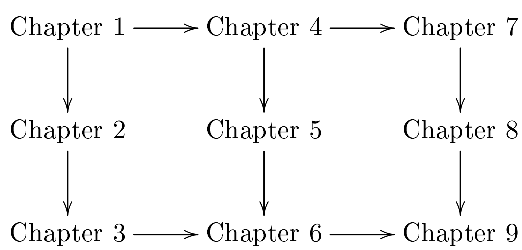
Morse Theory

M. M. Postnikov

Contents

Chapter dependency scheme

(Non-essential connections of a random nature are not reflected in the diagramme.)



Preface

A word from the transcriber

This is an English translation of "Morse Theory" by M. M. Postnikov originally written in Russian.

The preface by the original authour

When studying Morse theory, the main difficulty for beginners is the "synthetic" nature of this theory, i.e. the fact that it is located at the junction of at least three areas of mathematics (topology, analysis and geometry). The purpose of this book is to help the beginner overcome this difficulty. Actually, only a smaller part of the book is devoted to Morse theory: most of it is occupied with presenting the necessary information from topology and geometry.

The first three chapters are devoted to topology.

Since general topology is described in many Russian textbooks, the first chapter (general methodological issues) is written rather concisely and with special emphasis on facts that are usually not covered in textbooks. Some of the issues discussed in this chapter may also be of interest to a specialist.

The second chapter is entirely devoted to homotopy equivalence of topological spaces. As far as the author knows, this material (which is well known to specialists) has not been put together by anyone yet.

The third chapter discusses cellular decompositions. It seems surprising that, despite the main role that cellular decompositions play in modern topology, a coherent presentation of their theory has not yet been published anywhere.

The next two chapters are devoted to the theory of smooth manifolds. Although there are a number of expositions of this theory in Russian, none of them is suitable for our purposes. We construct a theory of smooth manifolds, following Chevalley, in a non-coordinate form; local coordinates are used only when it seems appropriate. With great regret, the author had to limit himself to finite-dimensional manifolds only - the coverage of infinite-dimensional manifolds would violate the elementary nature of the presentation. This is all the more regrettable because (as it has become clear recently) it is infinite-dimensional manifolds that represent the natural field of construction of Morse theory.

The sixth chapter, devoted to the theory of critical points of smooth functions, also belongs to the theory of smooth manifolds. The novelty here is the systematic consideration of not only non-degenerate critical points, but also non-degenerate critical manifolds. Morse numbers connecting the numbers of critical points of a given index with the Betti numbers of the manifold. This is the only place in the book (except for the directly related first half of paragraph 6 of Chapter 9) where we go beyond the topological material described in the first three chapters. However, all the necessary properties of Betty numbers in this supplement are clearly formulated.

The seventh chapter is devoted to the geometry of affine connectivity spaces- and Riemannian geometry. The presentation here is conducted mainly in the "classical" spirit, but in compliance with all modern requirements of rigour and from a global point of view. We are dealing here only with the very basics of Riemannian geometry, so much that is usually given in geometry courses remains outside the scope of our presentation. On the other hand, we had to include here some essential facts of Riemannian geometry (Whitehead's theorem on the existence of neighbourhoods normal with respect to any point, and the Hopf-Rinov theorem on complete Riemannian spaces), which are usually not considered in standard Riemannian geometry courses. Although we conduct the entire presentation of Riemannian geometry in the spirit of Cartan's ideas, nevertheless, due to the limitations of the tasks set, we managed to do without external differentiation (although in one place, namely, when deriving Cartan's basic equations, the concept of the external differential, although implicitly, still essentially appears).

The eighth chapter outlines the theory of the so-called "index form". We present here, adhering mainly to the original Morse construction, and only in the last paragraph we turn to a more modern interpretation related to the replacement of the length functional with the action functional. This, of course, somewhat lengthens and complicates the construction, but at the same time it is possible to preserve both the historical perspective and geometric clarity. In Appendix to this chapter, the "moving end" problem is treated in a similar way. Here, Morse's initial reasoning was significantly simplified.

In the ninth and final chapter, the main theorem of Morse theory is proved, describing the structure of the space of curves connecting two given points of the complete Riemannian space. In essence, this theorem is a fairly simple reworking of the main results of the previous chapter, with the aim of giving these results a more invariant appearance. In conclusion of this chapter, we give the simplest applications of Morse theory to the topology and geometry of Riemannian spaces.

A more detailed understanding of the contents of the book can be obtained from the table of contents and short summaries given to each chapter.

Formally, the reader is not required to have any knowledge beyond the first-year program of mathematical departments of universities and pedagogical colleges, although, of course, a certain level of mathematical culture and the ability to work with the book is assumed.

When working on the manuscript of this book, the authour allowed himself to make extensive use of his previous book (see the list of references of Postnikov [9]). In particular, chapters seven and eight reproduce the corresponding chapters of this book almost verbatim.

The list of references at the end of the book is provided with a "Historical and literary commentary" aimed at helping the reader to navigate the literature more easily. The list of references does not claim to be complete.

Addition during proofreading. The manuscript of this book was completed in 1965, and its printing was delayed. At present, the author would present many things in a completely different way, from a more modern perspective. But this would have postponed the publication of the book indefinitely, and therefore it was decided to leave it in its original form.

The Authour

Chapter 1

Necessary information from general topology

In this introductory chapter we present the information we need based on general set-theoretic topology. As a rule, we do not seek to exhaust a particular issue raised. Only questions for which there is no coherent presentation in the literature known to the authour are considered in detail.

§1.1 sets out the definition of topological spaces, introducing the classes of spaces distinguished by the axioms of countability and separability are defined, and the most important operations on topological spaces (free associations, topological sums and topological products) are described.

In §1.2, compact and closed spaces are considered (paracompact, finally compact, locally compact, etc.). The results of this point are mostly known, but so far were not collected together anywhere.

§1.3 presents classical theorems on continuous functions (the theorems of Urysohn and Tietze and the theorem on the existence of a partition of unity).

In §1.4, the presentation is concentrated mainly around Stone's theorem on the paracompactness of metric spaces. As a consequence of this theorem, in particular, we prove (based on the general results of §1.2 the well-known theorem of P. S. Aleksandrov on the separability of connected locally compact metric spaces.

In §1.5, the main attention is paid to the issue of constructing continuous maps. In this regard, identification maps and, in particular, their topological products are considered in detail.

§1.6 contains rather heterogeneous material, grouped around the concept of topology of identification. In particular, this section describes the construction of a cylinder of an arbitrary continuous map.

1.1 Topological spaces

Since the basic facts of general topology are widely known, in this chapter we only recall some concepts and clarify terminology.

A set X is said to be defined as a *topological space* or to have a *topology* introduced if some of its subsets are declared open and the following axioms hold:

1) the set X itself is open;

2) the empty set $\varnothing$ is open;

3) the union of any number of open sets is open;

4) the intersection of any finite number of open sets is open.

Let A be an arbitrary subset of the topological space X. Every open set U containing A is called its *neighbourhood*. A point a of a set A is called its *inner point* if some of its neighbourhood is contained in the set A. The set $\operatorname{int} A$ of all interior points of a set A is called its *interior*. The interior of $\operatorname{int} A$ is the largest open set contained in the set A. Therefore $\operatorname{int}(\operatorname{int} A) = \operatorname{int} A$. A set A is open if and only if $\operatorname{int} A = A$, that is, when any of its points is its interior point. If $A \subset B$, then $\operatorname{int} A \subset \operatorname{int} B$. For any family $\{A_\alpha\}$ of subsets of the space X the following inclusions occur

$$\operatorname{int} \cup_\alpha A_\alpha \supset \cup_\alpha \operatorname{int} A_\alpha, \quad \operatorname{int} \cap_\alpha A_\alpha \subset \cap_\alpha \operatorname{int} A_\alpha,$$

and for a finite family $\{A_\alpha\}$ equality takes place in the second of these relations.

A point x of a topological space X is called *isolated* if it, considered as a subset of the space X, is its open subset. A topological space in which all points are isolated is called *discrete*. The space X is discrete if and only if each of its subsets is open.

A family $\{U_\alpha\}$ of neighbourhoods U_α of a set A is called its *fundamental system of neighbourhoods* if for an arbitrary neighbourhood U of the set A there exists an α such that $U_\alpha \subset U$. A topological space X is called a *space of countable local weight*[1] if any of its points has a countable fundamental system of neighbourhoods. In this case we also say that the space X satisfies the *first axiom of countability*.

A *base* (or a *base of open sets*) of a topological space X is any family of its open sets that has the property that the unions of the sets of this family exhaust all the open sets of the space X. In order for some family of open sets of the space X to be a base, it is necessary and sufficient that for any point of the space X from this family it is possible to choose a fundamental system of neighbourhoods of this point. A family of open sets of a space X that has the property that all their possible finite intersections form a base of the space X is called its *prebase*. The assignment of some pre-base completely determines the

[1]Comment by the transcriber: The *weight* of a topological space is the smallest cardinality of an open base.

topology of the space X, and any family of subsets of the set X can be taken as a pre-base of some topology defined on this set. A space X is called a *space with a countable base* (or a *space of countable weight*) if it has at least one base (or, equivalently, a pre-base) containing at most a countable number of sets. In this case, we also say that the space X satisfies the *second axiom of countability.* Every space with a countable base has a countable local weight.

An example of a topological space with a countable base is the number line $\mathbb{R}$, considered in the usual *Euclidean topology.* By definition, the base of this topology is the family of all possible open intervals (a, b), $a < b$. To obtain a countable base, it is sufficient to restrict ourselves to intervals (a, b) with rational a and b.

The complements (in X) of open sets of a topological space X are called its *closed* subsets. A set X is completely defined as a topological space if it is specified which of its subsets are closed. Closed sets have the following basic properties, dual properties of open sets

1) the empty set $\varnothing$ is closed;

2) the entire set X is closed;

3) the intersection of any number of closed sets is closed;

4) the union of any finite number of closed sets is closed.

A point $x \in X$ is called a *contact point* (or an *adherent point*) of the set $a \subset X$ if any of its neighbourhoods intersects with A. The set $\overline{A}$ of all contact points of the set A is called its *closure*. It coincides with $X \subset \operatorname{int}(X \subset A)$ and therefore is the smallest closed set of the space X containing the set A. In particular, the set A is closed if and only if $\overline{A} = A$, that is, when it contains all its contact points. In addition, $\overline{\overline{A}} = A$ and $\overline{A} \subset \overline{B}$ if $A \subset B$. For any family $\{A_\alpha\}$ of subsets of the space X there are inclusions

$$\overline{\cup_\alpha A_\alpha} \supset \cup_\alpha \overline{A}_\alpha; \quad \overline{\cap_\alpha A_\alpha} \subset \cap_\alpha \overline{A}_\alpha,$$

and for a finite family $\{A_\alpha\}$ the first of these relations holds equality.

If $\overline{A} = X$ then the set A is called *everywhere dense* (in X). A space X is called *separable* (or *countable dense*) if there exists a countable everywhere dense set in it. Any space with a countable base is separable.

The set $A \subset X$ is called *nowhere dense* (in X) if $\operatorname{int} \overline{A} = \varnothing$, i.e. if the complement of its closure is everywhere dense ($\overline{X \setminus \overline{A}} = X$). A set $A \subset X$ is nowhere dense if and only if any non-empty open set $U \subset X$ contains a non-empty open subset disjoint from A (or, in other words, when any non-empty open subset X the closed set $F \subset X$ is contained in a closed set different from X that contains the set A). The closure of a nowhere dense set is nowhere dense, and an arbitrary subset of a nowhere dense set is also nowhere dense. The union of a finite number of nowhere dense sets is nowhere dense. A space X is called a *space of the first category* (in the sense of Baire-Hausdorff) if it is decomposable into the union of a countable number of nowhere dense sets (here, as elsewhere in the future, by a countable set we mean a finite or countable set). Otherwise, the space X is called a *space of the second category.*

A point a of a space X is called the *limit* of the sequence $\{x_n\}$ of points $x_n \in X$, $i = 1, 2, \dots$ if for each its neighbourhood U there is an integer N such that $x_n \in U$ for any $n \geq N$. In this case, the sequence $\{x_n\}$ is called *convergent.* Generally speaking, a convergent sequence can have several different limits. If all points x_n of the convergent sequence $\{x_n\}$ belong to the set $A \subset X$, then each limit a of this sequence belongs to the closure $\overline{A}$ of the set A. If the space X has a countable local weight, then the converse is also true, i.e.,

Proposition 1.1. *for an arbitrary subset $A \subset X$, each point $a \in A$ is the limit of a certain sequence $\{x_n\}$ of points from A.*

Proof. Indeed, let $\{U_n\}, n \geq 1$ be a countable fundamental system of neighbourhoods of a point a. Let us construct the sequence $\{x_n\}$, taking as a point x_n, $i = 1, 2, \dots$ an arbitrary point of the set A belonging to he intersection $V_n = \cap_{n=1}^{n} U_i$ (such a point necessarily exists, because $a \in \overline{A}$, and the intersection V_n is a neighbourhood of the point a). It is clear that the sequence $\{x_n\}$ constructed in this way converges to the point a. □

Each subset A of the topological space X is defined as a topological space if its open sets are taken to be intersections with A of the open sets of the space X. This topology of the set A is called the *induced topology*, and the set A itself, equipped with an induced topology, is a subspace of the space X. Any subspace B of the subspace A is a subspace of the space X. The closed sets of the subspace A are the intersections with A of the closed sets of the space X. A subspace A is open (resp. closed) in the space X if and only if any open (resp. closed) subset of it is open (resp. closed) in the space X. Each open subspace of a space of the first category is itself a space of the first category. Therefore, if the space X contains an open subspace that is a space of the second category, then it is itself a space of the second category. Each subspace of a space with a countable base (or a space of countable local weight) is a space with a countable base (or, respectively, a space of countable local weight). Every open subspace of a separable space is separable.

In particular, each subset of the number line $\mathbb{R}$, for example the unit segment $I = [0, 1]$, is a topological space with a countable base.

A topological space X is called *Hausdorff* if any two of its different points have disjoint neighbourhoods, and *completely Hausdorff* if any two of its different points have neighbourhoods with disjoint closures. Thus, every completely Hausdorff space is Hausdorff. Any point of a Hausdorff space is closed, i.e. it is a closed subset of it. Each convergent sequence of points in a Hausdorff space has a single limit. Any subspace of a Hausdorff (completely Hausdorff) space is Hausdorff (completely Hausdorff).

A Hausdorff space is called *regular* if any of its closed subsets and any point not belonging to this subset have disjoint neighbourhoods, and *normal* if any of its two disjoint closed subsets have disjoint neighbourhoods. Formally, a stronger condition for the existence of neighbourhoods with disjoint closures leads - in contrast to the case of Hausdorff property- to the same class of spaces. Thus, for example, in normal space, any two disjoint closed sets have neighbourhoods with disjoint closures. Any normal space is regular. Any regular space is completely Hausdorff. A Hausdorff space is regular (respectively normal) if and only if for any neighbourhood U of its arbitrary point (an arbitrary closed subset) there exists a neighbourhood V such that $\overline{V} \subset U$. Any subspace of a regular space is regular and any closed subspace of a normal space is normal. Every discrete space is normal. Any subspace of the number line $\mathbb{R}$, in particular, the unit segment I, is a normal space.

Let $\{X_\alpha; \alpha \in A\}$ be a family of subspaces X_α of the topological space X such that

$$X = \cup_{\alpha \in A} X_\alpha.$$

According to the definition of a subspace, for every closed (open) set $A \subset X$ and any $\alpha \in A$, the intersection $A \cap X_\alpha$ is closed (open) in the subspace X_α. In the case when the converse is true, i.e., when any set $A \subset X$ for which all sets $A \cap X_\alpha$ are closed (open) in the corresponding subspaces X_α, is itself closed (open) in the space X, we will say that the space x is a *free union of subspace* X_α. As is easy to see,

Proposition 1.2. *any closed or open subspace A of a free union X of subspaces X_α is a free union of subspaces $A \cap X_\alpha$.*

In addition,

Proposition 1.3. *the space $X = \cup_{\alpha \in A} X_\alpha$ is a free union of its subspaces X_α, $\alpha \in A$ if*

1) all subspaces X_α are open; or

2) all subspaces X_α are closed and their number is finite.

Proof. Indeed, if all subspaces X_α are open (closed), then every set that is open (closed) in one or another of the subspaces X_α is open (closed) throughout the space X. In particular, if for a set $A \subset X$ any of the intersections $A \cap X_\alpha)$ is open (closed) in the corresponding subspace X_α, then all these intersections are open (closed) in X, and therefore the set

$$A = \cup_{\alpha \in A}(A \cap X_\alpha)$$

is the union of open closed sets $A \cap X_\alpha$ in X. Consequently, A itself is open (closed, because the number of terms, by condition, is finite). □

A family $\{A_\alpha; \alpha \in A\}$ of subsets A_α of a space X will be called *locally finite* if any point $x \in X$ has a neighbourhood intersecting only with a finite number of subsets A_α. It turns out that in condition 2) of the previous proposition, the requirement of finiteness of the family $\{X_\alpha; \alpha \in A\}$ can be weakened to the requirement of local finiteness, i.e.

Proposition 1.4. *the space $X = \cup_\alpha X_\alpha$ is a free union of its subspaces X_α, $\alpha \in A$ if all these subspaces are closed and the family $\{X_\alpha; \alpha \in A\}$ is locally finite.*

Since the family of $\{A \cap X_\alpha; \alpha \in A\}$ is locally finite together with the family $\{X_\alpha; \alpha \in A\}$, it immediately follows from the above reasoning that it is enough for us to prove only the following (which has an independent interest) proposition:

Proposition 1.5. *The union $F = \cup_\alpha F_\alpha$ of any locally finite family $\{F_\alpha; \alpha \in A\}$ of closed sets $F_\alpha \subset X$ is closed.*

We will prove even more, namely that

Proposition 1.6. *for any locally finite family $\{A_\alpha; \alpha \in A\}$ of subsets of an arbitrary space X the foollowing equality holds*

$$\overline{\cup_\alpha A_\alpha} = \cup_\alpha \overline{A}_\alpha.$$

Proof. Indeed, since $\overline{\cup_\alpha A_\alpha} \supset \cup_\alpha \overline{A}_\alpha$, proof requires only the reverse inclusion

$$\overline{\cup_\alpha A_\alpha} \subset \cup_\alpha \overline{A}_\alpha. \tag{1.7}$$

Let $A = \cup_\alpha A_\alpha$ and let $x/in\overline{A}$. Due to the local finiteness of the family $\{A_\alpha$, the point x has a neighbourhood U_0 intersecting only with a finite number of sets A_{α_1}, ..., $A_{\alpha n}$ of this family. Therefore, for any neighbourhood U of the point x, its neighbourhood $V = U \cap U_0$ does not intersect with the sets A_α for $\alpha \neq \alpha_1, \ldots, \alpha_n$. But $x \in \overline{A}$ and therefore $V \cap A \neq \emptyset$. Therefore, $V \cap \cup_{i=1}^n A_{\alpha_i} \neq \emptyset$, i.e, $U \cap \cup_{i=1}^n A_{\alpha_i} \neq \emptyset$. Due to the arbitrariness of the neighbourhood U, it follows that $x \in \overline{\cup_i A_{\alpha_i}} = \cup_{i=0}^n \overline{A}_{\alpha_i} \subset \cup_\alpha \overline{A}_\alpha$. Thus, the inclusion (1.7) is proved. □

In connection with the proven proposition, it is useful to note that

Proposition 1.8. *a family $\{A_\alpha\}$ of arbitrary sets is locally finite if and only if the family $\{\overline{A}_\alpha\}$ consisting of the closures $\overline{A}_\alpha$ of sets A_α is locally finite.*

Proof. Indeed, it is clear that an open set U intersects some set A if and only if it intersects the closure $\overline{A}$ of this set. □

The concept of free union that we have considered had, so to speak, an "internal" character: it related to subspaces of a given, "already existing," topological space X. It can also be considered when the space X is not given in advance.

Let an arbitrary set X be represented as a union

$$X = \cup_{\alpha \in A} X_\alpha$$

of sets X_α, each of which is a topological space. It is easy to see that, by declaring closed (open) sets those and only those sets $A \subset X$ for which, for any $\alpha \in A$, the intersection $A \cap X_\alpha$ is closed (open) in the space X_α, we define the set X as a topological space. However, generally speaking, the spaces X_α will not be subspaces of this space. (To get an appropriate example, it is enough to consider the case when all the spaces X_α coincide as sets with the space X, but are equipped with different topologies.) In the case when each space X_α is a subspace of the constructed space X, we will say that the spaces X_α are *freely united*, and we will call the space X a *free union of the spaces* X_α. It will be a free union in the previously defined sense of its subspaces X_α.

It is clear that

Proposition 1.9. *if all intersections* $X_{\alpha_1} \cap X_{\alpha_2}$, $\alpha_1, \alpha_2 \in A$, *are closed (open) spaces of each spaces* X_{α_1} *and* X_{α_2}, *then the spaces* X_α *are freely united and are closed (respectively, open) spaces of their free union* X.

