e_{i_n} >$ is an isomorphism. But then by the inverse function theorem the projection of X to $< e_{i_1}, ..., e_{i_n} >$ is a local diffeomorphism at x because π is its derivative at x. Therefore, $x_{i_1}, x_{i_2}, ..., x_{i_n}$ form a local coordinate system around x.

By the second axiom of countability, there exist countably many open sets $U_i \subset X$ such that on each of them some n of the N coordinate functions of $\mathbb{R}^N$ form a coordinate system. Take one of these sets U_i and for convenience assume that $(x_1, ..., x_n)$ is a coordinate system on U_i. Let us prove that the set $S_i = \{a : f_a \text{ is not Morse on } U_i\}$ is of measure zero. For each $(N-n)$-tuple $c = (c_{n+1}, ..., c_N)$ consider the function

$$f_{(0,c)} = f + c_{n+1} x_{n+1} + ... + c_N x_N.$$

By Lemma 4.7 for almost all n-tuples $b = (b_1, ..., b_n)$ the function

$$f_{(b,c)} = f_{(0,c)} + b_1 x_1 + ...b_n x_n$$

is Morse on U_i. From the proof of Lemma 4.7 it follows that $a \in S_i$ if and only if $-a$ is a critical value of g. Since the set of critical values is a countable union of compact subset of $\mathbb{R}^n$, we can apply Fubini's theorem: the intersection of S_i with each "horizontal slice" $\mathbb{R}^n \times \{c\}$ has measure zero, therefore S_i has measure zero.

Since $\{a : f_a \text{ is not Morse on } X\} = \bigcup_i S_i$, it is of measure zero. □

5 Whitney embedding theorem

Whitney embedding theorem states that every n-dimensional manifold can be embedded in $\mathbb{R}^{2n}$. In this chapter we will prove a weaker version, namely, that it can be embedded in $\mathbb{R}^{2n+1}$. For the proof we will need the following object:

Definition 5.1. A **tangent bundle** $T(X)$ of a manifold $X \subset \mathbb{R}^N$ is a subset of $X \times \mathbb{R}^N$ defined by

$$T(X) = \{(x, v) \in X \times \mathbb{R}^N : v \in T_x(X)\}.$$

Any smooth map $f : X \to Y$ induces a global *derivative map* $df : T(X) \to T(Y)$ defined by

$$df(x, v) = (f(x), df_x(v)).$$

Naturally, this derivative map is smooth. To show this let $X \subset \mathbb{R}^N$ and $Y \subset \mathbb{R}^M$. The map $f : X \to \mathbb{R}^M$ is smooth, and therefore can be extended around each point of X to a smooth map $F : U \to \mathbb{R}^M$, where U is an open subset of $\mathbb{R}^N$. But then the derivative map $dF : T(U) \to \mathbb{R}^{2M}$ locally extends df. Since $T(U) = U \times \mathbb{R}^N$ is an open subset of $\mathbb{R}^{2N}$, we have a local extension of df to a smooth map on an open set, which means that df is smooth.

Tangent bundles of diffeomorphic manifolds are diffeomorphic. This follows from the fact that if $f : X \to Y$ is a diffeomorphism, then so is $df : T(X) \to T(Y)$.

Proposition 5.2. *The tangent bundle of a manifold X is a manifold, and dim $T(X) = 2 \cdot$ dim X.*

Proof. Let W be an open subset of X. Then $T(W) = T(X) \cap (W \times \mathbb{R}^N)$ is an open subset of $T(X)$, because $W \times \mathbb{R}^N$ is open in $\mathbb{R}^{2N}$. Now if $\phi : U \to W$ is a local parametrization of W, where $U \subset \mathbb{R}^n$ is open, then $d\phi : T(U) \to T(W)$ is a diffeomorhism. Since $T(U) = U \times \mathbb{R}^n$ is an open subset of $\mathbb{R}^{2n}$, $d\phi$ is a local parametrization of $T(W)$. Every point of $T(X)$ has such a neighbourhood $T(W)$, therefore $T(X)$ is a manifold of dimension $2n$. □

First we prove the following:

Theorem 5.3. *Every n-dimensional manifold admits an injective immersion in $\mathbb{R}^{2n+1}$.*

Proof. Let $X \subset \mathbb{R}^N$ be an n-dimensional manifold. We will prove that if $M > 2n+1$ and X admits an injective immersion $f : X \to \mathbb{R}^M$, then there exists a vector $a \in \mathbb{R}^M$ such that the composition of f with the projection π of $\mathbb{R}^M$ onto the orthogonal complement of a is still an injective immersion. The orthogonal complement of vector $a \in \mathbb{R}^M$ is isomorphic to $\mathbb{R}^{M-1}$. Hence, by induction we will produce a linear projection of $\mathbb{R}^N$ onto its $(2n+1)$-dimensional vector subspace that restricts to a one-to-one immersion of X.

Define a mapping $h : X \times X \times \mathbb{R} \to \mathbb{R}^M$ by $h(x, y, t) = t(f(x) - f(y))$ and a mapping $g : T(X) \to \mathbb{R}^M$ by $g(x, v) = df_x(v)$. Since $M > 2n+1$, all the values taken by h and g are critical, and therefore by Sard's theorem both $\mathrm{Im}(h)$ and $\mathrm{Im}(g)$ have measure zero in $\mathbb{R}^M$. Pick a nonzero vector $a \in \mathbb{R}^M$ which is in neither of the images.

Let π be a projection of $\mathbb{R}^M$ onto the orthogonal complement H of a. Let us show that $\pi \circ f : X \to H$ is a one-to-one immersion of X. Suppose it is not injective, i.e. $\pi \circ f(x) = \pi \circ f(y)$ for some $x \neq y$. Then $f(x) - f(y) = ta$ for some $t \in \mathbb{R}$. Since $x \neq y$ and f is injective, $f(x) \neq f(y)$ and therefore $t \neq 0$. But then $h(x, y, 1/t) = a$, which contradicts to the choice of a. Hence, $\pi \circ f$ is injective.

Suppose that $\pi \circ f$ is not an immersion. Then there exist $x \in X$ and a nonzero vector $v \in T_x(X)$ such that $d(\pi \circ f)_x(v) = 0$. By the chain rule and linearity of π, we have $d(\pi \circ f)_x = \pi \circ df_x$. Hence, $df_x(v) = ta$ for some $t \in \mathbb{R}$. Since f is an immersion, $t \neq 0$. But then $g(x, v/t) = df_x(v/t) = a$, which again contradicts to the choice of a. □

For compact manifolds embeddings are precisely one-to-one immersions, thus we have proved that any n-dimensional compact manifold embeds in $\mathbb{R}^{2n+1}$. For the noncompact case we need the immersion to be proper. We need the next two lemmas to prove Theorem 5.6. Its corollary (Corollary 5.7) will help us to construct the required proper injective immersion.

Lemma 5.4. *Any open set $W \subset \mathbb{R}^N$ admits an exhaustion by compact sets, i.e. a sequence $\{K_i\}_{i \in \mathbb{N}}$ of compact sets such that $K_i \subset$ Interior (K_{i+1}), and $\bigcup_i K_i = W$.*

Proof. Consider an arbitrary countable base of W (W is second countable), and choose those open sets $U_1, U_2, ...$ of this base that have a compact closure contained in W. These sets $U_1, U_2, ...$ cover W. Indeed, every point x of W has an open neighbourhood $U(x)$ whose closure is compact and is contained in W. This neighbourhood $U(x)$ is a union of some sets from the base, and clearly all of those sets should have a compact closure contained in W, as $U(x)$ does.

Now, define $K_1 = \bar{U}_1$. The set K_1 is compact and covered by $U_1, U_2, \ldots$. Let U_{i_1}, U_{i_2},..., U_{i_k} be a finite subcover of K_1, where $i_1 < i_2 < \ldots < i_k$. Set $K_2 = \bar{U}_1 \cup \bar{U}_2 \cup \ldots \cup \bar{U}_{i_k}$.

The sets $K_3, K_4, \ldots$ are built analogously. The conditions $K_i \subset$ Interior (K_{i+1}) and $\bigcup_i K_i = W$ are satisfied. □

Lemma 5.5. *Function $f : \mathbb{R} \to \mathbb{R}$ defined by*

$$f(x) = \begin{cases} e^{-1/x}, & \text{if } x > 0, \\ 0, & \text{if } x \leq 0 \end{cases}$$

is smooth[1].

Proof. For $x \neq 0$ the smoothness is obvious.