In particular, the spaces X_α are freely combined if they do not intersect in pairs. We will call the free union of disjoint topological spaces a *topological sum*. Each term X_α of the topological sum X is closed and simultaneously open in this sum. Conversely, if

$$X = \cup_{\alpha \in A} X_\alpha$$

and if each subspace of X_α is both closed and open in X, then the space X is the topological sum of the subspaces of X_α. If each term of a topological sum is Hausdorff, completely Hausdorff, regular or normal, then the topological sum also has the same property. The same remark applies to topological sums, the terms of which have a countable local weight. For spaces with a countable base and separable, a similar statement is true only when the number of terms of the topological sum is countable.

Along with the topological sum, we will also consider the *topological product* $X \times Y$ of any two (not necessarily disjoint) topological spaces X and Y. The points of this product are, by definition, all possible pairs of the form (x, y), where $x \in X$, $y \in Y$, and the base of its open sets is the collection of all subsets of the form $U \times V$, where U is an arbitrary open subset of the space X, and V is an arbitrary open subset of the space Y (the subset $U \times V$ consists, as its notation suggests, of all pairs (x, y), where $x \in U$, $y \in V$). The topological product of any (possibly infinite) number of topological spaces X_α, $a \in A$ is defined similarly (in this case, in the case of an infinite number of factors, when constructing open sets of the base of the product, one should choose open sets that coincide with the entire space in all but a finite number of factors). The topological product of any number of Hausdorff, completely Hausdorff, or regular spaces is, respectively, Hausdorff, completely Hausdorff, and regular. For spaces of countable local weight of spaces with a countable base or separable spaces, the corresponding statement holds if the number of factors is countable. The topological product of normal spaces is not, generally speaking, a normal space.

For example, there is a normal space X whose topological product $X \times I$ on the segment $I = [0, 1]$ of the number line $\mathbb{R}$ is no longer a normal space. The normal space X, for which the product $X \times I$ is also normal, we will call it *stably normal*.

1.2 Compact and some other similar spaces

A family $\{A_\alpha; \alpha \in A\}$ of subsets A_α of a topological space X is called a *covering* of the set $B \subset X$ if

$$B \subset \cup_{\alpha \in A} A_\alpha.$$

In particular (for $B = X$), a family $\{A_\alpha\}$ is called a *covering of the space* X if

$$X = \cup_{\alpha \in A} A_\alpha.$$

A covering $\{A_\alpha; \alpha \in A\}$ is called *open* if all sets A_α are open, and *closed* if all sets A_α are closed. In the future, we usually consider only open coverings.

The covering $\{A_\alpha; \alpha \in A\}$ of the space X is called *point finite* if any point $x \in X$ belongs to only a finite number of sets A_α, *locally finite* if any point $x \in X$ has a neighbourhood intersecting only with a finite number of sets A_α (cf. §refsect:1-1), and is *stellar finite* if any set A_α intersects only a finite number of other sets of covering. Any locally finite cover is point finite and any open stellar-finite cover is locally finite.

We will call the covering $\{B_\beta; \beta \in B\}$ a *subcovering* of the covering $\{A_\alpha; \alpha \in A\}$ if $B \subset A$ and $B_\beta = A_\beta$ for any $\beta \in B$. We will say that the cover $\{B_\beta; \beta \in B\}$ is a *refinement* of the cover $\{A_\alpha; \alpha \in A\}$ if for any $\beta \in B$ there exists an $\alpha \in A$ such that $B_\beta \subset A\alpha$. It is clear that any subcovering of the covering $\{A_\alpha; \alpha \in A\}$ is a refinement of this covering.

A space X is called *paracompact* if any of its open coverings can be refined into a locally finite open covering. Every closed subspace of a paracompact space is paracompact. The topological sum of any number of paracompact spaces is paracompact. Each discrete space is paracompact.

To prove deeper properties of paracompact spaces, it is useful to first prove that:

Proposition 1.10. *for a subset A and a closed subset F of a paracompact space X that does not intersect with it to have disjoint open neighbourhoods, it is sufficient that the set F have an open cover $\{U_\alpha; \alpha \in A\}$ such that $\overline{U}_\alpha \cap A = \emptyset$ for each $\alpha \in A$.*

Proof. Indeed, let us consider the covering $\{X \setminus F, U_\alpha\}$ of the entire space X. Let $\{W_\beta; \beta \in B\}$ be a locally finite covering refining the covering $\{X \setminus F, U_\alpha\}$, and let B' be the set of all indices $\beta \in B$ for which $W_\beta \cap F \neq \emptyset$. Since $W_\beta \not\subset X \setminus F$, for each index $\beta \in B'$ there exists an index $\alpha \in A$ such that $W_\beta \subset U_\alpha$ and therefore $\overline{W}_\beta \cap A = \emptyset$ (since $\overline{W}_\beta \subset \overline{U}_\alpha$). Therefore, the set $\cup_{\beta \in B'} \overline{W}_\beta$ does not intersect with the set A. But $\cup_{\beta \in B'} \overline{W}_\beta = \overline{\cup_{\beta \in B'} W_\beta}$ due to the local finiteness of the family $\{W_\beta\}$. Therefore, the set $U = X \setminus \cup_{\beta \in B'} \overline{W}_\beta$ is open and contains the set A, i.e. it is an open neighbourhood of the set A. To complete the proof, it remains to note that the set $W = \cup_{\beta \in B'} W_\beta$ is open, contains the set F and does not intersect with the set U. □

It follows easily from this proposition, in particular, that

Proposition 1.11. *any Hausdorff paracompact space X is normal.*

Proof. Indeed, let F_1 and F_2 be disjoint closed subsets of the space X. Since the space X is Hausdorff, for any two points $x_1 \in F_1$ and $x_2 \in F_2$ there are disjoint neighbourhoods $U(x_1)$ and $U(x_2)$. In particular, the point x_2 will not belong to the closure $\overline{U(x_1)}$ of the neighbourhood $U(x_1)$ of thee point x_1. In other words, the covering of the set F_1, consisting of the neighbourhoods $U(x_1$, $x_1 \in F_1$, satisfies the conditions of the proposition just proved (for $F = F_1$ and A consisting of the point x_2). Consequently, there are disjoint open sets V and $V(x_2$, respectively containing the set F_1 and the point x_2. In particular, the closure $\overline{V}(x_2)$ of the set $V(x_2)$ does not intersect with the set F_1. Therefore, the covering of the set F_2, consisting of the sets $V(X_2)$, $x_2 \in F_2$, also satisfies the conditions of our proposition (for $F = F_2$ and $A = F_1$). Consequently, the sets F_1 and F_2 can be separated by disjoint neighbourhoods. □

A space X (or, more generally, some subset of it) is called *compact* (respectively, *Lindelöf*) if any of its open covers can be refined with a finite (respectively, countable) open cover, or, equivalently, if any of its open covers can be subdivided with a finite (respectively, countable) cover.

Each compact space is Lindelöf. Any discrete compact (Lindelöf) space is finite (countable). Any finite (countable) space is compact (Lindelöf). A subset of an arbitrary space is compact (Lindelöf) if and only if it is compact (Lindelöf) in the induced topology. In a Hausdorff space X, the set of points of some convergent sequence together with the limit of this sequence constitutes a compact subset of the space X. The closed subspace A of a compact (Lindelöf) space X is compact (Lindelöf). In particular, the intersection of a compact set with a closed one is compact.

The classical Heine-Borel lemma states that any closed bounded subset of the real line $\mathbb{R}$ is compact. In particular, the interval $I = [0, 1]$ is compact.

Below we prove that in regular (and even Hausdorff) spaces all compact sets are closed. In this regard, it is useful to note that compact sets of regular spaces have a property that any closed sets have in a normal space, namely, for any neighbourhood U of a compact subset C of a regular space X, there exists a neighbourhood V such that $\overline{V} \subset U$.

Proof. Indeed, since the space X is regular, each point $x \in C$ has a neighbourhood V_x such that $\overline{V}_x \subset U$. Since the set C is compact, from its covering $\{V_x, x \in C\}$ one can choose a finite subcovering $V_{x_1}, \ldots, V_{x_n}$. The set

$$V = \cup_{i=1}^{n} V_{x_i}$$

is open and its closure $\overline{V}$ (which is the union of the closures $\overline{V}_{x_i}$ of the sets V_{x_i}) is contained in the neighbourhood U. □

A topological space X will be called *sequentially compact* if each sequence of its points contains a convergent subsequence. It is easy to see that

Proposition 1.12. *any Hausdorff compact space of countable weight is sequentially compact.*

Proof. Indeed, let this not be so, i.e. let there exist a compact space of countable local weight, containing a sequence $\{x_n\}$, with no subsequence of which converges. It is easy to see that the set of points of this sequence is, firstly, infinite, secondly, discrete, and, thirdly, closed. On the other hand, any closed discrete subset of a compact space is necessarily finite. □

It is equally easy to prove that

Proposition 1.13. *any Lindelöf and sequentially compact space is compact.*

Proof. To prove this statement, it suffices to show that from any countable open covering $\{U_i\}$ of a sequentially compact space X one can choose a finite subcovering. Consider the sets $V_n = U_1 \cup \cdots \cup U_n$, $n = 1, 2, \ldots$. If it is impossible to choose a finite subcovering from the covering $\{U_i\}$, then we can obviously assume that all sets V_n are distinct. Let $x_n \in V_{n+1} \setminus V_n$, $n = 1, 2, \ldots$. Since the point x_n can belong to the set U_i only when $n < i$, then none of the sets U_i contains the limit of any convergent subsequence of the sequence $\{x_n\}$. But this is impossible, because the sets Ui cover, according to the condition, the entire space X. The resulting contradiction shows that from the covering $\{U_i\}$ one can choose a finite subcovering. □

In connection with this statement, it is useful to keep in mind that

Proposition 1.14. *any space X with a countable base is Lindelöf.*

Proof. Indeed, by selecting for an arbitrary open covering Γ of the space X in a countable base of this space a subfamily of all its elements contained in the elements of the covering Γ, we obviously obtain a countable covering of the space X as a refinement in the covering Γ. □

The compactness of sequentially compact spaces follows not only from Lindelöf property, but also from paracompactness, i.e.

Proposition 1.15. *a paracompact and sequentially compact space is compact.*

To prove this statement it suffices to show that

Proposition 1.16. *any locally finite covering $\{U_\alpha\}$ of a sequentially compact space X is finite.*

Proof. Assume the contrary, i.e. assume that the covering $\{U_\alpha\}$ is infinite, and consider some of its countable subfamily $\{U_i\}$. Let, as above, $V_n = U_1 \cup \cdots \cup U_n$. From the local finiteness of the covering $\{U_\alpha\}$ it follows directly that the family $\{V_n\}$ contains infinitely many different sets Therefore, passing to some of its subfamily if necessary, we can assume that all sets V_n are distinct. Let $x_n \in V_n \setminus V_{n-1}$ and let x be the limit of some convergent subsequence of the sequence $\{x_n\}$. By definition, any neighbourhood of a point x contains infinitely many points of the sequence $\{x_n\}$ and therefore intersects with an infinite number of elements of the covering $\{U_\alpha\}$. Since this contradicts the local finiteness of the covering $\{U_\alpha\}$, the assumption that the covering $\{U_\alpha\}$ is infinite is false. □

The union of a finite (countable) number of compact (Lindelöf) subspaces of an arbitrary space is compact (Lindelöf). In particular, the union of countably many compact sets is Lindelöf. The topological sum of a finite (countable) number of compact (Lindelöf) spaces is a compact (Lindelöf) space. The topological product of any number of compact spaces is compact.

For two (and therefore any finite number of factors) this statement is obvious, since in each open covering of the product $X \times Y$ of factors one can inscribe a covering consisting of "rectangles" $U_\alpha \times V_\beta$, where U_α and V_β are elements of some coverings of the spaces X and Y, respectively.

For paracompact and Lindelöf spaces, the corresponding statement is generally false (even for the case of two factors). However, it can be shown that

Proposition 1.17. *the topological product of a paracompact and a compact space is paracompact.*

Proof. Indeed, let Γ be an arbitrary open covering of the product $X \times Y$ of a paracompact space X and a compact space Y. Without loss of generality, we can assume that the elements of the covering Γ have the form $U_\alpha \times V_\beta$ where u_α and V_β are some open coverings of the spaces X and Y, respectively. Since for each point $x \in X$ the "layer" $x \times Y$ of the product $X \times Y$ is obviously compact, from the covering Γ one can choose a finite subfamily $\{U_{\alpha_i,x} \times V_{\beta_i,x} | i = 1, 2, \ldots, n_x\}$ such that

$$x \times Y \subset \cup_{i=1}^{n_x} (U_{\alpha_i,x} \times V_{\beta_i,x}).$$

In this case, we can obviously assume that $x \in U_{\alpha_i,x}$ for all $i = 1, 2, \ldots, nx$. Let

$$U_x = \cap_{i=1}^{n_x} U_{\alpha_i,x}.$$

It is clear that the sets $U_x \times V_{\beta_i,x}$, $i = 1, 2, \ldots, n_x$, are open in $X \times Y$ and still cover the layer $x \times Y$:

$$x \times Y \subset \cup_{i=1}^{n_x} (U_x \times V_{\beta_i,x}).$$

Moreover, $x \in U_x$, so that the family $\{U_x | x \in X\}$ is an open covering of the space X. Let $\{W_\delta | \delta \in \Delta\}$ be a refined locally finite open covering of the covering $\{U_x | x \in X\}$. Having chosen for each $\delta \in \Delta$ a point $x_\delta \in X$ such that $W_\delta \in U_{x_\delta}$, we consider the sets $W_\delta \times^{V}_{\beta_i,x_\delta}$, $i = 1, 2, \ldots, n_{x_\delta}$. It is clear that

$$W_\delta \times Y \subset \cup_{i=1}^{n_{x_\delta}} (W_\delta \times V_{\beta_i,x_\delta}).$$

Consequently, the family $\{W\delta \times V_{\beta_i,x_\delta} | \delta \in \Delta, i = 1, 2, \ldots, n_{x_\delta}\}$ is an open covering of the space $X \times Y$, obviously a refinement the covering Γ. Therefore, to complete the proof, it only remains to show that this covering is locally finite. Let (x, y) be an arbitrary point of the space $X \times Y$. By the paracompactness of the space X, the point x has a neighbourhood $U(x)$ in X that intersects only a finite number of open sets W_δ. Consider a neighbourhood $U(x) \times V_{\beta_{i_0}, x_{\delta_0}}$ of a point (x, y) in the space $X \times Y$, where $V_{\beta_{i_0}, x_{\delta_0}}$ is one of the sets of the form V_{β_i, x_δ}, containing the point y. If this neighbourhood intersects some set $W_\delta \times V_{\beta_i, x_\delta}$ then $U(x) \cap W_\delta \neq \varnothing$ and therefore the number of such sets is finite. Consequently, the covering $W_\delta \times V_{\beta_{i_0}, x_{\delta_0}}$, $i = 1, 2, \ldots, n_{x_\delta}$ is locally finite. □

It immediately follows from the proved proposition that the topological product of a Hausdorff paracompact and a Hausdorff compact space is a Hausdorff paracompact and therefore a normal space. Since the segment $I = [0, 1]$ is a Hausdorff compact space, it follows, in particular, that

Proposition 1.18. *any Hausdorff paracompact space is stably normal.*

It is clear that any compact space is paracompact. Therefore,

Proposition 1.19. *any Hausdorff compact space is stably normal,*

Unlike compactness, Lindelöf-ness, generally speaking, does not ensure paracompactness. However,

Proposition 1.20. *any regular Lindelöf space X is paracompact (and hence normal).*

Proof. We will begin the proof of this statement by considering two disjoint closed sets F_1 and F_2 of the space X. Since the space X is regular, for any point $x \in X$, there exists a neighbourhood $U(x)$ such that $\overline{U}(x) \cap F_i \neq \varnothing$ if $x \notin F_i$. Since the set F_i, $i = 1, 2$, is a closed subset of the Lindelöf space X, it itself is Lindelöf and therefore from its covering $\{U(x) | x \in F_i\}$ one can choose a countable subcovering, i.e. in the set F_i there exists a countable system of points $x_{i,1}, \ldots, x_{i,n}, \ldots$ such that

$$F \subset \cup_{n=1}^{\infty} U(x_{i,n}), \quad i = 1, 2.$$

In this case $\overline{U}(x_{1,n}) \cap F_2 = \varnothing$ and similarly $\overline{U}(x_{22,n}) \cap F_1 = \varnothing$. Now we define by induction for any $n \geq 1$ the sets $V_{1,n}$ and $V_{2,n}$, setting

$$V_{1,n} = \overline{U}(x_{1,n}) \setminus \cup_{k=1}^{n} \overline{U}(x_{2,k}),$$
$$V_{2,n} = \overline{U}(x_{2,n}) \setminus \cup_{k=1}^{n} \overline{U}(x_{1,k}).$$

It is easy to see that $V_{1,n}$ and $V_{2,m}$ do not intersect. Indeed, if $n \leq m$, then

$$V_{1,n} \cap V_{2,m} \subset U(X_{1,n}) \cap (U(x_{2,m}) \setminus \overline{U}(x_{1,n}) = \varnothing,$$

and if $n > m$, then

$$V_{1,n} \cap V_{2,m} \subset (U(X_{1,n}) \setminus \overline{U}(x_{2,m}) \cap (U(x_{2,m}) = \varnothing.,$$

Consequently, the sets

$$V_1 = \cup_{n=1}^{\infty} V_{1,n}, \quad V_2 = \cup_{n=1}^{\infty} V_{2,n}$$

also do not intersect. On the other hand, they obviously contain the sets F_1 and F_2, respectively. Thus, we have proved that any two disjoint closed sets F_1 and F_2 of the space X can be enclosed in disjoint open sets V_1 and v_2. In other words, we have proved that the space X is normal. □

Now let Γ be an arbitrary open covering of X. Since X is regular by assumption, each point $x \in X$ has a neighbourhood $U(x)$ whose closure $\overline{U}(x)$ is contained in some element of Γ. Since X is also Lindelöf, there exists a countable system of points $x_1, \ldots, x_k, \ldots$ in it such that

$$X = \cup_{n=1}^{\infty} U(x_k).$$

Let U_k be an element of the covering Γ containing the set $\overline{U}(x_k)$. Using the already proven normality of the space X, for any $n \geq 1$ we can construct by induction an open set U_k^n such that

$$\overline{U}(x_k) \subset U_k^n \subset U_k \quad \text{and} \quad \overline{U}_k^n \subset U_k^{n+1}, \quad n, k = 1, 2, \ldots$$

For each $n \geq 1$ we now put

$$V^n = \cup_{k=1}^{n} U_k^n.$$

It is clear that $\overline{V}^n \subset V^{n+1}$ and that the sets V^n form an open covering of the space X. Consider the sets

$$H^n = V^n \setminus \overline{V}^{n-2}, \quad n = 1, 2, \ldots$$

(for $m \leq 0$ we conditionally assume that $V^m = \varnothing$). It is easy to see that the (obviously open) sets H^n form a covering of the space X. (Indeed, for any point $x \in X$ there exists an n such that $x \in V^n$ and $x \notin V^{n-1}$. But then $x \notin \overline{V}^{n-2}$ and therefore $x \in H^n$.) In addition, it is clear that this covering is star-finite (since $H^{n_1} \cap H^{n_2} = \varnothing$ for $n_2 \geq n_1 + 2$).

Let $k = 1, \ldots, n$. Put

$$W_k^n = H^n \cap U_k^n = U_k^n \setminus \overline{V}^{n-2}, \quad n = 1, 2, \ldots$$

Since $H^n \subset V^n = \cup_{k=1}^{n} U_k^n$, then $\cup_{k=1}^{n} W_k^n = H^n$. Consequently, the sets W_k^n constitute an (open) covering of the space X. This covering is a refinement of the covering Γ (since $W_k^n \subset U_k^n \subset U_k$) and is star-finite (since it is obtained from the star-finite covering $\{H^n\}$ by decomposing each of its elements into a finite number of sets). Since any star-finite open covering is locally finite, it is thus proved that any open covering of X can be refined into a locally finite open covering, i.e., that X is paracompact.

Let us now prove the fact mentioned above, namely, that

Proposition 1.21. *any compact subset C of a Hausdorff space X is closed,*

Proof. To this end, for any point $x \in X \setminus C$ and any point $c \in C$, we choose some disjoint neighbourhoods $U_c(x)$ and $U_x(c)$. For each fixed point $x \in X \setminus C$, the sets $U_x(c)$ $c \in C$ obviously form an open covering of the subset C. Therefore, there exists a finite system of points $c_1, \ldots, c_n \in C$ such that

$$C \subset \cup_{i=1}^{n} U_x(c_i).$$

Let

$$U(x) = \cup_{i=1}^{n} U_{c_i}(x).$$

The set $U(x)$ is open, contains the point x and does not intersect the set C. Thus, each point x of the set $X \setminus C$ is its interior point, i.e., this set is open. Therefore, the set C itself is closed. □

A subset A of a topological space X will be called *compactly closed* if its intersection with any compact subset $C \subset X$ is closed (in X). From the proposition just proved it follows immediately that

Proposition 1.22. *any closed subset of a Hausdorff space X is compactly closed.*

Thus, for Hausdorff spaces, the classes of compact, closed, and compact-closed subsets are related by a simple inclusion relation: each of these classes is wider than the previous one.

We will call a space X a *compactly generated*[2] space if it is a free union of all its compact subspaces, i.e. if its subset A is closed if and only if for any compact subset $C \subset X$ the intersection $A \cap C$ is closed in C. In a compactly generated space, any compactly closed set is closed. According to what was said above, if X is a Hausdorff space, then the converse is also true. Thus,

Proposition 1.23. *A Hausdorff space X is said to be a compactly generated space if and only if any of its compactly closed subsets is closed.*

Further, it is easy to see that

Proposition 1.24. *any closed (resp. open) subspace Y of a compactly generated (resp. compactly generated and regular) space X is also a compactly generated space.*

Proof. Indeed, let the set $A \subset Y$ have the property that for any compact set $C \subset Y$ the intersection $A \cap C$ is closed in C. We need to prove that then A is closed in Y. If Y is closed, then instead we will prove that A is closed in X. Since X is a compactly generated space, it suffices to prove that for any compact set $C \subset X$ the intersection $A \cap C$ is still closed in C. But this intersection coincides with the intersection $A \cap (Y \cap C)$, and the set $Y \cap C$, being the intersection of a compact and closed set, is compact. Therefore, by hypothesis, the intersection $A \cap (Y \cap C)$ is closed in C. □

[2]Transcriber's note: Postnikov has coined the word "kaonic."

Now let the subspace Y be open. Let us consider an arbitrary contact point a of the set A belonging to the subspace Y. Since the space X is, by assumption, regular, then this point has a neighbourhood U such that $\overline{U} \subset Y$. Since for any compact set $C \subset X$ the intersection $\overline{U} \cap C \subset Y$ is compact, the intersection $\overline{U} \cap C \cap A$ is closed in $\overline{U} \cap C$ and therefore in C. Consequently, the intersection $\overline{U} \cap A$ is closed in the space X. But it is clear that $a \in \overline{\overline{U} \cap A}$. Thus, $a \in \overline{U} \cap A$ and, therefore, $a \in A$. We have thus proved that each contact point of the set A belonging to the subspace Y lies in A. But this means that the set A is closed in the subspace Y.

It is clear that all compact spaces are compactly generated spaces. Moreover, it is easy to see that

Proposition 1.25. *any Hausdorff space X with countable local weight is compactly generated.*

Proof. Indeed, let a be an arbitrary contact point of some compactly closed subset A of the space X and let $\{a_n\}$ be an arbitrary sequence of points of the set A converging to the point a. Since this sequence together with its limit a forms a compact set, the intersection of this set with the set A is closed. But for $a \notin A$ this intersection would coincide with the sequence $\{a_n\}$ and would be obviously not closed (since the limit of this sequence $a \notin A$ would not belong to it). Consequently, $a \in A$. Thus, $\overline{A} = A$, i.e. the set A is closed. □

We will call a space X *countably compact* if it is a free union of some countable family $\{X_n, n \geq 0\}$ of its compact subspaces X_n. Without loss of generality, we can assume that $X_n \subset X_{n+1}$ for all $n \geq 0$. Every countably compact space is a compactly generated space and Lindelöf space. Any closed subspace of a countably compact space is countably compact. The topological sum of a countable number of countably compact spaces is a countably compact space. Moreover,

Proposition 1.26. *any Hausdorff countably compact space X is paracompact and, in particular, normal.*

Since the space X is Lindelöf, it suffices to establish its regularity. However, it will be more convenient for us to immediately prove its normality.

Let $F^{(1)}$ and $F^{(2)}$ be arbitrary disjoint closed subsets of X. Let us consider compact (and therefore closed and normal) subspaces

$$X_0 \subset X_1 \subset \cdots \subset X_n \subset X_{n+1} \subset \cdots,$$

whose free union is the space X. It turns out that for any $n \geq 0$ there exist open (in X_n) sets $U_n^{(i)}$, $i = 1, 2$ such that

$$X_n \cap F^{(i)} \subset U_n^{(i)}, \quad i = 1, 2,$$
$$\overline{U}_n^{(1)} \cap \overline{U}_n^{(2)} = \varnothing \quad \text{and (for } n > 0) \quad U_n^{(i)} \cap X_{n-1} = U_{n-1}^{(i)}.$$

For $n = 0$, the existence of such sets is ensured by the normality of the space X_0. Let for some $n \geq 0$ the sets $U_n^{(i)}$, $i = 1, 2$ have already been constructed. Let us consider in the subspace X_{n+1} the sets $(X_n \cap F^{(i)}) \cup \overline{U}_n^{(i)}$ $i = 1, 2$. These sets are closed and their intersection is empty. Therefore, they have neighbourhoods $V^{(1)}$ and $V^{(2)}$ with disjoint closures $\overline{V}^{(1)}$ and $\overline{V}^{(2)}$. On the other hand, since the sets $U_n^{(1)}$ and $U_n^{(2)}$ are open in the subspace X_n of the space X_{n+1}, then in the space X_{n+1} there exist open sets $W^{(1)}$ and $W^{(2)}$ such that

$$W^{(i)} \cap X_n = U_n^{(i)}, \quad i = 1, 2.$$

We will put

$$U_{n+1}^{(i)} = V^{(i)} \cap (W^{(i)} \cup (X_{n+1} \setminus X_n)), \quad i = 1, 2.$$

The sets $U_{n+1}^{(i)}$, $i = 1, 2$ constructed in this way obviously possess all the necessary properties.

Let us now consider the sets

$$U^{(i))} = \cup_{n=0}^{\infty} U_n^{(i)}, \quad i = 1, 2.$$

Since for any $n \geq 0$ the sets $U_n^{(i)} = U^{(i)} \cap X_n$ are open in the subspace X_n, the sets $U^{(i)}$, $i = 1, 2$ are open in the space X. In addition, they do not intersect and contain the sets $F^{(1)}$ and $F^{(2)}$, respectively. Thus, any two non-intersecting closed sets $F^{(1)}$ and $F^{(2)}$ of the space X have non-intersecting neighbourhoods $U^{(1)}$ and $U^{(2)}$. But this means that the space X is normal.

Remark 1.27. In the above proof we used only the normality and closedness of the subspaces X_n.

Therefore,

Proposition 1.28. *any space that is a free union of a skeletal increasing sequence of closed normal subspaces is itself a normal space.*

A space X is called *locally compact* (resp. emphlocally countably compact) if each of its points has a neighbourhood whose closure is compact (resp. countably compact). Any compact (resp. countably compact) space is locally compact (resp. locally countably compact). Any discrete space is locally compact. For each compact subset of a locally compact (resp. locally countably compact) space, there exists a neighbourhood whose closure is compact (resp. countably compact). Every Hausdorff locally countably compact and, in particular, locally compact space is regular (since a Hausdorff countably compact space is normal). Any closed subspace of a locally compact (resp. locally countably compact) space X is locally compact (resp. locally countably compact). The corresponding assertion for open subspaces is also true only if X is Hausdorff. In particular, any open subspace of a Hausdorff compact space is locally compact.

Theorem 1.29. *Conversely, any Hausdorff locally compact space X is an open subspace of some Hausdorff compact space X' and the space X' can be constructed so that the "growth" of $X \setminus X'$ consists of only one point (P. S. Alexandrov's theorem on one-point compactification).*

Proof. Indeed, let ω be an arbitrary point not belonging to the space X. Let us introduce a topology into the set $X^{'} = X \cup \omega$, considering its open sets to be those and only those sets $A^{'} \subset X^{'}$, for which the set $A = A^{'} \setminus \omega$ is open in the space X and - in the case when to $\omega \in A^{'}$ - is, in addition, the complement (in X) of some compact set. It is easy to verify that the space $X^{'}$ is compact, Hausdorff, and the topology induced on X by the topology of the space $X^{'}$ coincides with the original topology of the space X. □

From the theorem on one-point compactification, in particular, it follows directly that

Proposition 1.30. *any Hausdorff locally compact space X is a compactly generated space.*

This statement can easily be proved directly. Indeed, let A be an arbitrary compactly closed subset of a locally compact space X. Let us prove that A is closed.

Proof. Let $x \in X \setminus A$. Since X is locally compact, x has a neighbourhood U whose closure $\overline{U}$ is compact. Since A is compactly closed, $A \cap \overline{U}$ is closed, and hence its complement $V = X \setminus (A \cap \overline{U})$ is open. Thus, x has a neighbourhood V that does not intersect $A \cap \overline{U}$. But then $x \in U \cap V \subset U \cap V \subset X \setminus A$, i.e. $x \in \text{int}(X \setminus A)$. Thus, $X \setminus A$ is open, and hence A is closed. □

The last argument is almost literally preserved for locally countable compactly generated spaces. (The closedness of the set $A \cap \overline{U}$ follows in this case from the easily proven fact that the intersection of a compactly closed set with an arbitrary countably compact subspace is closed in this subspace.) Thus,

Proposition 1.31. *any Hausdorff locally countable compact space is a compactly generated space.*

Let us now deal exclusively with locally compact spaces (mainly Hausdorff). First of all, we will prove that

Proposition 1.32. *any Hausdorff locally compact space X is a space of the second category.*

Moreover,

Proposition 1.33. *any subspace of the first category of a Hausdorff locally compact space X has the property that its complement is everywhere dense in the space X.*

Proof. Indeed, let $\{A_n; n \geq 1\}$ be an arbitrary countable family of nowhere dense subsets of X. We must prove that in any neighbourhood U of an arbitrary point x of X there exist points that do not belong to any of the sets A_n. For this purpose, we construct in X a family $\{U_n; n \geq 0\}$ of nonempty open sets U_n such that

1) the closure of the set $\overline{U}_0$ is compact and is contained in the neighbourhood U;

2) for any $n \geq 1$ the relations

$$\overline{U}_n \subset U_{n-1}, \quad U_n \cap A_n = \varnothing$$

hold.

The existence of the set U_0 follows directly from the local compactness and regularity of the space X. Let for some $n \geq 1$ the set U_{n-1} has already been constructed. Since the set A_n is nowhere dense, in the open set U_{n-1} there exists a non-empty open set V_n such that $V_n \cap A_n = \varnothing$. Since the space X is regular, there exists a non-empty open set $U_n \subset V_n$ such that $\overline{U}_n \subset V_n$. It is clear that $\overline{U}_n \subset U_{n-1}$ and $\overline{U}_n \cap A_n = \varnothing$. Thus, the sets U_n have been constructed by induction for all $n \geq 0$.

Let us now consider the sets open in the subspace $\overline{U}_0$

$$W_n = \overline{U}_0 \setminus \overline{U}_n, \quad n \geq 1.$$

If the intersection of the sets $\overline{U}_n$, $n \geq 1$, is empty, then the sets W_n form a covering of the subspace $\overline{U}_0$. Since this subspace is compact by construction, a finite subcovering $\{W_{n_1}, \ldots, W_{n_k}\}$ can be chosen from this covering. The corresponding sets $\overline{U}_{n_1}, \ldots, \overline{U}_{n_k}$ will then have an empty intersection. The latter is impossible, since this intersection is equal to $\overline{U}_m$, where $m = \max(n_1, \ldots, n_k)$. Consequently, the intersection of sets $\overline{U}_n$, $n \geq 1$ is not empty. Since each point of the last intersection does not belong to any of the sets A_n, the above proposition is completely proved. □

Now recall that any Hausdorff locally compact space is regular. Therefore, if it is also Lindelöf, then it is also paracompact. Since the topological sum of paracompact spaces is a paracompact space, it follows immediately that

Proposition 1.34. *any Hausdorff locally compact space that is a topological sum of Linelöf subspaces is paracompact.*

It turns out that the converse is also true (and even without the Hausdorff proposition), i.e.

Proposition 1.35. *any locally compact paracompact space X is a topological sum of Lindelöf subspaces.*

Proof. Indeed, due to the local compactness of the space X, any point of it has a neighbourhood with compact closure. All these neighbourhoods form a certain covering Σ of the space X. Since the space X is paracompact, a locally finite covering Γ can be inscribed in the covering Σ. This covering (like the covering Σ) has the property that the closure of each of its elements U is compact. On the other hand, due to the local finiteness of the covering Γ, each point $x \in \overline{U}$ has a neighbourhood $V(x)$ that intersects only a finite number of elements of this covering. But, since the set $\overline{U}$ is compact, there exists a finite system of points $x_1, \ldots, x_n \in \overline{U}$ such that

$$\overline{U} \subset \cup_{i=1}^{n} V(x_i).$$

Consequently, the set $\overline{U}$, and therefore the set U, intersects only a finite number of elements of the covering Γ, i.e. the covering Γ is star-finite.

Having now chosen some element $U0$ of the covering Γ, we consider for any $n \geq 1$ the subset X_n of X consisting of all points $xinX$ for which there exist n sets $U_1, \ldots, U_n \in \Gamma$ such that $x \in U_n$ and

$$U_0 \cap U_1 \neq \emptyset, \quad U_1 \cap U_2 \neq \emptyset, \; \ldots, \; U_{n-1} \cap U_n \neq \emptyset.$$

It is clear that the subset X_n is not empty, open and - by the star-like finiteness of the covering Γ - is the union of some finite number of elements of this covering. Therefore, the set

$$X_\infty = \cup_{n=1}^{\infty} X_n$$

is also open and is the union of a countable number of elements of the covering Γ. The union X_∞^* of their closures is obviously contained in the closure $\overline{X}_\infty$ of the subspace X_∞ (in fact, it coincides with this closure). Being the union of a countable number of compact sets, the subspace X_∞^* is Lindelöf.

Since any element of the covering Γ is either contained in X_∞ or does not intersect X_∞, the complement $X \setminus X_\infty$ is a union of some elements of $U \in \Gamma$ and is therefore open. Consequently, the subspace X_∞ is not only open but also closed, and therefore $X_\infty^* = X_\infty$. Thus, the subspace X_∞ is Lindeöf.

To complete the proof, it remains to note that the subspaces X_∞ corresponding to different elements U_0 of the covering Γ either coincide or do not intersect and that the union of X is the entire space X. □

Connectedness

A topological space X is called *connected* if each of its non-empty closed and simultaneously open subspaces coincides with the entire space X. In other words, a space X is connected if it cannot be decomposed into a union of two (or more) non-empty disjoint open (or closed) sets. Any space X is a union, generally speaking, not free, of disjoint closed maximal connected subspaces, called the *connected components* of X. A connected Hausdorff space has no isolated points. In particular, a discrete space is connected if and only if it consists of a single point.

A set obtained from a connected set by adding some of its points of contact is connected. In particular, the closure of a connected set is connected. The union of connected sets with non-empty intersection is connected. For example, a space X is connected if any two of its points belong to a connected set. Any interval (a, b) or segment $[a, b]$ of the number line $\mathbb{R}$ is connected. In particular, the unit segment $I = [0, 1]$ is connected.

From the definition of a connected space and the theorem proved above it immediately follows that

Proposition 1.36. *any connected paracompact locally compact space is Lindelöf.*

1.3 Continuous functions

A numerical function f defined on a topological space X is called *continuous* if for any point $x_0 \in X$ and any $\varepsilon > 0$ there exists a neighbourhood U of the point x_0 that

$$|f(x) - f(x_0)| < \varepsilon$$

for all points $x \in U$. Such functions have many properties of continuous functions of a numerical argument. For example, any continuous function defined on a compact space X is bounded and takes its largest and smallest values. Any continuous function defined on a connected space X takes all intermediate values. The restriction $f|_A$ to an arbitrary subspace $A \subset X$ of any continuous function f on X is a function continuous on A.

For a numerical function f defined on a space X and an arbitrary number a, we will denote by the symbol $[f \leq a]$ the set of all points $x \in X$ for which $f(x) \leq a$. The notations $[f < a]$, $[f = a]$, $[a < f \leq b]$, etc. have a similar meaning. A function f is continuous if and only if for any $a \in \mathbb{R}$ the set $[f \leq a]$ is closed or, equivalently, if for any $a \in \mathbb{R}$ the set $[f < a]$ is open. From this, in particular, it follows that for any continuous function f and any number a the set $[f = a]$ is closed and the set $[f \neq a]$ is open. We will call the closure $\overline{[f \neq 0]}$ of the set $[f \neq 0]$ the *support* of the function f.

The sets $[f = a]$ are a special case of the *coincidence sets* $[f = g]$, defined for any pair of continuous (on the space X) functions f and g and consisting, by definition, of all points $x \in X$ for which $f(x) = g(x)$. It is easy to see that

Proposition 1.37. *for any two continuous functions f and g the set $[f = g]$ is closed.*

Proof. Indeed, if $x \notin [f = g]$ then $f(x) \neq g(x)$ and therefore the points $f(x)$ and $g(x)$ have disjoint neighbourhoods U and V. The set $f^{-1}(U) \cap g^{-1}(V)$ is open, contains the point x and does not intersect the set $[f = g]$. It is thus proved that the complement $X \setminus [f = g]$ of the set $[f = g]$ is open. Consequently, the set $[f = g]$ itself is closed. □

Continuous functions exist on any topological space X. Indeed, any constant function, i.e. a function that takes the same value at all points $x \in X$, is obviously continuous. However, non-constant continuous functions, generally speaking, may not exist. For their existence, it is sufficient that the space X be normal (and contain more than one point). Namely, it can be shown that for any two distinct points x_0 and x_1 of a normal space X, there exists a continuous function f on X such that $f(x_0) = 0$ and $f(x_1) = 1$. Moreover, as P. S. Uryson first proved,

Proposition 1.38. *for any closed set A of a normal space X and any of its neighbourhoods U, there exists on the space X a continuous function f that takes values from the interval $[0,1]$, is equal to one on A and zero outside U, i.e. such that*

$$A \subset [f = 1], \quad U \supset [f \neq 0].$$

This statement is known as Urysohn's lemma. Note that here we do not exclude cases when $A = \varnothing$ or $U = X$ (however, in these cases the lemma is trivial). We will call each function f satisfying the conditions of the Urysohn lemma a *Urysohn function* of the pair (U, A).

Proof. To prove Urysohn's lemma, we construct in the space X a family $\{V_r\}$ of neighbourhoods V_r of the set A contained in a neighbourhood U, numbered by binary-rational numbers $r \in (0,1)$ (i.e., numbers of the form $\frac{m}{2^n}$, where $n = 1, 2, \ldots$ and $m = 1, 2, \ldots, 2^n - 1$) and possessing the property that

$$\overline{V}_r \subset V_{r'}, \quad \text{if} \quad r < r'.$$

Such neighbourhoods V_r are easily constructed by induction on n. For $n = 1$, for the neighbourhood $V_{1/2}$ we should take an arbitrary neighbourhood V of the set A for which $\overline{V} \subset U$ (such a neighbourhood exists by the normality of the space X). Let for some $n \geq 1$ the neighbourhoods $V_{m/2^n}$ have already been constructed ($m = 1, 2, \ldots, 2^n - 1$). Let us take for each number $r = (2s+1)/2^{n+1}$, $s = 0, 1, \ldots, 2^n - 1$, the neighbourhood V_r to be an open set V (existing due to the normality of the space X) for which $\overline{V}_{s/2^n} \subset V$ and $\overline{V} \subset \overline{V}_{(s+1)/2^n}$ (for $s = 0$, the set $\overline{V}_{s/2^n}$ means the set A, and for $s = 2^n - 1$, the neighbourhood $V_{(s+1)/2^n}$ means the neighbourhood U). Thus, we obtain neighbourhoods V_r for $n + 1$ as well.

Having constructed the neighbourhoods V_r, we define the function g on the space X by the formula

$$g(x) = \sup_{x \notin V_r} r, \quad x \in X.$$

In other words, the value $g(x)$ of the function g at a point of $x \in X$ is equal to the upper line of all numbers r, for which $x \notin V_r$. The function g is unequivocally defined and continuous (for for any $a \in \mathbb{R}$ there are a lot of $[g < a]$ coincides with the many $\cup_{r<a} V_r$ and therefore open). Its values belong to a segment of $[0,1]$ and it is equal to zero on a lot of A and to unity outside the vicinity of U. Therefore, the function $f = 1 - g$ is the desired function of Uryson. □

Remark 1.39. Urysohn's lemma does not assert that $A = [f = 1]$ or that $U = [f \neq 0]$. Generally speaking, a function f for which at least one of these equalities holds may not exist. Closed sets of the form $[f = \text{const}]$ are sometimes called *functionally closed*. Similarly, open sets of the form $[f \neq \text{const}]$ are called *functionally open*. It can be proved that a closed (resp. open) set of a normal space is functionally closed (resp. open) if and only if it can be represented as the intersection (resp. union) of a countable family of open (resp. closed) sets.

We will not need this fact and will leave it without proof.

Let us now show that

Proposition 1.40. *for any continuous function g defined on a closed subspace A of a normal space X, there exists on the space X a continuous function f such that*

$$f(x) = g(x)$$

for each point $x \in A$.

This statement is known as Tietze's theorem (sometimes it is also called the Brouwer-Urysohn theorem). The function f provided by this theorem we will call the extension of the function g (from the subspace A to the entire space X).

Proof. First, we prove Tietze's theorem under the additional assumption that

$$|g(x)| \leq 1 \tag{1.41}$$

for all $x \in A$. We define by induction on the set A the sequence $\{g_n\}$ of continuous functions, setting

$$g_0 = g,$$
$$g_{n+1} = g_n + \frac{2^n}{3^{n+1}}(2h_n - 1), \quad n \geq 0,$$

where h_n is the Urysohn function on the space X, constructed for the closed (in A and therefore in X) set $[g_n \leq -\frac{2^n}{3^{n+1}}]$ and its neighbourhood $X \setminus [g_n \geq \frac{2^n}{3^{n+1}}]$. By induction, for any point $x \in A$, the estimate

$$|g_n(x)| \leq \left(\frac{2}{3}\right)^n$$

(it should be borne in mind that $g_{n+1}(x) = g_n(x) + \frac{2^n}{3^{n+1}}$ when $g_n(x) \leq -\frac{2^n}{3^{n+1}}$ and $g_{n+1}(x) = g_n(x) - \frac{2^n}{3^{n+1}}$ for $g_n(x) \geq \frac{2^n}{3^{n+1}}$.) Therefore

$$\lim_{n\to\infty} g_n x) = 0$$

for any point $x \in A$.