Now, since $\lim_{x \searrow 0} e^{-1/x} = 0$, f is continuous at 0. Notice that

$$(e^{-1/x})' = \frac{e^{-1/x}}{x^2}$$

and

$$\left(\frac{e^{-1/x}}{x^m}\right)' = \frac{e^{-1/x}}{x^{m+2}} - m\frac{e^{-1/x}}{x^{m+1}}.$$

Hence, if we show that

$$\lim_{x \searrow 0} \frac{e^{-1/x}}{x^m} = 0 \tag{5.1}$$

for all integers $m \geq 0$, it will be proved by induction that $\lim_{x \searrow 0} f^{(n)}(x) = 0$ for all $n \geq 0$.

The Taylor's expansion of the exponential function gives:

$$e^{1/x} = \sum_{i=0}^{\infty} \frac{1}{i!}\left(\frac{1}{x^i}\right) \geq \frac{1}{(m+1)!}\left(\frac{1}{x^{m+1}}\right)$$

for all $m \geq 0$ and $x > 0$. Therefore, $0 \leq \frac{e^{-1/x}}{x^m} \leq (m+1)!x$, and since $\lim_{x \searrow 0}(m+1)!x = 0$, (5.1) is proved. □

Theorem 5.6. *Let X be an arbitrary subset of $\mathbb{R}^N$. For any cover of X by relatively open subsets $\{U_\alpha; \alpha \in A\}$ there exists a sequence $\{\theta_i\}_{i\in\mathbb{N}}$ of smooth functions from X to $\mathbb{R}$, called a* ***partition of unity*** *subordinate to the open cover $\{U_\alpha; \alpha \in A\}$, with the following properties:*

(a) For all i, $0 \leq \theta_i(x) \leq 1$ for all $x \in X$, and θ_i is identically zero except on some open set whose closure is contained in one of the sets U_α.

(b) Each $x \in X$ has a neighbourhood on which all but finitely many functions θ_i are identically zero, and $\sum_i \theta_i(x) = 1$.

[1] This is an example of a smooth non-analytic function.

Proof. Each set $U_\alpha \subset X$ can be extended to an open set $W_\alpha \subset \mathbb{R}^N$. Let $W = \bigcup_\alpha W_\alpha$. We construct a sequence of open balls $\{B_i\}_{i\in\mathbb{N}}$ such that each of them has a closure contained in at least one of the sets W_α, and each $x \in W$ is contained in finitely many and in at least one of these balls. Let $\{K_i\}_{i\in\mathbb{N}}$ be an exhaustion of W by compact sets (see Lemma 5.4). Consider the collection of open balls of $\mathbb{R}^N$ whose closures belong to at least one W_α. This is an open cover of W. Select a finite number of these balls $B_1, B_2, ..., B_r$ covering the set K_2. For each $j \geq 3$ the compact set $K_j \setminus \text{Interior}(K_{j-1})$ is contained in the open set $W \setminus K_{j-2}$. Open balls whose closures belong to both $W \setminus K_{j-2}$ and at least one W_α form an open cover of $K_j \setminus \text{Interior}(K_{j-1})$. Select a finite subcover and add it to our sequence $\{B_i\}_{i\in\mathbb{N}}$. By construction, $\bigcup_i B_i = W$ and each set K_j intersects with finitely many balls B_i. Since each point of W is contained in some K_j, it belongs to a finite number of balls.

Next, observe that for the function f defined as in Lemma 5.5 and for $a \in \mathbb{R}^N, r > 0$ the function $\phi_{a,r} : \mathbb{R}^N \to \mathbb{R}$ defined by

$$\phi_{a,r}(x) = f(r^2 - ||x||^2)$$

is smooth, equals zero outside the open ball $B_{a,r}$ with radius r and center at a, and is strictly positive on this ball.

Construct the sequence of smooth functions $\{\eta_i\}_{i\in\mathbb{N}}$ such that η_i is positive on B_i and zero outside. All of the functions η_i are smooth and for each $x \in W$ the sum $\sum_i \eta_i(x)$ is finite and positive, hence $\sum_{i=1}^\infty \eta_i$ is a smooth positive function on W and the quotient

$$\frac{\eta_i}{\sum_{i=1}^\infty \eta_i}$$

is a well-defined smooth function on W. Let θ_i be the restriction of this to X, and we are done. □

Corollary 5.7. *On an arbitrary manifold X there exists a proper map $\rho : X \to \mathbb{R}$.*