Now we compose from functions continuous on the space X

$$f_n(x) = -\frac{2^n}{3^{n+1}}(2h_n(x) - 1)$$

an infinite series

$$f_0(x) + f_1(x) + \cdots + f_n(x) + \cdots \tag{1.42}$$

Since for any point $x \in X$ the inequality $|f_n(x)| \leq \frac{2^n}{3^{n+1}}$ holds and the numerical series

$$\frac{1}{3} + \frac{2}{9} + \cdots + \frac{2^n}{3^{n+1}} + \cdots \tag{1.43}$$

converges, the functional series (1.42) also converges (at all points $x \in X$ and its sum $f(x)$ is continuous on X. Moreover, for $x \in A$, for the n-th partial sum $s_n(x) = f_0(x) + \cdots ; f_n(x)$ of Series (1.42), the equality

$$s_n(x) = g_0(x) - g_{n+1}(x)$$

holds. Passing to the limit in this relation (for $n \to \infty$), we immediately obtain that for any point $x \in A$, the equality

$$f(x) = g(x)$$

holds. Thus, Tietze's theorem under the additional assumption (1.41) is completely proved. In this case, the constructed function f satisfies the inequality

$$|f(x)| \leq 1, \quad x \in X$$

(since the sum of series (1.43) is equal to unity).

Let us now consider the case when for all points $x \in A$ the strict inequality

$$|g(x)| < 1, \tag{1.44}$$

is satisfied and we will show that then there exists an extension f of the function g for which a similar strict inequality

$$|f(x)| < 1, \tag{1.45}$$

is satisfied at all points $x \in X$.

Indeed, as has been proved, there exists an extension f_0 of the function g that satisfies the inequality $|f_0(x)| \leq 1$ at any point $x \in X$. Let $A_1 = A \cup [f_0 = 1] \cup [f_0 = 1]$. Let us define a numerical function g_1 on the set A_1, assuming

$$g_1(x) = \begin{cases} g(x), & \text{if} \quad x \in A, \\ 1, & \text{if} \quad x \in [f_0 = -1], \\ -1, & \text{if} \quad x \in [f_0 = 1]. \end{cases}$$

Since the sets A, $[f_0 = -1]$ and $[f_0 = 1]$ are closed and pairwise disjoint, the function g_1 is uniquely defined and continuous. Moreover,

$$|g_1(x)| \leq 1$$

for any point $x \in A$. Consequently, as proved above, on the space X there exists an extension f_1 of the function g_1 such that

$$|f_1(x)| \leq 1$$

for any point $x \in X$. But then the function

$$f(x) = \frac{f_0(x) + f_1(x)}{2}$$

will be an extension on X of the function g_1 satisfying the condition (1.45).

To prove Tietze's theorem in the general case, it is now sufficient to note that for any function $g(x)$ continuous on A, the function $g*(x) = \frac{2}{\pi} \arctan g(x)$ is continuous on A and satisfies the condition (1.44). Let $f_*(x)$ be its extension satisfying condition (1.45). It is clear that the function $f(x) = \tan\{\frac{\pi}{2} f(x)\}$ is continuous on X and is an extension of the function $g(x)$.

Thus, Tietze's theorem is completely proven. □

In addition to Tietze's theorem, it is useful to note that

Proposition 1.46. *if for any point $x \in A$ the inequalities*

$$a \leq g(x) \leq b$$

(resp. the inequalities $a < g(x) < b$) hold, then the extension f of the function g can be constructed so that for any point $x \in X$ the inequalities

$$a \leq f(x) \leq b$$

(resp. the inequalities $a < f(x) < b$) hold.

Proof. Indeed, for $a = -1$, $b = 1$ this statement has in fact already been proven above. The general case is reduced to this special one by an obvious linear transformation. □

Now let $\Gamma = \{U_\alpha; \alpha \in A\}$ be an arbitrary open locally finite covering of X. A family $\{f_\alpha; \alpha \in A\}$ of continuous functions f_α on X will be called a *partition of unity subordinate to the covering* Γ if

1) for any $\alpha \in A$ the function f_α is non-negative and its support $[\overline{f_\alpha \neq 0}]$ is contained in the element U_α of the covering Γ;

2) for any point $x \in X$ the equality

$$\sum_{\alpha \in A} f_\alpha(x) = 1$$

holds.

(due to the local, and therefore point, finiteness of the covering Γ, in the last sum for any point $x \in X$ only a finite number of terms are non-zero).

It turns out that

Proposition 1.47. *for any locally finite covering $\Gamma = \{U_\alpha; \alpha \in A\}$ of a normal space X there exists a subordinate partition of unity $\{f_\alpha; \alpha \in A\}$.*

To prove this statement, it suffices to prove that

Proposition 1.48. *for any locally finite covering $\Gamma = \{U_\alpha; \alpha \in A\}$ of a normal space X, there exists a covering $\Delta = \{V_\alpha; \alpha \in A\}$ such that*

$$\overline{V}_\alpha \subset U_\alpha$$

for any $\alpha \in A$.

Proof. Indeed, since the space X is normal, for any element V_α of the covering Δ there exists an open set W_α such that

$$\overline{V}_\alpha \subset W_\alpha, \quad \overline{W}_\alpha \subset U_\alpha$$

But then, as is easy to see, the functions

$$f_\alpha(x) = \frac{g_\alpha(x)}{\sum_{\alpha\in A} g_\alpha(x)}, \quad \alpha \in A, \quad x \in X,$$

where g_α is the Urysohn function of the pair $(W_\alpha, \overline{V}_\alpha)$, are uniquely defined, continuous (due to the local finiteness of the covering Γ) and constitute a partition of unity subordinate to the covering Γ.

All that remains for us, therefore, is to construct the covering Δ. For this purpose, we consider the set $\mathfrak{G}$ of all (open) coverings $\{G_\alpha; \alpha \in A\}$ of the space X for which for any $\alpha \in A$ either $\overline{G}_\alpha \subset U_\alpha$ or $G_\alpha = U_\alpha$. This set is not empty (for example, it contains the given covering Γ). For any such covering $\{G_\alpha; \alpha \in A\}$ we will denote by the symbol A_0 the set of all $\alpha \in A$ for which $\overline{G}_\alpha \subset U_\alpha$.

Let us introduce a partial ordering into the set $\mathfrak{G}$, assuming that $\{G_\alpha\} < \{G'_\alpha\}$ if $A_0 \subset A'_0$ and if $G_\alpha = G'_\alpha$ for any $\alpha \in A$. Roughly speaking, to obtain a "larger" covering $\{G'_\alpha\}$ from a covering $\{G_\alpha\}Oa$, we need, without touching the already constructed sets G_α, to replace some U_α with smaller sets G_α.

It is clear that this relation is indeed a partial ordering relation (i.e. it has the property of transitivity).

Let us now consider an arbitrary chain (= linearly ordered subset) in the set $\mathfrak{G}$. Denoting the elements of this subset by the symbols Γ^β, where β runs through some set of indices B, and setting $\Gamma^\beta = \{G^\beta_\alpha; \alpha \in A\}$, we define for each $\alpha \in A$ the set $G_\alpha \subset X$ by the formula

$$G_\alpha = \cap_{\beta\in B} G^\beta_\alpha.$$

It is clear that for any $\alpha \in A$ all sets $G^\beta_\alpha \neq U_\alpha$ (if such sets exist) coincide with each other. Therefore, there exists an index $\beta_\alpha \in B$ such that $G_\alpha = G^{\beta_\alpha}_\alpha$. Consequently, the set G_α is open and has the property that either $G_\alpha = U_\alpha$ or $\overline{G}_\alpha \subset U_\alpha$. We will show that the family $\{G_\alpha; \alpha \in A\}$ is a covering of the space X.

Let x be an arbitrary point in X and let A_x be a subset of A consisting of all indices $\alpha \in A$ for which $x \in U_\alpha$. Since the set A_x is finite due to the local finiteness of the covering Γ, among the indices β_α, $\alpha \in A_x$, there exists the largest (with respect to the ordering of the set B induced by the ordering of the set $\mathfrak{G}$) index β_{α_0}. Since the family $\Gamma^\beta_{\alpha_0} = \{G^{\beta_{\alpha_0}}_\alpha \alpha \in A\}$ is a covering, there exists an index $\alpha_* \in A$ such that $x \in \Gamma^{\beta_{\alpha_0}}_{\alpha_*}$. Since $G^{\beta_{\alpha_0}}_{\alpha_*} \subset U_{\alpha_*}$, the index α_* belongs to the subset A_x. But it is clear that for any index $\alpha \in A_x$ the equality $G^{\beta_{\alpha_0}}_\alpha = G^{\beta_\alpha}_\alpha = G_\alpha$ holds. Consequently, $x \in G_{\alpha_*}$.

The constructed covering $\{G_\alpha; \alpha \in A\}$ belongs, according to what was said above, to the set $\mathfrak{G}$ and is, as is easy to see, its smallest element, following all the elements of Γ^β, i.e., it represents the upper bound of the considered chain (= linearly ordered subset.)

Thus, we have proved that any chain (= linearly ordered subset) of the set $\mathfrak{G}$ has an upper bound. Consequently, according to Zorn's lemma[3], this set contains at least one maximal element $\{V_\alpha; \alpha \in A\}$.

To complete the proof, it remains to show that the covering $\Delta = \{V_\alpha; \alpha \in A\}$ has the property that for any of its elements V_α the inclusion holds

$$\overline{V}_\alpha \subset U_\alpha.$$

Suppose not. Then there exists an index $\alpha_0 \in A$ such that $V_{\alpha_0} = U_{\alpha_0}$. Consider the closed set

$$F = X \setminus \cup_{\alpha \neq \alpha_0} U_\alpha.$$

It is clear that $F \subset U_{\alpha_0}$. Therefore, since the space X is normal, there exists an open set V such that

$$F \subset V, \quad \overline{V} \subset U_{\alpha_0}.$$

It is easy to see that the family $\Delta' = \{V, V_\alpha; \alpha \in A \setminus \alpha_0\}$ is a covering of the space X, belongs to the set $\mathfrak{G}$, and is distinct from the covering Δ and has the property that $\Delta < \Delta'$. But in view of the maximality of the covering Δ this is impossible. □

Thus, the theorem formulated above is completely proven.

Remark 1.49. We needed the local finiteness of the covering Γ to ensure that all sets A_x were finite. Consequently, the covering Δ also exists for point-finite coverings Γ.

1.4 Metric spaces

A set X is said to be defined as a *metric space* or to have a *metric* introduced into it if any two of its points $x, y \in X$ are assigned a non-negative real number $\rho(x, y)$ (called the *distance* between these points), and the following axioms are satisfied:

1) the equality $\rho(x, y) = 0$ holds if and only if $x = y$;

2) for any two points $x, y \in X$ the equality holds

$$\rho(x, y) = \rho(y, x);$$

3) for any three points $x, y, z \in X$ the inequality holds

$$\rho(x, z) \leq \rho(x, y) + \rho(y, z).$$

[3]This lemma states that a partially ordered set in which every chain (= linearly ordered subset) has an upper bound contains a maximal element.

Axiom 1) is called the *axiom of non-degeneracy* of the metric, axiom 2) is called the *axiom of symmetry*, and axiom 3) is called the *triangle axiom*.

An example of a metric space is the n-dimensional arithmetic space $\mathbb{R}^n$ with the usual Euclidean metric (in this metric, the distance $\rho(\boldsymbol{u}, \boldsymbol{v})$ between the points $\boldsymbol{u} = (u_1, \ldots, u_n) \in \mathbb{R}^n$ and $\boldsymbol{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n$ is taken to be the length $|\boldsymbol{u} - \boldsymbol{v}| = \sqrt{(u_1 - v_1)^2 + \cdots + (u_n - v_n)^2}$ of the vector $\boldsymbol{u} - \boldsymbol{v}$).

The *distance* $\rho(A, B)$ between two subsets A and B of a metric space X is the greatest lower bound $\inf \rho(x, y$ of the distances $\rho(x, y)$, where x and y are arbitrary points of the subsets A and B, respectively. This distance can be equal to zero even without the sets A and B intersecting. For any $\varepsilon > 0$, the set $S_\varepsilon(A)$ of all points $x \in X$ for which $\rho(x, A) < \varepsilon$ is called a *spherical ε-neighbourhood* of the subset $A \subset X$. In any metric space, one can introduce one and only one topology in which all spherical neighbourhoods of each of its points are open and constitute a fundamental system of its neighbourhoods. In this topology, a set $A \subset X$ is closed if and only if any point $x \in X$ for which $\rho(x, A) = 0$ belongs to this set. We will call this topology the *natural topology* of the metric space X and in what follows we will consider each metric space (for example, the Euclidean space $\mathbb{R}^n$) as a topological space with this natural topology.

The topology of the space $\mathbb{R}^n$ can be described in another way, noting that it is a topological product of n copies of the number line $\mathbb{R}$. Therefore, the base of this space is, for example, the family of all parallelepipeds

$$a^1 < t^1 < b^1, \ \ldots, \ a^n < t^n < b^n, \quad (t^1, \ \ldots, \ t^n) \in \mathbb{R}^n,$$

where a^i and $b^i > a^i$, $i = 1, \ldots, n$ are arbitrary real numbers. The parallelepipeds for which the weight of the number a^i and b^i are rational obviously also form the base of the space $\mathbb{R}^n$.

By definition, the natural topology of a metric space X has the property that for any point $x \in X$, the system of all possible spherical neighbourhoods $S_\varepsilon(x)$ is a fundamental system of neighbourhoods of the point x. A similar statement also holds for compact subsets, i.e., for any compact subset $C \subset X$, the neighbourhoods $S_\varepsilon(C)$ constitute a fundamental system of its neighbourhoods. In other words, for any open set $U \supset C$, there exists $\varepsilon > 0$ such that $S_\varepsilon(c) \subset U$. For arbitrary closed subsets, the corresponding statement is, generally speaking, false.

In the natural topology of X, the metric $\rho(x, y)$ is a continuous function on the product $X \times X$. Moreover, for any point $x \in X$ and any closed set $A \subset X$, the function $\rho(x, A)$ is continuous on X. For any closed set $A \subset X$ and any neighbourhood U, the function

$$f(x) = \frac{\rho(x, X \setminus U)}{\rho(x, A) + \rho(x, X \setminus U)}$$

is the Urysohn function of the pair (U, A). For this function, the set $[f = 1]$ exactly coincides with the set A (and the set $[f \neq 0]$ with the set U). The set $[f > \frac{1}{2}]$ is a neighbourhood of the set A whose closure $[f \geq \frac{1}{2}]$ is contained in the neighbourhood U. Thus, any metric space, being obviously Hausdorff, is normal. Moreover, unlike arbitrary normal spaces, in the metric space X any closed set is functionally closed and any open set is functionally open.

A topological space X is called *metrisable* if there exists on it (generally speaking, not a unique) metric that is consistent with the topology of the space X, i.e. such that the natural topology defined by it coincides with the topology of the space X.

Proposition 1.50. *This metric can always be chosen so that the resulting metric space is bounded, i.e. so that the distance between any two of its points does not exceed some fixed number $K > 0$.*

Proof. Indeed, any metric ρ can be transformed into a bounded (by a number K) metric ρ', defining the same topology, by setting

$$\rho'(x, y) = K\frac{\rho(x, y)}{1 + \rho(x, y)}$$

for any points $x, y \in X$. □

Proposition 1.51. *Every discrete space is metrisable. The topological sum of any number and the topological product of a countable number of metrisable spaces are metrisable.*

Proof. The metric in the topological product of a countable number of metrisable spaces is introduced by means of an obvious infinite series. In order for this series to converge, it is sufficient to choose a metric in the n-th factor, bounded, say, by the number $1/n^2$. Verifying that the metric constructed in this way is consistent with the topology of the product does not present any difficulties, and we will omit it. □

From the metrisability of the product of metric spaces it immediately follows, in particular, that any metric space is stably normal.

Each subset A of a metric space X is naturally defined as a metric space, and the metric of the subset A is consistent with its topology induced by the natural topology of the entire space X.

In particular, any subset of the Euclidean space $\mathbb{R}^n$ is a metric space. Among these subsets, the unit ball $\mathbb{E}^n$ of the space $\mathbb{R}^n$, consisting of all points $\boldsymbol{v} \in \mathbb{R}^n$ for which $|\boldsymbol{v}| \leq 1$, will play a special role for us in what follows (here and below we identify points of the space $\mathbb{R}^n$ with their radius vectors). The ball $\mathbb{E}^n$ is closed in the space $\mathbb{R}^n$ and is the *closure of the open unit ball E^n* consisting of points $\boldsymbol{v} \in \mathbb{R}^n$ for which $|\boldsymbol{v}| < 1$. The boundary of the ball $\mathbb{E}^n$ is the unit sphere $\mathbb{S}^{n-1}$ consisting of points $\boldsymbol{u} \in \mathbb{R}^n$ for which $|\boldsymbol{u}| = 1$. Each point of the ball $\mathbb{E}^n$ has the form $v\boldsymbol{u}$, where $0 \leq v \leq 1$ and $\boldsymbol{u} \in \mathbb{S}^{n-1}$, and for $v \neq 0$ this representation is unambiguous.

In the definitions presented it was assumed that $n > 0$. Sometimes it will be convenient for us to extend them to the case $n = 0$, assuming that the ball $\mathbb{E}^0$ consists of one point. In this case we will assume that $E^0 = \mathbb{E}^0$ and that $\mathbb{S}^{-1} = \varnothing$.

It is easy to show that the ball $\mathbb{E}^n$ (and also the open ball E^n) and the sphere $\mathbb{S}^{n-1}$ (for $n > 1$) are connected. Moreover, the ball $\mathbb{E}^n$ and the sphere $\mathbb{S}^{n-1}$ are compact, and the open ball E^n is locally compact. Finally, the space $\mathbb{R}^n$ (and therefore all its subspaces) is a space with a countable base and, therefore, has countable local weight, is separable and Lindelöf.

It is easy, however, to see that

Proposition 1.52. *any metric space X is a space of countable local weight.*

Proof. Indeed, for each point $x \in X$ the neighbourhoods $S_{1/n}(x)$ obviously constitute a fundamental system of neighbourhoods. □

In particular,

Proposition 1.53. *every metric space is a compactly generated space.*

Moreover,

Proposition 1.54. *for any metric space X the following properties are equivalent:*

1) the space X has a countable base;

2) the space X is separable;

3) the space X is Lindelöf.

Proof. Indeed, implications 1) $\Rightarrow$ 2) and 1) $\Rightarrow$ 3) hold for any topological spaces.
To prove implication 2) $\Rightarrow$ 1) it suffices to note that for any countable everywhere dense set $\{x_n\}$ the open sets of the form $S_{1/m}(x_n)$, where $n, m = 1, 2, \ldots$, form a base of the space X.
To prove implication 3 $\Rightarrow$ 2), we, having chosen an arbitrary $n > 0$, consider the family of all subsets of the space X, each of which has the property that the pairwise distances between any two of its points are not less than $1/n$. It is clear that this family (partially ordered by inclusion) satisfies the conditions of Zorn's lemma and, therefore, it contains a maximal subset $C_{1/n}$. This subset is discrete and therefore (by the Lindelöf property) countable. Consequently, the set

$$C = \cup_{n=1}^{\infty} C_{1/n}$$

is also countable. To complete the proof, it remains to note that the set C is everywhere dense (since for any point $x \in X$ and any $n > 0$ in the set $C_{1/n}$, and - due to its maximality - therefore in the set C, there exists a point y such that $\rho(x, y) < 1/n$). □

Since every compact space is finally compact, this theorem implies that

Proposition 1.55. *every compact metric space is separable.*

Local compactness, however, is no longer sufficient for separability of a metric space. Nevertheless,

Proposition 1.56. *any connected locally compact metric space is separable.*

This theorem is due to P. S. Alexandrov. It follows immediately from the result of 2) above and the fact that

Proposition 1.57. *any metric space X is paracompact.*

The last proposition was first proven by Stone.

Proof. To prove it, we must establish that any covering $\{U_\alpha; \alpha \in A\}$ of X can be refined into a locally finite covering. For this purpose, we assume that the set of indices A of the covering under consideration is well-ordered. Let α_0 be its first element. Denoting for any set $A \subset X$ and any $n > 0$, by the symbol $[A]_n$, the set (obviously closed) of all points $x \in X$ for which $S_{2^{-n}}(x) \subset A$, we construct by induction for each $\alpha \in A$ and any $n > 0$ the closed set F_α^n, by

$$F_{\alpha_0}^n = [U_{\alpha_0}]_n, \quad F_\alpha^n = [U_\alpha \setminus \cup_{\beta<\alpha} F_\beta^n]_n.$$

Let x be an arbitrary point in X and let α be the smallest index in A for which $x \in U_\alpha$, and n be a number such that $S_{2^{-n}}(x)) \subset U_\alpha$. If $x \notin F_\alpha^n$, then $S_{2^{-n}}(x) \not\subset U_\alpha \setminus \cup_{\beta<\alpha} F_\beta^n$, and therefore there exists an index $\alpha_1 < \alpha$ such that $S_{2^{-n}}(x) \cap F_{\alpha_1}^n \neq \emptyset$. But then

$$x \in S_{2^{-n}}(F_{\alpha_1}^n) = S_{2^{-n}}([U_{\alpha_1} \setminus \cup_{\beta<\alpha} F_\beta^n]_n) \subset U_{\alpha_1} \setminus \cup_{\beta<\alpha} F_\beta^n \subset U_{\alpha_1},$$

which is impossible due to the choice of index α. Consequently, $x \in F_\alpha^n$. It is thus proved that

$$X = \cup_{n,\alpha} F_\alpha^n.$$

Let us further consider the sets

$$\Phi_\alpha^n = \overline{S}_{2^{-(n+3}}(F_\alpha^n), \quad G_\alpha^n S_{2^{-(n+2)}}(F_\alpha^n).$$

It is clear that the closed set Φ_α^n is contained in the open set G_α^n. Further, since

$$\overline{S}_{2^{-n}}(F_\alpha^n) \subset U_\alpha \setminus \cup_{\beta<\alpha} F_\alpha^n,$$

then for any $\beta < \alpha$ the intersection $S_{2^{-n}}(F_\alpha^n) \cap F_\beta^n$ is empty because $\rho(F_\alpha^n, F_\beta^n) \geq 2^{-n}$. Therefore, $\rho(G_\alpha^n, G_\beta^n) \geq 2^{-(n+1)}$. Since $\Phi_\alpha^n \subset G_\alpha^n$, it follows that for any $n > 0$ the set

$$\Phi^n = \cup_\alpha \Phi_\alpha^n$$

is closed.

Now we define by induction the open sets V_α^n, setting

$$V_\alpha^1 = G_\alpha^1, \quad V_\alpha^n = G_\alpha^n \setminus \cup_{m<n} \Phi^m.$$

Let $x \in X$. Since $\cup_{n,\alpha} F_\alpha^n = X$ and, therefore, $\cup_{n,\alpha} \Phi_\alpha^n = X$, there exist $n \geq 0$ and $\alpha \in A$ such that $x \in \Phi_\alpha^n$. If n is chosen (for a given α) to be the smallest possible, then

$$x \in \Phi_\alpha^n \setminus \cup_{m<n} \Phi_\alpha^m = \Phi_\alpha^n \setminus \cup_{m<n} \Phi^m$$

(as we know, $\Phi_\alpha^n \subset G_\alpha^n$ and therefore $\Phi_\alpha^m \cap \Phi_\beta^m \neq \varnothing$, if $\alpha \neq \beta$). Hence

$$x \in G_\alpha^n \setminus_{m<n} \Phi^m = V_\alpha^n.$$

This shows that the sets V_α^n form a covering of the space X.

Since

$$V_\alpha^n \subset G_\alpha^n \subset S_{2^{-n}}(F_\alpha^n) \subset U_\alpha \setminus \cap_{\beta<\alpha} F_\beta^n \subset U_\alpha$$

then the covering $\{V_\alpha^n\}$ is a refinement of the covering $\{U_\alpha\}$.

Finally, let $x \in X$ and let $x \in F_{\alpha'}^n$. Since

$$S_{2^{-(n+3)}}(x) \subset \Phi_{\alpha'}^n \subset \Phi^n$$

then

$$S_{2^{-(n+3)}}(x) \cap V_\alpha^m = \varnothing$$

for any $\alpha \in A$ and any $m > n$. On the other hand, since $\rho(G_\alpha^m, G_\beta^m) \geq 2^{-(m+1)}$ for $\alpha \neq \beta$, and $2 \cdot 2^{-(m+3)} < 2^{-(n+1)}$ for $m \leq n$, then for each $m \leq n$ there is at most one index $\alpha \in A$ for which $S_{2^{-(n+3)}}(x) \cap G_\alpha^m \neq \varnothing$, and hence at most one index $\alpha \in A$ for which $S_{2^{-(n+3)}}(x) \cap V_\alpha^m = \varnothing$. Thus, the neighbourhood $S_{2^{-(n+3)}}(x)$ of the point x intersects at most n elements of the covering $\{V_\alpha^n\}$. Consequently, this covering is locally finite.

Thus, Stone's theorem is completely proved. □

From Stone's theorem and the results of §1.2 it immediately follows, in particular, that

Proposition 1.58. *a metric space is compact if and only if it is sequentially compact.*

All the properties of metric spaces considered above are related not so much to metric spaces as to metrisable spaces and therefore had, in essence, a topological character. Let us now consider some "metric" properties of metric spaces, i.e. properties that essentially depend on the metric given in the space.

A sequence $\{x_n\}$ of points of a metric space X is called *fundamental* if for any $\varepsilon > 0$ there exists $N > 0$ such that $\rho(x_n, x_m) < \varepsilon$ when $n, m > N$. It is clear that any convergent sequence is fundamental. If the converse is true, i.e. if any fundamental sequence of points of X converges, then this space is called a *complete space*. Obviously, every closed subspace of a complete space is itself a complete space.

Proposition 1.59. *If any closed bounded subspace of a metric space X is compact, then X is complete.*

Proof. Indeed, let $\{x_n\}$ be an arbitrary fundamental sequence of points in X. If among its points there are only a finite number of distinct ones, then this sequence obviously converges. Let among the points x_n there be infinitely many distinct ones. The set of these points, being obviously discrete and bounded, cannot be closed (for otherwise it would be compact, and any discrete compact set is finite). Therefore, for this set there exists an adherent point a, that does not belong to it. This point is the limit of some convergent subsequence of the sequence $\{x_n\}$, and therefore of the entire sequence $\{x_n\}$ (since this sequence, by assumption, is fundamental). □

The conditions of this theorem are satisfied, in particular, by the Euclidean space $\mathbb{R}^n$. Therefore, the space $\mathbb{R}^n$ (and consequently any of its closed subspaces) is complete.

On the other hand, every compact metric space also satisfies these conditions. Consequently,

Proposition 1.60. *any compact metric space is complete.*

The requirement of local compactness for completeness is no longer sufficient, even if we additionally assume the existence of a countable base. However, in any locally compact topological space X with a countable base, one can introduce a metric consistent with the topology of this space, with respect to which the space X is a complete metric space. We will not need this result, and therefore we will leave it without proof.

It can be shown that any complete metric space is a space of the second category. The proof of this statement essentially repeats the proof of a similar proposition for Hausdorff locally compact spaces (see §1.2). It is only necessary to require that instead of the compactness of the sets $\overline{U}_n$, their diameters tend to zero. Since we will not need this statement either, we will not give this proof in detail.

1.5 Continuous maps

Let X and Y be arbitrary topological spaces. It is easy to see that

Proposition 1.61. *for any single-valued map*

$$f : X \to Y$$

of X into Y the following properties are equivalent:

1) *the complete preimage $f^{-1}(B)$ under the map f of an arbitrary closed set $B \subset Y$ is closed in X;*

2) *the complete preimage $f^{-1}(V)$ under the map f of an arbitrary open set $V \subset Y$ is open in X;*

2') *the complete preimage $f^{-1}(V)$ under the map f of an arbitrary element V of some prebase of open sets of Y is open in X;*

3) *for any point $x \in X$ and any neighbourhood V of $f(x)$) in Y there exists in X a neighbourhood U of x such that $f(U) \subset V$;*

3') *for any point $x \in x$ and any neighbourhood V of $f(x)$ in the space Y, belonging to some fundamental system of neighbourhoods of this point, there exists in the space X a neighbourhood U of the point x such that $f(U) \in V$;*

4) *for any set $A \subset X$ the inclusion holds*

$$f(\overline{A}) \subset \overline{f(A)}$$

Maps $f : X \to Y$ that have these properties are called *continuous*. These include, in particular, all constant maps, i.e. maps that take the entire space X to some fixed point y_0 of the space Y.

The continuous functions considered in §1.3 are nothing more than continuous maps of the space X into the real numbers $\mathbb{R}$.

The definition of the *coincidence set* $[f = g]$ is literally transferred from continuous functions to any continuous maps $f, g : X \to Y$. However, unlike the case of numerical functions, this set may not be closed. We can only assert that

Proposition 1.62. *if the space Y is Hausdorff, then for any two continuous maps $f, g : X \to Y$ the set $[f = g]$ is closed.*

Indeed, in the proof for continuous functions given in §1.3, only the Hausdorff property of the real line was used.

A special case of a coincidence set is the set $[f = \mathrm{id}_X]$ of all fixed points of the map $f : X \to X$ (i.e., the points $x \in X$ for which $f(x) = x$). According to what has just been said,

Proposition 1.63. *the set of fixed points of an arbitrary continuous map $f : X \to X$ of a Hausdorff space X is closed in itself in this space.*

The composition[4]

$$g \circ f : X \to Z$$

of any two continuous maps $f : X \to Y$, $g : Y \to Z$ is also a continuous map. Moreover, for any space X the identity map

$$\mathrm{id}_X : X \to X$$

(defined by the formula $\mathrm{id}_X(x) = x$) is continuous. In the language of category theory, these statements mean that the totality $\mathcal{X}$ of all topological spaces and all their continuous maps forms a category. Isomorphisms of this category, i.e. bijective[5] continuous maps $f : X \to Y$ for which the inverse map $f^{-1} : Y \to X$ is also continuous, are called *homeomorphisms*. Spaces X and Y for which there exists at least one homeomorphism $X \to Y$ are called *homeomorphic*. As a rule, we will further consider homeomorphic spaces as identical.

[4]A map $h : X \to Z$ is called the *composition* of the maps $f : X \to Y$ and $g : Y \to Z$ if $h(x) = g(f(x))$ for any point $x \in X$. The composition of the maps f and g is denoted by the symbol $g \circ f$.

[5]A map $f : X \to Y$ is called *injective* if $f(x_1) \neq f(x_2)$ when $x_1 \neq x_2$, *sujective* if $f(X) = Y$, and bijective if it is both injective and surjective.

Remark 1.64. In §1.1 we defined the concept of the topological sum of any non-intersecting spaces. The convention just introduced allows us to define the *topological sum* of arbitrary spaces, including intersecting ones, as the topological sum of spaces homeomorphic to them, but non-intersecting. The topological sum constructed in this way is defined up to homeomorphism.

A space Y is called a *continuous image* of a space X if there exists at least one continuous surjective map $X \to Y$. It is easy to see that a continuous image of any compact (resp. Lindelöf or countably compact) space is compact (resp. Lindelöf or countably compact). Similarly, a continuous image of a connected space is connected.

Continuous maps

$$u : I \to X$$

into the space X of the unit segment $I = [0, 1]$ we will call *paths* of the space X. We will call the points $x_0 = u(0)$ and $x_1 = u(1)$ respectively the *beginning* and *end* of the path u and we will say that the path u *connects* the point x_0 with the point x_1. We will call the space X *path-connected* if any two of its points can be connected by at least one path.

Proposition 1.65. *Any path-connected space X is connected.*

Proof. Indeed, for each path $u : I \to X$ the set $u(II) \subset X$ is connected (since the segment I is connected). Thus, any two points of a linearly connected space X belong to a connected set. Therefore, the space X is connected. □

An arbitrary topological space X decomposes into a union of disjoint linearly connected subspaces, called the *components of linear connectivity* of the space X. We will denote the set of all components of linear connectivity of the space X by the symbol $\pi_0(X)$.

Let the space X be represented as the union of some family $\{X_\mu;$ $mu \in M\}$ of its subspaces:

$$X = \cup_{\mu \in M} X_\mu$$

Then for any continuous map

$$f : X \to Y$$

and each $\mu \in M$ the partial map

$$f|_{X_\mu} : X_\mu \to Y,$$

as we know, is continuous. The converse, generally speaking, is not true, i.e. the continuity of the partial maps $f|_{X_\mu}$ does not imply the continuity of the map f. However, it is easy to see that

Proposition 1.66. *if the space X is a free union of its subspaces X_μ, $\mu \in M$, then for any space Y the map*

$$f : X \to Y$$

is continuous if and only if all partial maps

$$f|_{X_\mu} : X_\mu \to Y,$$

are continuous.

This proposition allows us to construct continuous maps $X \to Y$ from continuous maps

$$f_\mu : X_\mu \to Y$$

of subspaces X_μ, $x \in M$, of the space X, of which it is a free union. In this case, it is only required that the maps f_μ be
emphconsistent, i.e. that

$$f_{\mu_1}|_{X_{\mu_1} \cap X_{\mu_2}} = f_{\mu_2}|_{X_{\mu_1} \cap X_{\mu_2}}$$

for any $\mu_1, \mu 2 \in M$. Indeed, by putting

$$f(x) = f_\mu(x), \quad \text{if} \quad x \in X_\mu,$$

where X is an arbitrary point in the space X, we obtain a single-valued map

$$f : X \to Y$$

for which

$$f|_{X_\mu} = f_\mu$$

for any $\mu \in M$. According to the previous proposition, the map f constructed in this way is continuous.

As a rule, we will use this construction only in the case when the number of subspaces X_μ is finite and each of them is closed. As we know, in this case the space X is automatically a free union of subspaces X_μ.

Another (even more important) method of constructing continuous maps uses the map α^{-1} (generally speaking, multi-valued), the inverse of a given surjective continuous map

$$\alpha : P \to X.$$

Let $g : P \to Y$ be a continuous map such that the composite map

$$f = g \circ \alpha^{-1} : X \to Y$$

is single-valued. Under what conditions is the map f continuous?

In order to give (even if only an incomplete) answer to this question, we shall call a continuous surjective map $\alpha : P \to X$ an identification map if each subset $A \subset X$ whose complete preimage $\alpha^{-1}(A)$ under the map α is closed (resp. open) in P is itself closed (resp. open) in X. It is clear that

Proposition 1.67. *if the map α is an identification map, then the map $f = g \circ \alpha^{-1}$ (assumed to be single-valued) is continuous if and only if the map $g = f \circ \alpha$ is contnuous.*

We will call a subset $S \subset P$ *saturated* with respect to the map $\alpha :\to X$ if it is a complete preimage of its image, i.e. if

$$S = \alpha^{-1}(\alpha(S)).$$

For the set S to be saturated, it is sufficient to require that $S \supset \alpha^{-1}(\alpha(S))$, since always $S \subset \alpha^{-1}(\alpha(S))$. The union and intersection of saturated sets are saturated. In addition, the complete preimage $\alpha^{-1}(A)$ of an arbitrary set $A \subset X$ is saturated with respect to the map $\alpha : P \to X$.

For any continuous surjective map $\alpha : P \to X$, each closed (esp. open) set $A \subset X$ is the image of some closed (resp. open) saturated set $S \subset P$ (for example, the set $S = \alpha^{-1}(A)$). It turns out that the inverse property characterises identification maps, i.e.

Proposition 1.68. *a continuous surjective map $\alpha : P \to X$ is an identification map if and only if for any closed (resp. open) saturated set $S \subset P$ the set $\alpha(S)$ is closed (resp. open) in X.*

Proof. Indeed, if α is an identification map, then for any saturated closed (resp. open) set S the set $\alpha(S)$ is closed (resp. open), since the set $\alpha^{-1}(\alpha(S)) = S$ is closed (open). Conversely, let for any closed (resp. open) saturated set S the set $\alpha(S)$ be closed (resp. open). Consider an arbitrary set $S \subset X$ for which the set $S = \alpha^{-1}(A)$ is closed (resp. open) in P. Then, since the set S is saturated, the set $A = \alpha(S)$ will, by assumption, be closed (resp. open), and therefore the map α will be an identification map. □

It follows directly from the proved statement that

Proposition 1.69. *any continuous map α of a compact space P onto a Hausdorff space X is an identification map.*

Proof. Indeed, let the closed saturated set $S \subset P$ be the complete preimage of the subset $A \subset x$. Being a closed subset of the compact space P, the set S is compact and therefore its image $\alpha(S) = A$ is also compact, and therefore closed (since the space X is Hausdorff). □

In general, a continuous bijective map may not be a homeomorphism. However,

Proposition 1.70. *any bijective identification map $\alpha \to X$ is a homeomorphism.*

Proof. Indeed, any set $T \subset P$ is saturated with respect to the bijective map α and therefore, if it is closed (in P), then the set $(\alpha^{-1})^{-1}T = \alpha(T)$ is also closed (in X). But this also means that the inverse map $\alpha^{-1} : X \to P$ is continuous. □

In particular,

Proposition 1.71. *any bijective continuous map α of a compact space P onto a Hausdorff space X is a homeomorphism.*

Let, for example, P be an arbitrary bounded convex body lying in the Euclidean space $\mathbb{R}^n$, i.e., an arbitrary bounded closed (and therefore compact) convex subset of the space $\mathbb{R}^n$ containing interior points. It is easy to see that any such body is stellar with respect to any of its internal points $\boldsymbol{x}_0$, i.e., each ray

$$\boldsymbol{x}_0 + \boldsymbol{u}t, \quad \boldsymbol{u} \in \mathbb{S}^{n-1}, 0 \leq t < \infty$$

starting from the point $\boldsymbol{x}_0$ intersects the boundary of the body P at one point. In other words, for any vector $\boldsymbol{u} \in \mathbb{S}^{n-1}$ there exists a number $\varphi(\boldsymbol{u}) > 0$ such that $\boldsymbol{x}_0 + \boldsymbol{u}t \in P$ for $0 \leq t \leq \varphi(\boldsymbol{u})$ and $\boldsymbol{x}_0 + \boldsymbol{u}t \notin P$ for $t > \varphi(\boldsymbol{u})$. Moreover, elementary geometric considerations show that the function $\varphi(\boldsymbol{u})$ of the vector $\boldsymbol{u} \in \mathbb{S}^{n-1}$ is continuous on the sphere $\mathbb{S}^{n-1}$. Since, in addition, any point $\boldsymbol{x} \in P$ has the form $\boldsymbol{x}_0 + \boldsymbol{u}t$, where $\boldsymbol{u} \in \mathbb{S}^{n-1}$, $0 \leq t \leq \varphi(\boldsymbol{u})$, and for $\boldsymbol{x} \neq \boldsymbol{x}_0$ this representation is single-valued, then the formula

$$\alpha(\boldsymbol{x}_0 + \boldsymbol{u}t) = \frac{t}{\varphi(\boldsymbol{u})}\boldsymbol{u}$$

defines a bijective continuous map

$$\alpha : P \to \mathbb{E}^n$$

of the body P onto the unit ball $\mathbb{E}^n$. Since the body P is compact and the ball $\mathbb{E}^n$ is Hausdorff, this map is homeomorphic. Thus, we have proved that

Proposition 1.72. *any bounded convex body $P \subset \mathbb{R}^n$ is homeomorphic to the ball $\mathbb{E}^n$.*

From this, in particular, it follows that

Proposition 1.73. *for any $n \geq 0$ and $m \geq 0$ the product $\mathbb{E}^n \times \mathbb{E}^m$ of the balls $\mathbb{E}^n$ and $\mathbb{E}^m$ is homeomorphic to the ball $\mathbb{E}^{n+m}$.*

Proof. Indeed, the product $\mathbb{E}^n \times \mathbb{E}^m$ is obviously a bounded convex body of the space $\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$. □

Note, by the way, that

Proposition 1.74. *for $n \neq m$ the balls $\mathbb{E}^n$ and $\mathbb{E}^m$ are not homeomorphic.*

The proof of this statement, despite its obvious "obviousness", is not at all simple and requires a deep study of the topology of Euclidean spaces. Since we essentially do not need this statement, we will not prove it here.

We will call a continuous map $f : P \to X$ a *map of compact character* if the preimage $f^{-1}(C)$ of any compact set $C \subset X$ is a compact subset of the space P. It turns out that

Proposition 1.75. *if the space X is Hausdorff and compactly generated, then for any map $f : P \to X$ of compact character the subset $f(P)$ is closed in the space X, and if, in addition, the space P is also Hausdorff, then the map f, considered as a map onto the subset $f(P)$, is an identification map.*

Proof. Indeed, let C be an arbitrary compact subset of X. Since the set $f^{-1}(C)$ is compact by assumption, the set $f(f^{-1}(C))$ is also compact and therefore closed. On the other hand, since

$$f^{-1}(C) = f^{-1}(C \cap f(P)),$$

then $f(f^{-1}(C)) = C \cap f(P)$. Thus, the subset $f(P)$ is compactly closed. Consequently, it is closed. □

Similarly, if for some subset $A \subset f(P)$ the subset $f^{-1}(A) \subset P$ is closed, then for any compact set $C \subset S$ the set $ff^{-1}(A \cap C) = f^{-1}(A) \cap f^{-1}(C)$ is compact, and therefore the set $A \cap C = f(f^{-1}(A \cap C))$ is also compact. Thus, the set A is compactly closed and, therefore, closed.

In particular, we see that

Proposition 1.76. *an injective map of a compact character of a Hausdorff space to a Hausdorff compactly generated space is a homeomorphism onto a closed subspace.*

Let P, Q, X, Y be arbitrary topological spaces and

$$\alpha : P \to X, \quad \beta : Q \to Y$$

be arbitrary continuous maps. It is easy to see that the formula

$$(\alpha \times \beta)(p, q) = (\alpha(p), \beta(q)), \quad p \in P, q \in Q,$$

uniquely determines some continuous map

$$\alpha \times \beta : P \times Q \to X \times Y.$$

We will call this map the *topological product* of maps α and β. It is surjective (esp. injective) if maps α and β are surjective (resp. injective).