Proof. Let $\{U_\alpha; \alpha \in A\}$ be an open cover of X by sets with compact closures, and let $\{\theta_i\}_{i\in\mathbb{N}}$ be a subordinate partition of unity. Then

$$\rho = \sum_{i=1}^\infty i\theta_i$$

is a well-defined smooth positive function. If $\rho(x) \leq j$, then, since $\sum_i \theta_i(x) = 1$, at least one of the first j functions $\theta_1, ..., \theta_j$ should be nonzero at x: if we assume that $\theta_1(x) = ... = \theta_j(x) = 0$, then $j \geq \rho(x) = \sum_{i=1}^\infty i\theta_i(x) \geq (j+1)\sum_{i=j+1}^\infty \theta_i(x) = j+1$, a contradiction. Hence, the preimage $\rho^{-1}([-j, j])$ is contained in the set

$$\bigcup_{i=1}^{j} \{x; \theta_i(x) \neq 0\},$$

which has a compact closure, because the closure of $\{x; \theta_i(x) \neq 0\}$ is contained in some U_α having a compact closure. Since each compact set in $\mathbb{R}$ is contained in some interval $[-j, j]$, this implies that the preimage of every compact set is compact. □

We can prove now the embedding theorem for the noncompact case.

Theorem 5.8 (Whitney theorem). *Every n-dimensional manifold embeds in $\mathbb{R}^{2n+1}$.*

Proof. Let X be an n-dimensional manifold. We already know that X admits an injective immersion in $\mathbb{R}^{2n+1}$. Composing this immersion with any diffeomorphism of $\mathbb{R}^{2n+1}$ into its unit ball, for example with $z \mapsto \frac{z}{\sqrt{1+||z||^2}}$, we obtain another injective immersion $f : X \to \mathbb{R}^{2n+1}$ such that

$$||f(x)|| < 1 \text{ for all } x \in X. \tag{5.2}$$

Let $\rho : X \to \mathbb{R}$ be a proper function on X. Then $F : X \to \mathbb{R}^{2n+2}$ defined by $F(x) = (f(x), \rho(x))$ is an injective immersion of X in $\mathbb{R}^{2n+2}$. Similarly as in Theorem 5.3, for almost all $a \in \mathbb{R}^{2n+2}$ the composition of F with the projection π of $\mathbb{R}^{2n+2}$ onto the orthogonal complement H of a is an injective immersion. Let us choose a so that it is a unit vector and different from the north and south poles of S^{2n+1}, i.e. vectors $(0, ..., 0, \pm 1)$. We will show that $\pi \circ F : X \to H$ is proper.

It suffices to show that for any constant $c > 0$ there exists another constant d such that $||\pi \circ F(x)|| \leq c$ implies $|\rho(x)| \leq d$: since an arbitrary compact set $K \subset \mathbb{R}^{2n+1}$ is contained in some ball $\bar{B}_{0,c}$, this will mean that its preimage $(\pi \circ F)^{-1}(K)$, which is closed by continuity of $\pi \circ F$, is contained in $\rho^{-1}([-d; d])$, which is compact because ρ is proper, and therefore $(\pi \circ F)^{-1}(K)$ is compact itself showing that the preimage of every compact set is compact, i.e. $\pi \circ F$ is proper.

Suppose that for some constant c there exists a sequence $\{x_i\}_{i \in \mathbb{N}}$ of points in X such that $||\pi \circ F(x_i)|| \leq c$ but $\lim_{i \to \infty} |\rho(x_i)| = \infty$. By construction, $F(x_i) - \pi \circ F(x_i)$ is a multiple of a, therefore the vector

$$w_i = \frac{1}{\rho(x_i)}[F(x_i) - \pi \circ F(x_i)]$$

is also a multiple of a for each $i \in \mathbb{N}$. But $||\frac{\pi \circ F(x_i)}{\rho(x_i)}|| \leq \frac{c}{|\rho(x_i)|} \to 0$ as $i \to \infty$, and

$$\frac{F(x_i)}{\rho(x_i)} = \left(\frac{f(x_i)}{\rho(x_i)}, 1\right) \to (0, ..., 0, 1)$$

as $i \to \infty$ by (5.2). Hence, $w_i \to (0, ..., 0, 1)$. Since each w_i is a multiple of a, so must be the limit $(0, ..., 0, 1)$, but this contradicts to the choice of a. □

6 Conclusion

In this project paper we prove Sard's theorem, existence of Morse functions and version of Whitney embedding theorem. The main source for writing this project paper was [1]. The book [5] was helpful in understanding some parts of the content. The interested reader is referred to [3] for introduction to differential topology.