Along with the product $\alpha \times \beta$, we will also consider the topological sum

$$\alpha \cup \beta : P \cup Q \to X \cup Y$$

of the maps α and β. This sum is defined when the spaces $P \cup Q$ and $X \cup Y$ are defined (i.e. when $P \cap Q = \varnothing$ and $X \cap Y = \varnothing$), and, by definition, is a map that coincides on P with the map α, and on Q with the map β. It is also surjective (resp. injective) when the maps α and β are surjective (resp. injective). Furthermore, if the maps α and β are identification maps, then their sum $\alpha \cup \beta$ is also an identification map.

For the map $\alpha \times \beta$ the analogue of the last statement, generally speaking, does not apply. Having in mind to indicate sufficient conditions under which the product $\alpha \times \beta$ of two identification maps is also an identification map, we will call a continuous surjective map $\alpha : P \to X$ *locally compact* (resp. *locally countably compact*) if any point $p \in P$ has a neighbourhood saturated with respect to the map α, the closure of which is contained in a closed, saturated and compact (resp. countably compact) set $C \subset P$. If the space X is locally compact, then any surjective map $P \to X$ of compact character is locally compact. If a locally compact (resp. locally countable-compact) map$P \to X$ exists, then the space P is necessarily locally compact (resp. locally countable-compact). On the other hand, if the space P is compact (resp. countably compact), then for every space X any continuous sujective map $P \to X$ is locally compact (resp. locally countably compact).

It is easy to see that

Proposition 1.77. *if the space X is regular and the map $\alpha : P \to X$ is locally compact (resp. locally countably compact), then for any point $p \in P$ and any of its neighbourhoods S^* saturated with respect to the map α, there exists a saturated neighbourhood S of the point p whose closure $\overline{S}$ is compact (resp. countably compact) and is contained in the neighbourhood S^*.*

Proof. Indeed, by hypothesis, the point p has a saturated neighbourhood T, whose closure $\overline{T}$ is contained in a compact (resp. countably compact) set and is therefore itself compact (resp. countably compact). Let us consider the set $U = \alpha(T \cap S^*)$. Since the set $T \cap S^*$ is saturated with respect to the map α, the set U is open, i.e. it is a neighbourhood of the point $\alpha(p)$ in the space X. But then, due to the regularity of the space X, the point $\alpha(p)$ has a neighbourhood V such that $\overline{V} \subset U$. Let $S = \alpha^{-1}(V)$. The set S is saturated and is a neighbourhood of the point p. Moreover, since $\overline{S} \subset \alpha^{-1}(\overline{V}) \subset \alpha^{-1}(U) = T \cap S^*$, then, firstly, $\overline{S} \subset S^*$ and, secondly, $\overline{S} \subset T$, so that the set $\overline{S}$ is compact (resp. countably compact) and is contained in S^*. □

Let us now show that

Proposition 1.78. *if the identification map $\alpha : P \to X$ is locally compact and the space X is regular, then for any identification map $\beta : Q \to Y$ the map*

$$\alpha \times \beta : P \times Q \to X \times Y$$

is also an identification map.

Proof. Indeed, let W be an arbitrary subset of the product $X \times Y$ whose preimage $(\alpha \times \beta)^{-1}(W)$ under the map $\alpha \times \beta$ is open in the space $P \times Q$. We must prove that each such set W is open in the space $X \times Y$, i.e., that any of its points (x_0, y_0) is its interior point. In other words, we must prove that in the spaces X and Y there exist neighbourhoods U and V of the points x_0 and y_0, respectively, such that

$$(x_0, y_0) \in U \times V \subset W.$$

Let p_0 and q_0 be points in the spaces P and Q, respectively, such that $\alpha(p_0) = x_0$, $\beta(q_0) = y_0$. It is clear that to prove the existence of neighbourhoods U and V it is sufficient to prove that the points p_0 and q_0 have (in the spaces P and Q, respectively) saturated (with respect to the maps α and β, respectively) neighbourhoods S and T such that

$$(p_0, q_0) \in S \times T \subset (\alpha \times \beta)^{-1} W.$$

For this purpose, we consider the set S^* of all points $p \in P$ for which the following holds

$$(p, q_0) \in (\alpha \times \beta))^{-1}(W).$$

It is clear that the set S^* contains the point p_0, is open in the space P and is saturated with respect to the map α. Therefore, by the proposition proved above, the point p_0 has a neighbourhood S that is saturated with respect to the map α, whose closure $\overline{S}$ is compact and is contained in the neighbourhood S^*.

Let, further, T be the set of all points $q \in Q$ with the property that for any point $p \in \overline{S}$ the following inclusion holds

$$(p, q) \in (\alpha \times \beta)^{-1}(W).$$

In other words, T is the maximal subset of the space Q for which the following inclusion holds

$$\overline{S} \times T \subset (\alpha \times \beta)^{-1}(W).$$

In particular, we have

$$S \times T \subset (\alpha \times \beta)^{-1}(W).$$

Moreover, $q_0 \in T$ (since $\overline{S} \subset S^*$). Further,

$$\overline{S} \times \beta^{-1}(\beta(T)) \subset \alpha^{-1}(\alpha(\overline{S})) \times \beta^{-1}(\beta(T)) = (\alpha \times \beta)^{-1}(\alpha \times \beta)(\overline{S} \times T) \subset (\alpha \times \beta)^{-1}(W),$$

whence, in view of the maximality of the set T, it follows that $\beta^{-1}(\beta(T)) \subset T$, i.e., that the set T is saturated with respect to the map β.

Therefore, to complete the proof, we only need to prove that the set T is open in the space Q.

To this end, we note that since the set $(\alpha \times \beta)^{-1}(W)$ is, by assumption, open, each of its points is its interior point and therefore for points p and q there exist (in the spaces P and Q, respectively) neighbourhoods $S_q(p)$ and $T_p(q)$ such that

$$(p, q) \in S_q(p) \times T_p(q) \subset (\alpha \times \beta)^{-1}(W).$$

For each point $q \in T$, all possible sets of the form $S_q(P) \cap \overline{S}$, $p \in \overline{S}$, obviously constitute an open covering of the subspace $\overline{S} \subset P$. Therefore, since this subspace is, by construction, compact, there exists a finite system of points $p_1, \ \ldots, \ p_n \in \overline{S}$, such that

$$\overline{S} \subset \cup_{i=1}^n S_q(p_i).$$

Let

$$T(q) = \cap_{i=1}^n T_{p_i}(q).$$

It is clear that the set $T(q)$ is open (in Q), contains the point q and has the property

$$p \times T(q) \subset (\alpha \times \beta)^{-1}(W)$$

for any point $p \in \cup_{i1}^n S_q(p_i)$, and therefore, in particular, for any point $p \in \overline{S}$. Therefore, $T(q) \subset T$, i.e. $q \in \operatorname{int} T$. Thus, the set T, as stated, is open. Thus, the above statement is completely proven. □

It is possible to specify other conditions under which the product of identification maps is an identification map. For example,

Proposition 1.79. *if the identification maps $\alpha : P \to X$ and $\beta : Q \to Y$ are locally countably compact, the spaces P and Q are Hausdorff, and the spaces X and Y are regular, then the map*

$$\alpha \times \beta : P \times Q \to X \times Y$$

is an identification map.

Proof. Indeed, in this case the points p_0 and q_0 have (in the spaces P and Q, respectively) saturated neighbourhoods S^* and T^* with respect to the maps α and β, whose closures $\overline{S}^*$ and $\overline{T}^*$ are contained in the closed saturated countably compact sets $C \subset P$ and $D \subset Q$. Let

$$C_0 \subset C_1 \subset \cdots \subset C_n \subset \cdots .$$
$$D_0 \subset D_1 \subset \cdots \subset D_n \subset \cdots$$

be increasing sequences of compact subsets $C_n \subset P$ and $D_n \subset Q$, whose free unions are the sets C and D, respectively. Without loss of generality we can assume that

$$p_0 \in C_0, \quad q_0 \in D_0.$$

Note also that, since the spaces P and Q are, by assumption, Hausdorff, all sets C_n and D_n are closed.

First of all, for any $n \geq 0$ we will construct open sets $S_n \subset C_n$ and $T_n \subset D_n$ saturated with respect to the maps $\alpha|_{C_n}$ and $\beta|_{D_n}$ (in the spaces C_n and D_n, respectively) such that

$$p_0 \in S_n, \quad q_0 \in T_n, \quad \overline{S}_n \subset S_{n+1}, \quad \overline{T}_n \subset T_{n+1}$$

and

$$\overline{S}_n \times \overline{T}_n \subset W_n, \tag{1.80}$$

where

$$W_n = (\alpha|_{C_n} \times \beta|_{D_n})^{-1} = (C_n \times D_n) \cap (\alpha \times \beta)^{-1}(W).$$

To this end, we note that since the space C_n is compact, the map $\alpha|_{C_n}$ is locally compact. Therefore, the maps $\alpha|_{C_n} : C_n \to X$ and $\beta|_{D_n} : D_n \to Y$ satisfy the conditions of the previous proposition. Therefore, for any point $(p, q) \in W_n$ in the spaces C_n and D_n there exist saturated (with respect to the maps $\alpha|_{C_n}$ and $\beta|_{D_n}$) neighbourhoods $S^*_{q,n}(p)$ and $T^*_{p,n}(q)$ of the points p and q such that

$$(p, q) \in S^*_{q,n}(p) \times T^*_{p,n}(q) \subset W_n.$$

Furthermore, since the spaces X and Y are regular by assumption, the points p and q have in the spaces C_n and D_n saturated neighbourhoods $S_{q,n}(p)$ and $T_{p,n}(q)$ such that $\overline{S}_{q,n}(p) \subset S^*_{q,n}(p)$ and $\overline{T}_{p,n}(q) \subset T^*_{p,n}(q)$ (here we mean closures in the spaces C_n and D_n; however, since the subspaces C_n and D_n are closed, these closures coincide with the closures in the spaces P and Q). Thus, for any point $(p,q) \in W_n$ we have constructed in the spaces C_n and D_n saturated neighbourhoods $S_{q,n}(p)$ and $T_{p,n}(q)$ of the points p and q with respect to the maps $\alpha|_{C_n}$ and $\beta|_{D_n}$ such that

$$(p,q) \in \overline{S}_{q,n}(p) \times \overline{T}_{p,n}(q) \subset W_n.$$

In particular, for $n = 0$, $p = p_0$, $q = q_0$ we obtain neighbourhoods

$$S_0 = S_{q,0}(p), \quad T_0 = T_{p,0}(q)$$

of the points p_0 and q_0 (in the spaces C_0 and D_0) that have property (1.80).

Reasoning by induction, we assume that for some $n > 0$ neighbourhoods S_{n-1} and T_{n-1} possessing property (1.80) have already been constructed. It is clear that for any point $q \in \overline{T}_{n-1}$ all sets of the form $\overline{S}_{n-1} \cap S_{q,n}(p)$, $p \in \overline{S}_{n-1}$, form an open covering of the space $\overline{S}_{n-1}$. Since this space, being a closed subspace of the compact space C_{n-1}, is compact, there exists a finite system of points $p_1, \ldots, p_n \in \overline{S}_{n-1}$ such that

$$\overline{S}_{n-1} \subset S_{q,n},$$

where

$$S_{q,n}, = \cup_{i=1}^{s} S_{q,n}(p_i).$$

Let

$$T_n(q) = \cap_{i=1}^{s} T_{p_i,n}(q).$$

It is clear that the sets $S_{q,n}$ and $T_n(q)$ are saturated (with respect to the maps $\alpha|_{C_n}$ and $\beta|_{D_n}$), open (in the spaces C_n and D_n) and have the property that

$$\overline{S}_{q,n} \times \overline{T}_n(q) \subset W_n.$$

Moreover, $q \in T_n(q)$, so that all possible sets of the form $\overline{T}_{n-1} \cap \overline{T}_n(q)$, $q \in \overline{T}_{n-1}$ form an open covering of the space $\overline{T}_{n-1}$. Since the space $\overline{T}_{n-1}$ (similar to the space $\overline{S}_{n-1}$) is compact, there exists a finite system of points $q_1, \ldots, q_n \in \overline{T}_{n-1}$ such that

$$\overline{T}_{n-1} \subset T_n,$$

where

$$T_n = \cup_{i=1}^{t} T_n(q_i).$$

It is clear that the set T_n together with the set

$$S_n = \cap_{i=1}^{n} S_{q_i},$$

satisfies all the conditions imposed on the neighbourhoods of S_n and T_n.

Thus, the neighbourhoods S_n and T_n are constructed for all $n \geq 0$.

Now let

$$S_\infty = \cup_{i=1}^{\infty} S_n, \quad T_\infty = \cup_{i=1}^{\infty} T_n.$$

It is obvious that the sets S_∞ and T_∞ are saturated (with respect to the maps α and β, respectively), since the sets C and D are saturated by assumption. In addition, $p_0 \in S_\infty$, $q_0 \in T_\infty$, and $S_\infty \times T_\infty \subset (\alpha\times\beta)^{-1}(W)$. Therefore, the sets

$$S = S^* \cap S_\infty, \quad T = T^* \cap T_\infty$$

are also saturated and have the property that

$$(p_0, q_0) \in S \times T \subset (\alpha\times\beta)^{-1} W_n.$$

Therefore, to complete the proof, it remains only for us to show that the sets S and T are open (in the spaces P and Q, respectively).

mLet us first consider the set S_∞. Since the set S_n is open in the subspace C_n, then for any $m \leq n$ the set $S_n \cap C_m$ is open in the subspace $C_m \subset C_n$. On the other hand, $S_n \cap Cm = S_n \subset S_m \cap C_m$ for $m > n$. Therefore,

$$S_\infty \cap C_m = \cup_{n=m}^{\infty}(S_n \cap C_m)$$

for any $m \geq 0$ and this set is open in the subspace C_m. Since the space C is a free union of subspaces C_m, it follows that the set S_∞ is open in the space C and, therefore, the set $S = S^* \cap S_\infty$ is open in the subspace $S^* \subset C$. But the last subspace is open in the space P. Consequently, the set S is also open in the space P.

The fact that the set T is open in the space Q is proved similarly.

Thus, the statement formulated above is completely proven. □

1.6 Topologies of identification, glued spaces, relative homeomorphisms

Let P be an arbitrary topological space and let

$$\alpha : P \longrightarrow X$$

be an arbitrary map of the space P onto some set X. We introduce a topology into the set X, considering a subset $A \subset X$ to be open (resp. closed) if and only if its complete preimage $\alpha^{-1}(A)$ is open (resp. closed). We will call this topology the *identification topology* (defined by the map α). It is the weakest (i.e. containing the largest number of closed sets) topology of the space X in which the map α is continuous.

It is clear that

Proposition 1.81. *if the space X is equipped with the identification topology, then the map α is an identification map in the sense of §1.5.*

Conversely,

Proposition 1.82. *if a topological space X has the property that a given surjective map $\alpha : P \longrightarrow X$ is an identification map then the topology of this space is the identification topology defined by the map α.*

In other words, the identification topology of the space X is uniquely determined by the requirement that the given map α be an identification map.

In most applications, the set X is the set of all classes under some equivalence relation defined in the space P, and the map $\alpha : P \to X$ is the natural projection that associates with each point $p \in P$ its equivalence class. In this case, the set x, equipped with the identification topology, is called the *factor space* of the space P with respect to the given equivalence relation. However, the difference between factor spaces and any spaces equipped with the identification topology is essentially purely formal, since for any identification map $\alpha : P \to X$ there exists on P an equivalence relation such that the corresponding factor space is naturally homeomorphic to the space X. In this equivalence relation, the points $p_1, p_2 \in P$ are equivalent if and only if

$$\alpha(p_1) = \alpha(p_2).$$

An important example of a factor space arises when considering an arbitrary continuous map

$$f : A \to Y$$

of a closed subspace A of some topological space Z into a given space Y. Assuming that the spaces X and Y do not intersect, we introduce in their topological sum

$$P = X \cup Y$$

an equivalence relation, considering that

1) points $x_1, x_2 \in X$ are equivalent if and only if either $x_1 = x_2$ or $x_1, x_2 \in A$ and $f(x_1) = f(x_2)$;

2) points $x \in X$ and $y \in Y$ are equivalent if and only if $y = f(x)$;

3) points $y_1, y_2 \in Y$ are equivalent if and only if $y_1 = y_2$.

We will denote the corresponding factor space of the space P by the symbol $X \cup_f Y$ and we will say that *it is obtained by gluing the space X along the subspace A to the space Y by means of the map F.*

The natural projection

$$\alpha : P \to X \cup_f Y$$

is a homeomorphism on the subspace $Y \subset P$. Therefore, in the future we will, as a rule, identify the space Y with its image $\alpha(Y)$, i.e. we will consider the space Y as a subspace of the space $X \cup_f Y$:

$$Y \subset X \cup_f Y.$$

It is easy to verify that the space Y is closed in the space $X \cup_f Y$.

Further, as is easy to see, the natural projection a is a homeomorphism on the open set $X \setminus A$. Therefore, we can also assume that

$$X \setminus A \subset X \cup_f Y.$$

In this case

$$X \setminus A = (X \cup_f Y) \setminus Y,$$

so that the space $X \setminus A$ is open in the space $X \cup_f Y$.

Thus, the space $X \cup_f Y$ can be considered as the union

$$X \cup_f Y = (X \setminus A) \cup Y$$

of two mutually complementary spaces $X \setminus A$ and Y, the first of which is open and the second is closed. In this connection, we will sometimes say that the space $X \cup_f Y$ is obtained *by gluing* the spaces $X \setminus A$ and Y.

On the subspace A the natural projection α coincides with the map f.

It is customary to call a *pair of spaces* (or simply a *pair*) an arbitrary pair (X, A) consisting of some topological space X and some of its subspace A. We will call a pair (X, A) *Hausdorff* if the space X is Hausdorff, and *compact* if the space X is compact and its subspace A is closed.

Let (X, A) and (Z, Y) be arbitrary pairs. By a *map*

$$g : (X, A) \to (Z, Y) \tag{1.83}$$

of a pair (X, A) into a pair (Z, Y) we mean an arbitrary continuous map g of the space X into the space Y for which $g(A) \subset Y$. We will call the map (1.83) a *relative homeomorphism* if it homeomorphically maps the subspace $X \setminus A$ onto the subspace $Z \setminus Y$. From the above properties of the natural projection

$$\alpha : P \to X \cup fY$$

it immediately follows that

Proposition 1.84. *the restriction*

$$\alpha|_X : X \to X \cup_f Y$$

of the projection α on the space X is a relative homeomorphism of the pair (X, A) onto the pair $(X \cup_f Y, Y)$.

Relative homeomorphisms of the form $\alpha|_X$ are identification maps, and the subspace Y is closed for them. It turns out that in the class of all relative homeomorphisms, maps of the form $\alpha|_X$ are uniquely characterised by these properties up to homeomorphism, i.e.

Proposition 1.85. *for every relative homeomorphism*

$$g : (X, A) - \to (Z, Y),$$

for which the subspace Y is closed in the space Z and the map $g : X \to Z$ is an identification map, there exists a homeomorphism

$$h : X \cup_f Y \to Z, \quad f = g|_A,$$

identical on the subspace Y such that

$$g = h \circ \alpha|_X.$$

Thus, if the specified conditions are met, the space Z can be considered as the result of gluing the space X to the space Y along the subspace A by means of the map $f = g|_A$.

Remark 1.86. This formulation implies that $X \cap Y = \varnothing$. Otherwise, the space X should be replaced by a homeomorphic space that already has this property.

To prove the formulated proposition, we first note that the subspace A, being the prreimae under a continuous map of the closed subspace Y, is itself closed. Therefore (in view of the remark made above), the space $X \cup_f Y$ is defined. Let's consider the map

$$h : X \cup_f Y \to Z,$$

coinciding on $X \setminus A$ with the map g and identical on Y. It is clear that this map is bijective and has the property that the map $h \circ \alpha$ coincides on X with the map g (and is the identity map on Y). From this, firstly, it follows that the map h is continuous (since the map $h \circ \alpha$ is continuous, and the projection α is an identification map). Secondly, since for any closed set $C \subset X \cup_f Y$ the intersection of the closed set $(h \circ \alpha)^{-1}(h(C)) = \alpha^{-1}(C) \subset C \cup Y$ with the space X coincides with the set $g^{-1}(h(C)$, then, since the map g is an identification map, the set $g(C)$ is closed in Z. Consequently, the map h^{-1} is also continuous. Thus, the map h, as stated, is homeomorphic.

As we know, the condition on the map g is automatically satisfied if the space X is compact and the space Z is Hausdorff. Moreover, it is clear that in this case, for the space Y to be closed, it is sufficient that the space A is closed. Thus,

Proposition 1.87. *for any Hausdorff pair (Z, Y) relatively homeomorphic to a compact pair (X, A), the space Z is homeomorphic to the space obtained by gluing the space X to the space Y along the subspace A.*

Generally speaking, the space $X \cup_f Y$ may not be Hausdorff, even if the spaces X and Y are Hausdorff. However,

Proposition 1.88. *the space $X \cup_f Y$ is Hausdorff if*

1) the space $X \setminus A$ is Hausdorff;

2) each point of the subspace $X \setminus A$ has a neighbourhood whose closure does not intersect the subspace A;
and either

3) the space Y is Hausdorff and any two disjoint open (in A) subsets of the subspace A that are saturated with respect to the map f are cut out on A by disjoint open subsets of the space X;
or

4) the space Y is completely Hausdorff and any two disjoint closed subsets of the subspace A have disjoint neighbourhoods in X.

Proof. Indeed, condition 1) obviously ensures the Hausdorff property (the existence of disjoint neighbourhoods) for any pair of distinct points of the space $X \cup_f Y$ belonging to the subspace $X \setminus A$, and condition 2) ensures the Hausdorff property for any two points $x \in X \setminus A$ and $y \in Y$. Therefore, we need to check the Hausdorff property only for (distinct) points $y_1, y_2 \in Y$.

Let condition 3) be satisfied. Then the points y_1 and y_2 have in Y non-intersecting neighbourhoods V_1 and V_2. The preimages $f^{-1}(V_1)$ and $f^{-1}(V_2)$ (possibly empty) of these neighbourhoods also do not intersect and are open sets in A, saturated with respect to the map f. Therefore, according to the condition, in X there exist open disjoint sets U_1 and U_2 such that

$$U_1 \cap A = f^{-1}(V_1), \quad U_2 \cap A = f^{-1}(V_2).$$

It is clear that the sets

$$(U_1 \setminus A) \cup V_1, \quad (U_2 \setminus A) \cup V_2$$

are open in the space $X \cup_f Y$ and do not intersect (since they serve as images under the natural projection $\alpha : X \cup Y \to X \cup_f Y$ of non-intersecting, open, and saturated sets with respect to the map α of $U_1 \cup V_1$ and $U2 \cup V_2$). Since these sets contain the points y_1 and y_2, the Hausdorff property for these points is thus completely proved.

Let condition 4) be satisfied. Then the points y_1 and y_2 have neighbourhoods V_1 and V_2 in Y, the closures of which $\overline{V}_1$ and $\overline{V}_2$ do not intersect. Let us consider the sets $V_1' = f^{-1}(V_1)$ and $V_2' = f^{-1}(V_2)$. Since $\overline{V}_1' \subset f^{-1}(\overline{V}_1)$, $\overline{V}_2' \subset f^{-1}(\overline{V}_2)$ and $f^{-1}(\overline{V}_1) \cap f^{-1}(\overline{V}_2 = \varnothing$, then we have

$$\overline{V}_1' \cap \overline{V}_2' = \varnothing.$$

Thus, the sets $\overline{V}_1'$ and $\overline{V}_2'$ are disjoint closed subsets of the subspace A (recall that A is assumed to be closed). Therefore, according to the condition, these sets have disjoint neighbourhoods U_1 and U_2 in X. It is clear that the sets

$$(U_1 \setminus A) \cup V_1, \quad (U_2 \setminus A) \cup V_2$$

are open in the space $X \cup_f Y$ and do not intersect (since they serve as images under the natural projection α of non-intersecting, open and saturated sets $(U_1 \setminus A) \cup V_1' \cup V_1$ and $(U_2 \setminus A) \cup V_2' \cup V_2$ with respect to the map α.) Since these sets contain the points y_1 and y_2, the Hausdorff property is thus proved in this case as well. □

It is clear that conditions 1), 2) and 4) (in the part concerning the subspace A) are automatically satisfied if the space X is normal. Therefore,

Proposition 1.89. *if the space X is normal and the space Y is completely Hausdorff, then the space $X \cup_f Y$ is Hausdorff.*

Remark 1.90. It can be shown that if the spaces X and Y are normal, then the space $X \cup_f Y$ is also normal. We will not need this fact.

In the special case when the space Y consists of only one point y_0 (and, consequently, the map $f : A \to Y$ automatically turns out to be constant), the space $X \cup_f Y$ is denoted by the symbol X/A and is called the result of the *contraction of the subspace A to the point y_0*. In this case, condition 4) is obviously satisfied. As for conditions 1) and 2), they are certainly satisfied if the space X is regular. Thus,

Proposition 1.91. *if the space X is regular, then for any of its closed subspaces A the space X/A is Hausdorff.*

Another important special case of gluing arises when considering an arbitrary continuous map

$$f : X \to Y.$$

Let $X \times 1$ be a subspace of the product $X \times I$, where I is the unit segment $[0, 1]$ consisting of all points of the form $(x, 1)$, $x \in X$, and let

$$f_1 : X \times 1 \to Y$$

be the map into the space Y, defined by the formula

$$f_1(x, 1) = f(x).$$

Let us consider the space

$$Z_f = (X \times I) \cup_{f_1} Y.$$

This space is called the *mapping cylinder by f* and, as we shall see later, plays a fundamental role in the study of the homotopy properties of this map. Each of its points either has the form (x, t), where $x \in X$, $0 \leq t < 1$, or is a point y of the space Y.

It is clear that condition 2) for the space Z_f is always satisfied (a neighbourhood of the point (x, t), $x \in X$, $0 \leq t < 1$, the closure of which does not intersect the subspace $X \times 1$ is, for example, any neighbourhood of the form $X \times [0, t + \varepsilon)$, where ε is any positive number less than $1 - t$), and condition 1) is satisfied if the space X is Hausdorff. Moreover, condition 4) (in the part not related to the space Y) is also obviously always satisfied (open sets of the subspace $X \times 1$ have the form $G \times 1$, where G are open sets of the space X; they are cut out from $X \times 1$ by open sets $G \times I$ in $X \times I$; if the sets $G_1 \times 1$ and $G_2 \times 1$ do not intersect, then the sets $G_1 \times 1$ and $G_2 \times I$ also do not intersect). Consequently,

Proposition 1.92. *if the spaces X and Y are Hausdorff, then for any map $f : X \to Y$ the space Z_f is also Hausdorff.*

Chapter 2

Homotopy equivalences

This chapter mainly sets out various criteria that allow, in some cases, to judge whether a given continuous map will be a homotopy equivalence.

In the introductory §2.1, the basic concepts of the homotopy theory of continuous maps are presented and the simplest connections between these concepts are established. In particular, a simple but useful lemma is proved here, establishing conditions under which two homotopic maps that coincide on some subspace are homotopic relattive to this subspace. At the end of this section, the concept of a m-connected space is considered and some elementary properties of such spaces are proved.

In §2.2, after a number of simple remarks on homotopy equivalences, a well-known characteristic property of their cylinders is proved.

In §2.3 it is proved that (under certain conditions) the homotopy type of the glued space $X \cup_f Y$ depends only on the homotopy type of the space Y and the homotopy equivalence class of the map f.

In §2.4 the concept of weak homotopy equivalence is introduced and in connection with this a number of properties of homotopy groups are presented. However, the detailed theory of homotopy groups remains almost completely outside the scope of our exposition (it is enough to say that we do not even use their group operation here).

In §2.5 the concept of homotopy limit is considered (in both the "weak" and "strong" versions) and it is proved that the limit of homotopy equivalences is also a homotopy equivalence.

2.1 Homotopies and extensions of continuous maps

Each family

$$f_t : X \to Y, \quad 0 \leq t \leq 1, \tag{2.1}$$

of continuous maps of a topological space X into a topological space Y defines by the formula

$$F(x,t) = f_t(x), \quad x \in X, t \in I,$$

a certain map

$$F : X \times I \to Y \tag{2.2}$$

into the space Y of the product $X \times I$ of the space X and the unit segment $I = [0, 1]$. We will call the family (2.1) a *homotopy* of maps of the space X into the space Y if the corresponding map (2.2) is continuous. It is clear that, conversely, any continuous map (2.2) defines some homotopy (2.1).

We will call the map f_0 the *initial* map, and the map f_1 the *final* map of homotopy (2.1).

In what follows, we will often have to consider not separate spaces X and Y, but pairs (X, A) and (Y, B), where A and B are some subspaces of the spaces X and Y respectively. In this case, we will be interested, as a rule, only in homotopies of the form

$$f_t : (X, A) \to (Y, B),$$

i.e., homotopies (2.1) for which

$$f_t(A) \subset B$$

for any $t \in I$. We will call such homotopies *homotopies of pair maps.*

Maps

$$f, g : X \to Y$$

(or maps $f, g : (X, A) \to (Y, B)$) we will call *homotopic* (notation, $f \sim g$) if there exists a homotopy

$$f_t : X \to Y, \quad 0 \le t \le 1,$$

(or, correspondingly, a homotopy $f_t : (X, A) \to (Y, B)$ such that

$$f_0 = f, \quad f_1 = g$$

i.e., that the map f is its initial, and the map g is its final map. In this case we will also say that the maps f and g are related by a homotopy f_t (notation $f_t : f \sim g$).

Each continuous map

$$f : X \to Y$$

defines a certain homotopy

$$(1_f)_t : X \to Y, \quad 0 \le t \le 1,$$

for which

$$(1_f)_t(x) = f(x)$$

for any $x \in X$ and $t \in I$. We will call this homotopy a *stationary homotopy* of the map f.

The homotopy relation is, as is easy to see, an equivalence relation, i.e. it is

reflexive since $(1_f)_t : f \sim f$,

symmetric if $f_t f \sim g$, then $f_{1-t} : g \sim f$,

transitive if $f_t : f \sim g$ and $g_t : g \sim h$, then $h_t : f \sim h$, where

$$h_t = \begin{cases} f_{2t} & \text{for} \quad 0 \le t \le 1/2, \\ g_{2t-1} & \text{for} \quad 1/2 \le t \le 1. \end{cases}$$

Therefore, the set of all continuous maps $X \to Y$ (respectively, the set of all continuous maps $(X, A) \to (Y, B)$) splits into disjoint homotopy classes consisting of pairwise homotopic maps. We will denote the set of all homotopy classes of maps $X \to Y$ by the symbol $[X, Y]$, and maps $(X, A) \to (Y, B)$ by the symbol $[(X, A), (Y, B)]$. We will denote the class containing a given map f by the symbol $[f]$.

For any two homotopies

$$f_t : X \to Y, \quad g_t : Y \to Z, \qquad 0 \le t \le 1,$$

the family of maps

$$h_t = g_t \circ f_t : X \to Z, \quad 0 \le t \le 1,$$

is also, obviously, a homotopy. It follows that

Proposition 2.3. *if the maps $f_0 : X \to Y$ and $g_0 : Y \to Z$ are homotopic, respectively, to the maps $f_1 : X \to Y$ and $g_1 : Y \to Z$, then the map $g_0 \circ f_0 : X \to Z$ is homotopic to the map $g_1 : X \to Z$.*

In other words,

Proposition 2.4. *for any maps*

$$f : X \to Y, \quad g : Y \to Z$$

the homotopy class $[g \circ f] \in [X, Z]$ of the map

$$g \circ f : X \to Z$$

depends only on the homotopy classes $[f] \in [X, Y]$ and $[g] \in [Y, Z]$ of the maps f and g, respectively.

In particular, for any space Z and any continuous map

$$f : X \to Y$$

the formula

$$f_*[\varphi] = [f \circ \varphi], \quad \varphi : Z \to X,$$

uniquely determines some map of sets

$$f_* : [Z, X] \to [Z, Y].$$

We will say that the map f_* *is induced* by the continuous map f.

Similar statements hold, of course, for maps of pairs.

Recall that the map

$$f : X \to Y$$

is called the *extension* of the map

$$g : A \to Y,$$

where A is some subspace of the space X, if

$$f(a) = g(a)$$

for any point $a \in A$. In this case, we also say that the map g is a *restriction* of the map f on the subspace A and write

$$g = f|_A.$$

Otherwise, we can say that

$$g = f \circ i,$$

where

$$i : A \to X$$

is an inclusion map, i.e. a map defined by the formula

$$i(a) = a$$

for any point $a \in A$. In what follows, we will indicate that some map $i : A \to X$ is an inlusion map by replacing the symbol "$\to$" with the symbol "$\subset$", i.e. instead of $i : A \to X$ we will write

$$i : A \subset X.$$

A map

$$r : X \to A$$

of a space X onto its subspace A is called *retractive* if it is an extension of the identity map

$$1_A : A \to A,$$

i.e., if

$$r \circ i = 1_A,$$

or, in other words, if

$$r(a) = a$$

for any point $a \in A$. In this case, we will write

$$r : X \supset A.$$

Subspaces A of X for which retractive map $r : X \supset A$ exist are called its *retracts*. For a retractive map (or *retraction*, for short) $r : X \to A$ the set A coincides with the set of all fixed points of the map $i \circ r$. Therefore,

Proposition 2.5. *every retract A of a Hausdorff space X is closed in X.*

It is easy to see that

Proposition 2.6. *a space A of a space X is a retract of it if and only if for any space Y each map $g : A \to Y$ admits an extension $f : X \to Y$.*

Proof. Indeed, if any map $g : A \to Y$ can be extended, then, in particular, the identity map $1_A : A \to A$ can also be extended. Conversely, if there exists a retractive map $r : X \to A$, then for any map $g : A \to Y$ the composition $g \circ r : X \to Y$ will be an extension of the map g. □

The concept of the extension of maps is closely related to the concept of a homotopy of maps, since for any homotopy $f_t : X \to Y$ connecting the map $f : X \to Y$ with the map $g : X \to Y$, the corresponding map $F : X \times I \to Y$ is an extension to the space $X \times I$ of the map $(X \times 0) \cup (X \times 1) \to Y$, which maps each point $(x, 0)$, $s \in X$, to the point $f(x)$, and each point $(x, 1)$, $x \in X$, to the point $g(x)$.

A pair (X, A) is said to satisfy the *axiom of homotopy extension* if for any space Y, any map $f : X \to Y$ and any homotopy $g_t : A \to Y$ of the map $g = f|_A$ there exists a homotopy $f_t : X \to Y$ such that $f_0 = f$ and $f_t|_A = g_t$ for any $t \in I$. The significance of this axiom is primarily that for pairs (X, A) subject to it, the property of the map $g : A \to Y$ to allow the extension $f : X \to Y$ depends only on the homotopy class of the map g, i.e., together with the map g, each homotopic map $g' : A \to Y$ can be extended to the entire space X.

Examples of pairs satisfying the axiom of homotopy extension are given below.

It is clear that

Proposition 2.7. *if the pair (X, A) satisfies the axiom of homotopy extension, then the subspace $(X \times 0) \cup (A \times I)$ of the product $X \times I$ is its retract.*

Proof. Indeed, the homotopy f_t constructed for the identity map $1_X : X \to X$ and the stationary homotopy $(1_i)_t : S \to X$ of the inclusion map $i : A \subset X$ obviously defines a retracting map $X \times I \to (X \times 0) \cup (A \times I)$. □

If the subspace A is closed in the space X, then the converse is also true, i.e.

Proposition 2.8. *if the subspace $(X \times 0) \cup (A \times I)$ of the product $X \times I$ is its retract, and the subspace A is closed in the space X, then the pair (X, A) satisfies the axiom of homotopy extension.*

Proof. Indeed, the problem of constructing a homotopy f_t is equivalent to the problem of extending to the entire space $X \times I$ the map

$$G : (X \times 0) \cup (A \times I) \to Y,$$

defined by the formula

$$G(x, t) = \begin{cases} f(x), & \text{if} \quad t = 0, \\ g_t(x), & \text{if} \quad x \in A \end{cases}$$

(for closed A the space $(X\times 0)\cup(A\times I)$ is obviously a free union of the spaces $X\times 0$ and $A\times I$ and therefore this map is continuous). Therefore, if the space $(X\times 0)\cup(A\times I)$ is a retract of the space $X\times I$, then the homotopy f_t exists. □

In connection with the last statement, it is useful to keep in mind that

Proposition 2.9. *if the space X is Hausdorff and the subspace $(X\times 0)\cup(A\times I)$ of the product $X\times I$ is its retract, then the subspace S is closed in the space X.*

Proof. Indeed, then the subspace $(X\times 0)\cup(A\times I)$ is closed in the product $X\times I$ and therefore the set

$$A\times 1 = [(X\times 0)\cup(A\times I)]\cap(X\times 1)$$

is closed in the space $X\times 1$. □

Thus,

Proposition 2.10. *if the space X is Hausdorff or if the subspace A is closed, then the pair (X,A) satisfies the axiom of homotopy extension if and only if the subspace $(X\times 0)\cup(A\times I)$ is a retract of the product $X\times I$.*

Remark 2.11. The property of a pair (X,A) to satisfy the axiom of homotopy propagation is mainly local in nature, i.e., it is essentially determined by the structure of the space X in some neighborhood of the subspace A. The precise meaning of this statement can be given in many different ways. For example, it is easy to prove that

Proposition 2.12. *a pair (X,A) (with closed A) satisfies the axiom of homotopy extension if and only if the subspace A is functionally closed and there exists a homotopy $f_t : X\to Y$ and a function $\varphi : X\to I$ equal to zero on the subspace A such that*

$$\begin{aligned} f_0(x) &= x, \quad x\in X,\\ f_t(a) &= a, \quad (a,t)\in A\times I,\\ f_1(x) &\in A, \quad \text{if} \quad \varphi(x)<1. \end{aligned}$$

It can also be shown that

Proposition 2.13. *a pair (X,A) (with closed A) satisfies the axiom of homotopy extension if and only if there exists on the space X a continuous non-negative function φ equal to zero on the subspace A, and a map F into the space X of the subspace of the product $X\times I$ consisting of all points (x,t) for which*

$$0\le t\le\varphi(x),$$

having the following properties:

$$\begin{aligned} F(x,0) &= x, \quad \text{for any point} \quad x\in X,\\ F(x,\varphi(x)) &\in A, \quad \text{if} \quad \varphi(x)\le 1. \end{aligned}$$

We will not need these statements and therefore we will not prove them here.

A homotopy

$$f_t : X \to Y, \quad 0 \le t \le 1,$$

will be called a *homotopy relative to a subspace* $A \subset X$ if this homotopy is stationary on A, i.e. if

$$f_t(a) = f_0(a)$$

for any point $a \in A$ and any $t \in I$. Accordingly, we will call the two maps

$$f, g : X \to Y$$

homotopic relative to A (notation $f \sim g \operatorname{rel} A$) if they are related by some homotopy relative to A. Of course, for this it is necessary that

$$f|_A = g|_A,$$

i.e. that the maps f and g coincide on A. It is clear that homotopy relative to A is also an equivalence relation, and therefore the set of all continuous maps $X \to Y$ that coincide on A splits into disjoint *homotopy classes relative to* A. The class containing the given map $f : X \to Y$ we will denote by the symbol $[f] \operatorname{rel} A$, and the set of all such classes - by the symbol $[X, Y] \operatorname{rel} A$.

The properties of relative homotopy classes are similar to the properties of the "absolute" classes discussed above. For example, any continuous map $f : Y \to Z$ defines by the formula

$$f_*([\varphi] \operatorname{rel} A) = [f \circ \varphi] \operatorname{rel} A, \quad \varphi : X \to Y,$$

some map

$$f_* : [X, Y] \operatorname{rel} A \to [X, Z] \operatorname{rel} A.$$

If the space Y is Hausdorff, then any homotopy $f_t : X \to Y$ that is stationary on A will obviously also be stationary on the closure $\overline{A}$ of the subspace A. Consequently, in this case we can, without loss of generality, consider the subspace A to be closed.

The problem of constructing a homotopy relative to a closed subspace A is equivalent to the problem of extending to the entire space $X \times I$ the map

$$F : (X \times 0) \cup (A \times I) \to (X \times 1) \to Y,$$

defined by the formula

$$F(x, t) = \begin{cases} f(x), & \text{if} \quad t = 0 \text{ for } x \in A, \\ g(x), & \text{if} \quad t = 1. \end{cases}$$

In what follows, to simplify the formulae, for any pair (X, A) we will denote the subspace $(X \times 0) \cup (A \times I) \cup (X \times 1)$ of the space $X \times I$ (see Fig. 2.1) by the symbol $I(X, A)$. When A is closed, it is closed.