7 Povzetek naloge v slovenskem jeziku

V prvem poglavju zaključne naloge definiramo diferenciabilne mnogoterosti, njene tangentne prostore v dani točki ter odvod preslikav definiranih na mnogoterosti. Prav tako podamo posplošitev klasičnega izreka o odvajanju kompozituma funkcij ter izreka o inverzni funkciji za primer, ko so definicijska območja splošne mnogoterosti in ne zgolj odprte podmnožice Evklidksega prostora.

V drugem poglavju definiramo imerzijo, vložitev ter submerzijo. Pokažemo, da sta lokalno imerzija in submerzija ekvivalentni kanonični imerziji oziroma kanonični submerziji. Pokažemo tudi, da je vsaka injektivna imerzija z lastnostjo, da so praslike kompaktnih množic zopet kompakti, topološka vložitev. Definiramo tudi regularne in kritične točke ter regularne in kritične vrednosti preslikav na mnogoterostih. Dokažemo, da je praslika regularne vrednosti vedno podmnogoterost v domeni funkcije.

V tretjem poglavju definiramo podmnožice mnogoterosti z Lebesgueovo mero nič in dokazujemo poseben primer Fubinijevega izreka. Nato dokažemo Sardov izrek, ki pravi da ima množica kritičnih vrednosti gladke preslikave $f : X \to Y$ med mnogoterostima X in Y Lebesgueovo mero nič v Y.

V četrtem poglavju definiramo neizrojene kritične točke funkcije na mnogoterostih. To predstavlja posplošitev definicije neizrojenih kritičnih točk, torej točk, kjer je odvod ničeln in Hessejeva matrika neizrojena, funkcij, definiranih na odprtih podmnožicah Evklidskega prostora. Brez dokaza navedemo Morsejevo lemo, ki popolnoma opisuje obnašanje funkcije v neizrojeni kritični točki. Dokažemo pa, da je za poljubno gladko funkcijo f na mnogoterosti $X \subset \mathbb{R}^N$ pri skoraj vsakem $a \in \mathbb{R}^N$ funkcija $f_a(x) := f(x)+a\cdot x$ Morsejeva, kar po definiciji pomeni, da so vse njene kritične točke neizrojene. Najprej to dokažemo za odprte podmnožice Evklidskega prostora in potem posplošimo za poljubne mnogoterosti.

V petem poglavju dokažemo, da lahko vsako n-razsežno diferenciabilno mnogoterost vložimo v $\mathbb{R}^{2n+1}$. Tukaj velja sponmiti, da smo mnogoterost definirali kot že vloženo v dovolj velik Evklidski prostor. V dokazu tega izreka definiramo pojem tangentnega svežnja mnogoterosti. Najprej dokažemo da za vsako n-dimenzionalno diferenciabilno mnogoterost X obstaja injektivna imerzija v $\mathbb{R}^{2n+1}$. Glavna ideja dokaza je, da če ima mnogoterost X injektivno imerzijo f v $\mathbb{R}^M$ in je $M > 2n+1$, potem obstaja tak vektor $a \in \mathbb{R}^M$, da je kompozitum f in projekcije $\mathbb{R}^M$ na ortogonalni komplement H vektorja a še vedno injektivna imerzija X. Za kompaktne mnogoterosti je injektivna imerzija natanko vložitev. Za nekompakten primer pa potrebujemo, da je injektivna imerzija prava, torej da so praslike kompaktnih množic zopet kompakti. Dokažemo tudi, da na poljubni mnogoterosti X obstaja prava funkcija in s pomočjo te funkcije konstruiramo potrebno imerzijo. Za dokaz obstoja prave funkcije na X uporabimo izrek o obstoju razčlenitve enote na poljubnem odprtem pokritju X.

8 Bibliography

[1] V. Guillemin and A. Pollack, *Differential topology.* Prentice-Hall, 1974. *(Cited on pages 1, 11, and 27.)*

[2] J. Milnor, *Morse theory*, Annals of Mathematics Studies, Number 51, Princeton University Press, 1965. *(Cited on pages 19 and 20.)*

[3] J. Milnor, *Topology from the differentiable viewpoint*, The University Press of Virginia, Charlottesville, 1965. *(Cited on page 27.)*

[4] M. Spivak, *Calculus on manifolds*, Addison-Wesley, 1965. *(Cited on pages 2, 3, and 5.)*

[5] В. В. Прасолов, *Элементы комбинаторной и дифференциальной топологии*, Москва, Издательство МЦНМО, 2004. *(Cited on page 27.)*