Note that if $f|_A = g|_A$ and $f \sim g$, then, generally speaking, it cannot be asserted that $f \sim \operatorname{rel} A$. However,

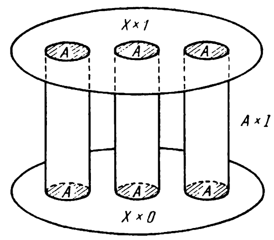


Figure 2.1:

Proposition 2.14. *if the pair $(X \times I, I(X,A))$ satisfies the axiom of homotopy extension and if*

$$H_{A\times I} \sim F\,\mathrm{rel}(A \times \cup A \times 1),$$

where H is the map $X \times I \to Y$ corresponding to the homotopy $h_t : X \times I \to Y$, connecting the maps f and g, and F is the map $A \times I \to Y$ corresponding to the stationary homotopy $(1_f)_t : A \to Y$, then

$$f \sim g\,\mathrm{rel}\,A$$

Proof. Indeed, we can extend the homotopy relative to $A\times 0 \cup A \times 1$ connecting the maps $H|_{A\times I}$ and F to some homotopy

$$H_t : I(X,A) \to Y \quad 0 \le t \le 1,$$

assuming that

$$H_t(x,0) = f(x), \quad H_t(x,1) = g(x)$$

for all $t \in I$. It is clear that

$$H_0|_{A\times I} = H|_{A\times I}.$$

Therefore, since the pair $X \times I, I(X,A))$ satisfies, by condition, the axiom of homotopy extension, there exists a homotopy

$$F_t : X \times I \to Y, \quad 0 \le t \le 1,$$

such that

$$F_0 = H, \quad F_t|_{I(X,A)} = H_t, \quad 0 \le t \le 1.$$

Therefore, the map

$$F_1 : X \times I \to Y,$$

defines (by the formula $f_t(x) = F_1(x,t)$) a homotopy

$$f_t : X \to Y, \quad 0 \le t \le 1,$$

relative to A, such that $f_0 = f$ and $f_1 = g$. Consequently, $f \sim g\,\mathrm{rel}\,A$. □

In what follows, continuous maps into a given space X of the unit n-dimensional sphere $\mathbb{S}^n$, $n \geq 0$ will be of great importance. In particular, we will be interested in the conditions under which any such map is homotopic to a constant map, i.e., a map to one point. In this connection, we first note that

Proposition 2.15. *a map $f : \mathbb{S}^n \to x$ is homotopic to a constant map if and only if the map f can be extended to some map $\mathbb{E}^{n+1} \to X$.*

Proof. Indeed, any homotopy

$$f_t : \mathbb{S}^n \to X,$$

for which $f_0 = f$ and $f_1(\boldsymbol{u}) = x_0$ for any point $\boldsymbol{u} \in \mathbb{S}^n$, where x_0 is some fixed point of the space X defines by the formula

$$F(v\boldsymbol{u}) = f_{1-v}(\boldsymbol{u}), \quad 0 \leq v \leq 1, \ \boldsymbol{u} \in \mathbb{S}^n,$$

a continuous map $F : \mathbb{E}^{n+1} \to X$, for which

$$F|_{\mathbb{S}^n} = f.$$

Conversely, any such map $F : \mathbb{E}^{n+1} \to X$ defines by the formula

$$f_t(\boldsymbol{u})F((1-t)\boldsymbol{u}), \quad 0 \leq t \leq 1, \ \boldsymbol{u} \in \mathbb{S}^n,$$

a homotopy $f_t : \mathbb{S}^n \to X$ for which $f_0 = f$ and $f_1(\boldsymbol{u}) = x_0$, where $x_0 = F(\boldsymbol{0})$. □

Thus,

Proposition 2.16. *for any space X the following properties are equivalent:*

1) any map $f : \mathbb{S}^n \to X$ is homotopic to a constant map;

2) any map $f : \mathbb{S}^n \to X$ can be extended to some map $F : \mathbb{E}^{n+1} \to X$.

Spaces with these properties we will call *n-aspherical*. Property 2) for $n = 0$ is obviously equivalent to the path-connectedness. Thus,

Proposition 2.17. *a space X is 0-aspherical if and only if it is path-connected.*

Spaces that are n-aspherical for all non-negative $n \leq m$ will be called *m-connected*.

It is easy to see that

Proposition 2.18. *the open unit interval $(0, 1)$ (as well as the closed segment $I = [0, 1]$) is an m-connected space for any $m \geq 0$.*

This statement follows immediately from Tietze's theorem (see §1.3), since the ball $\mathbb{E}^{n+1}$ is a normal space. However, it is easy to see that for each $n \geq 0$ the extension $F : \mathbb{E}^{n+1} \to (0, 1)$ of an arbitrary map $f : \mathbb{S}^n \to (0, 1)$ can be defined by the formula

$$F(v\boldsymbol{u}) = \frac{1 + (2f(\boldsymbol{u}) - 1)v}{2}, \quad 0 \leq v \leq 1, \ \boldsymbol{u} \in \mathbb{S}^n.$$

It is equally easy to see that

Proposition 2.19. *the topological product $X \times Y$ of two m-connected spaces X and Y is also an m-connected space.*

Proof. Indeed, let
$$f : \mathbb{S}^n \to X \times Y$$
be an arbitrary map of the n-dimensional ($n \leq m$) sphere $\mathbb{S}^n$ into the product $X \times Y$. Assuming for any point $\boldsymbol{u} \in \mathbb{S}^n$,

$$f(\boldsymbol{u}) = (f_1(\boldsymbol{u}), f_2(\boldsymbol{u}), \quad f_1(\boldsymbol{u}) \in X,\ f_2(\boldsymbol{u}) \in Y,$$

we obtain two (obviously continuous) maps

$$f_1 : \mathbb{S}^n \to X, \quad f_2 : \mathbb{S}^n \to Y.$$

By the condition, these maps can be extended to maps

$$F_1 : \mathbb{E}^{n+1} \to X, \quad F_2 : \mathbb{E}^{n+1} \to Y.$$

It is clear that the formula

$$F(\boldsymbol{v}) = (F_1(\boldsymbol{v}), F_2(\boldsymbol{v})), \quad \boldsymbol{v} \in \mathbb{E}^{n+1},$$

then defines a continuous map

$$F : \mathbb{E}^{n+1} \to X \times Y,$$

which is an extension of the map f. □

Comparing the proved statements, we obtain, in particular, that

Proposition 2.20. *for any m-connected space X the space $X \times (0, 1)$ is also m-connected.*

2.2 Homotopy equivalences and deformation retracts

Let X be an arbitrary space. Each homotopy

$$\xi_t : X \to X, \quad 0 \leq t \leq 1,$$

for which $\xi_0 = 1_X$, we will call the *deformation* of the space X. A continuous map
$$h : X \to X$$
we will call *homotopically identical* if it is homotopic to the identity map 1_X of the space X, i.e. if there exists a deformation $\xi_t : X \to X$ of the space X such that $\xi_1 = h$. A continuous map

$$f : X \to Y$$

will be called a *homotopy equivalence* if there exists a continuous map

$$g : Y \to X,$$

such that both maps

$$g \circ f : X \to X, \quad f \circ g : Y \to Y$$

are homotopically identical. In this case, the map g is also a homotopy equivalence. We will call it the homotopy equivalence *inverse* to the equivalence f. Since the equivalence f is in turn inverse to the equivalence g, we will sometimes call the equivalences f and g *mutually inverse*. It is clear that any map that is homotopic to a homotopy equivalence is also a homotopy equivalence. We will call spaces X and Y *homotopically equivalent* if there exists at least one homotopy equivalence $f : X \to Y$.

It is obvious that the composition

$$f_2 \circ f_1 : X \to Z$$

of two homotopy equivalences $f_1 : X \to Y$ and $f_2 : Y \to Z$ is also a homotopy equivalence. Therefore, the relation of homotopy equivalence of spaces is transitive. Since it is obviously reflexive and symmetric, the totality of all topological spaces decomposes into *homotopy types* of pairwise homotopy equivalent spaces.

A continuous map

$$f : X \to Y$$

we will call *homotopically injective* (resp. *homotopically surjective*) if there exists a continuous map such that the composition $g \circ f : X \to X$ (resp. the composition $f \circ : Y \to Y$) is homotopically identical. It is clear that any homotopy equivalence is a map that is both homotopy injective and homotopy surjective. It turns out that the converse is also true, i.e.

Proposition 2.21. *a map $f : X \to Y$ is a homotopy equivalence if and only if it is homotopy injective and homotopy surjective.*

Proof. Indeed, let the map f be homotopically injective and simultaneously homotopically surjective, i.e. let there exist maps

$$g_1 : Y \to X, \quad g_2 : Y \to X,$$

such that

$$g_1 \circ f \sim 1_X, \quad f \circ g_2 \sim 1_Y.$$

Then

$$f \circ g_1 \sim (f \circ g_1) \circ (f \circ g_2) = f \circ (g_1 \circ f) \circ g_2 \sim f \circ g_2 \sim 1_Y$$

and similarly

$$g_2 \circ f \sim 1_X.$$

Therefore, the map f is a homotopy equivalence and each of the maps g_1 and g_2 is a homotopy equivalence inverse to the equivalence f. □

Note that the above argument has a general, “purely categorical” character. Similar general considerations show that if for the map $f : X \to Y$ there exist a space Z and a map $h : Z \to X$ such that the composition $f \circ h$ is a homotopy equivalence, then the map f is homotopy surjective. Similarly, if there exist a space Z_1 and a map $h_1 : Y \to Z_1$ such that the composition $h_1 \circ f$ is a homotopy equivalence, then the map f is homotopy injective. Thus,

Proof. if for the map $f : X \to Y$ there exist maps $h : Z \to X$ and $h_1 : Y \to Z_1$ such that the compositions $f \circ h$ and $h_1 \circ f$ are homotopy equivalences, then the map f is also a homotopy equivalence. □

Let the spaces X and Y be homotopically equivalent to the spaces $X^{'}$ and $Y^{'}$, respectively. Continuous maps

$$f : X \to Y, \quad f^{'} : X^{'} \to Y^{'}$$

we will call ***homotopically equivalent*** if there exist homotopic equivalences

$$\varphi : X \to X^{'}, \quad \psi : Y \to Y^{'},$$

such that

$$\psi \circ f \sim f^{'} \circ \varphi$$

i.e. if the diagramme

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ {\scriptstyle\varphi}\downarrow & & \downarrow{\scriptstyle\psi} \\ X^{'} & \xrightarrow[f^{'}]{} & Y^{'} \end{array}$$

is homotopically commutative. Denoting by

$$\varphi^{'} : X^{'} \to X, \quad \psi^{'} : Y^{'} \to Y$$

the homotopy equivalences inverse to the equivalences φ and ψ, respectively, we immediately obtain that this condition is equivalent to both the condition

$$f \sim \psi^{'} \circ f^{'} \circ \varphi$$

and the condition

$$\psi \circ f \circ \varphi^{'} \sim f^{'}$$

It is clear that a map that is homotopically equivalent to a homotopically injective (resp. surjective) map is also homotopically injective (resp. surjective).

A trivial example of homotopy equivalence is an arbitrary homeomorphic map $f : X \to Y$. This shows that all the concepts introduced above are topologically invariant.

Less trivial make-ups of homotopy equivalences arise when considering retractive maps

$$f : X \supset A.$$

A subspace A is called a *deformation retract* of the space X if there exists a retractive map $r : X \supset A$ that is a homotopy equivalence. Since, by definition, $r \circ i = 1_A$, the retractive map r is necessarily homotopy injective. Therefore, it is a homotopy equivalence if and only if it is homotopy injective, i.e., when there exists a map $k : A \to X$ such that $k \circ r \sim 1_X$. As we have seen, in this case both maps i and k are necessarily homotopy equivalences, inverse to the homotopy equivalence r. In particular, the condition $i \circ r \sim 1_X$ will be satisfied. Thus,

Proposition 2.22. *for a retracting map r to be a homotopy equivalence, it is necessary and sufficient that the composite map*

$$i \circ r : X \to X$$

be homotopy identical.

Moreover,

Proposition 2.23. *for any deformation retract A the inclusion map $i : A \subset X$ is a homotopy equivalence.*

Note that the converse is generally not true.

We obtain an important example of a deformation retract by considering (see §1.6) the cylinder Z_f of an arbitrary continuous map

$$f : X \to Y.$$

Namely, as we will now show,

Proposition 2.24. *the subspace Y of the space Z_f is its deformation retract.*

Proof. To this end, for any $\tau \in I$ we define the map

$$p_\tau : Z_f \to Z_f, \quad 0 \le \tau \le 1,$$

of the space Z_f into itself, putting

$$p_\tau(x,t) = (x, t + \tau - \tau t), \quad x \in X,\ 0 \le t \le 1,$$
$$p_\tau(y) = y, \quad y \in Y.$$

(In the first of these formulae, as in similar cases below, the symbol $(x,1)$ is understood to mean the point $f(x) \in Y$.) The maps p_τ constructed in this way obviously constitute a homotopy (i.e. the corresponding map $Z_f \times I \to Z_f$ is unique and continuous). In addition, $p_0 = 1_{Z_f}$, i.e. this homotopy is a deformation of the space Z_f. Finally, for the map

$$p_1 : Z_f \to Z_f$$

the formula

$$p_1 = j \circ p,$$

holds, where

$$j : Y \subset Z_f,$$

is the inclusion map, and

$$p : Z_f \supset Y$$

is the retracting map of the space Z_f onto its subspace Y, defined by the formulae

$$p(x,t) = f(x), \quad x \in X, \ 0 \leq t \leq 1,$$
$$p(y) = y, \quad y \in Y.$$

Since $p \circ j = 1_Y$, the above statement is thus completely proven. □

Another remarkable property of the space Z_f is that

Proposition 2.25. *the map*

$$j \circ f : X \to Z_f$$

is homotopic to the map

$$i : X \to Zf,$$

defined by the formula

$$i(x) = (x, 0), \quad x \in X.$$

Proof. Indeed, the formula

$$i_t(x) = (x, 0), \quad x \in X, \ 0 \leq t \leq 1,$$

defines, as is easy to see, a homotopy

$$i_t : X \to Z_f,$$

connecting the map f with the map $j \circ f$. □

Since the inclusion map j is, as proved, a homotopy equivalence, this proposition means that the maps i and f are homotopy equivalent. On the other hand, identifying each point $x \in X$ with the corresponding point $i(x) = (x, 0) \in Z_f$, we can assume that

$$i : X \subset Z_f.$$

Thus, it is proved that

Proposition 2.26. *for any map $f : X \to Y$ there exists a space Z_f homotopically equivalent to the space Y and containing the space X such that the map $f : X \to Y$ is homotopically equivalent to the inlclusion map $i : X \subset Z_f$.*

Since a map homotopically equivalent to a homotopy equivalence is itself a homotopy equivalence, it follows from this statement that

Proposition 2.27. *if a subspace X of Z_f is its deformation retract, then the map f is a homotopy equivalence.*

It turns out that the converse is also true, i.e.

Proposition 2.28. *for any homotopy equivalence $f : X \to Y$ the space X is a deformation retract of the space Z_f.*

Proof. Indeed, let

$$g : Y \to X$$

be a homotopy equivalence inverse to the homotopy equivalence f, and let

$$\xi_t : X \to X, \quad \eta_t : Y \to Y$$

be deformations of the spaces X and Y, respectively, such that

$$\xi_1 = g \circ f, \quad \eta_1 = f \circ g.$$

For any point $z \in Z_f$ and any $\tau \in I$ we set

$$q_\tau(z) = \begin{cases} (x, t + 4\tau(1-t)) & \text{for} \quad 0 \le \tau \le 1/4, \\ \eta_{4\tau-1}(f(x)) & \text{for} \quad 1/4 \le \tau \le 1/2, \\ (g(f(x)), 3 - 4\tau) & \text{for} \quad 1/2 \le \tau \le 3/4, \\ \xi_{1+(1-t)(3-4\tau)} & \text{for} \quad 3/4 \le \tau \le 1, \end{cases}$$

if $z = (x, t)$, $xinX$, $0 \le t \le 1$, and

$$q_\tau(z) = \begin{cases} y & \text{for} \quad 0 \le \tau \le 1/4, \\ \eta_{4\tau-1}(y) & \text{for} \quad 1/4 \le \tau \le 1/2, \\ (g(y), 3 - 4\tau) & \text{for} \quad 1/2 \le \tau \le 3/4, \\ g(y) & \text{for} \quad 3/4 \le \tau \le 1, \end{cases}$$

if $z = y \in Y$.

It is easy to verify that the family

$$q_\tau ; Z_f \to Z_f, \quad 0 \le \tau \le 1,$$

defined in this way is a homotopy (recall that according to the results of §1.6, the map $\alpha \times 1_I : [(X \times I \cup Y] \times I \to Z_f \times I$, where $\alpha : (X \times I) \cup Y \to Z_f$ is the natural projection, is an identification map). Since, in addition,

$$q_0 = 1_{Z_f},$$

thus the homotopy q_τ is a deformation of the space Z_f.

On the other hand, the map

$$q_1 : Z_f \to Z_f$$

obviously has the form

$$q_1 = i \circ q,$$

where q is the map $Z_f \to X$ defined by the formulae

$$q(x, t) = \xi_t(x), \quad x \in X, \; 0 \le t \le 1,$$
$$q(y) = g(y), \quad y \in Y.$$

Since $q \circ i = 1_X$ the above statement is thus completely proven. □

2.3 Homotopy type of glued spaces

Let X and Y be topological spaces, A be a closed subspace of X and $f : A \to Y$ be a continuous map of the subspace A into the space Y. Then (see §1.6) the space $X \cup_f Y$ is defined, obtained by gluing the space X along the subspace A to the space Y by means of the map f. In this section we will show that for a sufficiently "good" pair (X, A) the homotopy type of the space $X \cup_f Y$ depends only on the homotopy class of the map f and the homotopy type of the space Y. In other words, for any map $g : A \to Y$ homotopic to the map f, the space $X \cup_g Y$ is homotopically equivalent to the space $X \cup_f Y$, and for ny space Z homotopically equivalent to the space Y, the space $X \cup_{h \circ f} Z$, where h is an arbitrary homotopy equivalence $Y \to Z$, is homotopically equivalent to the space $X \cup_f Y$.

First of all, we will show that

Proposition 2.29. *if a pair (X, A) has the property that both it and the pair $(X \times I, I(X, A))$ satisfy the axiom of homotopy extension, then for any two homotopic maps*

$$f, g : A \to Y$$

the spaces $X \cup_f Y$ and $X \cup_g Y$ are homotopy equivalent, and the homotopy equivalence $X \cup_f Y \to X \cup_g Y$ can be chosen in such a way that on the space Y it is the identity map.

Proof. Let

$$f_t : A \to Y, \quad 0 \le t \le 1,$$

be an arbitrary homotopy connecting the map f with the map g. First, we extend this homotopy to some homotopy of the space $\cup Y$ into the space $X \cup_f Y$. Since the natural projection

$$\alpha : X \cup Y \to X \cup_f Y$$

on the subspace $A \subset Y$ coincides with the map $f : A \to Y$ and since the pair (X, A) satisfies, by hypothesis, the axiom of homotopy extension, then there exists a homotopy

$$\alpha_t^* : X \to X \cup_f Y, \quad 0 \le t \le 1,$$

such that

$$\alpha_0^* = \alpha|_X, \quad \alpha_t^* = f_t.$$

Putting

$$\alpha_t = \begin{cases} \alpha_t & \text{on} \quad X, \\ 1_Y & \text{on} \quad Y, \end{cases} \quad 0 \le t \le 1,$$

we obviously obtain a homotopy

$$\alpha_t : X \cup Y \to X \cup_f Y, \quad 0 \le t \le 1,$$

for which

$$\alpha_0 = \alpha, \quad \alpha_t|_A = f_t.$$

Since the homotopy α_t has the property that

$$\alpha_1|_A = g,$$

the map

$$\eta = \alpha_1 \circ \beta^{-1} : X \cup_g Y \to X \cup_f Y,$$

where

$$\beta : X \cup Y \to X \cup_g Y$$

is the natural projection, is a single-valued map. Moreover, since the map

$$\alpha_1 = \eta \circ \beta$$

is continuous, and the map β is an identification map, the map η is continuous (see §1.5). On the space Y this map is identical:

$$\eta|_Y = 1_Y.$$

Let us now construct the "inverse" map $X \cup_g Y \to X \cup_f Y$. By the same considerations as above, there exists a homotopy

$$\beta_t : X \cup Y \to X \cup_g Y, \quad 0 \le t \le 1,$$

such that

$$\beta_0 = \beta, \quad \beta_t|_A = f_{1-t},$$

and

$$\beta_t|_Y = 1_Y.$$

Since $\beta_1|_A = f$, then, setting

$$\xi = \beta_1 \circ \alpha^{-1},$$

we obtain a single-valued continuous map

$$\xi : X \cup_f Y \to X \cup_g Y,$$

for which, as for η,

$$\xi|_Y = 1_Y.$$

The above statement will obviously be proved if we show that the maps ξ and η are mutually inverse homotopy equivalences. With this in mind, for any number $t \in I$ we set

$$\gamma_t = \begin{cases} \alpha_{2t}, & \text{if} \quad 0 \le t \le 1/2, \\ \eta \circ \beta_{2t-1}, & \text{if} \quad 1/2 \le t \le 1. \end{cases}$$

Since $\alpha_1 = \eta \circ \beta_0$, then we thereby obtain some homotopy

$$\gamma_t : X \cup Y \to X \cup_f Y, \quad 0 \le t \le 1,$$

for which

$$\gamma_0 = \alpha, \quad \gamma_1 = \eta \circ \xi \circ \alpha.$$

The homotopy γ_t is *not*, generally speaking, a homotopy relative to A, because

$$\gamma_t|_A = \begin{cases} f_{2t}, & \text{if} \quad 0 \leq t \leq 1/2, \\ f_{2t-1}, & \text{if} \quad 1/2 \leq t \leq 1. \end{cases}$$

Nevertheless, the map $A \times I \to X \cup_f Y$ corresponding to the homotopy $\gamma_t|_A$, as is easy to see, is homotopic relative to $A \times 0 \cup A \times 1$ to the map $A \times I \to X \cup_f Y$, which maps each point $(a,t) \in A \times I$ to the point $\alpha(a) \in X \cup_f Y$, i.e., which defines a stationary homotopy of the map $\alpha|_A$. The corresponding homotopy

$$H_\tau : A \times I \to X \cup_f Y, \quad 0 \leq \tau \leq 1,$$

can, for example, be defined by the formula

$$H_\tau(a,t) = \begin{cases} f_{2t(1-\tau)}(a), & \text{if} \quad 0 \leq t \leq 1/2, \\ f_{2(1-t)(1-\tau)}(a), & \text{if} \quad 1/2 \leq t \leq 1. \end{cases}$$

where (a,t) is an arbitrary point in space $A \times I$. Therefore, according to the statement proved in 2.14 (the conditions of applicability of which are fulfilled), the maps α and $\eta \circ \xi\alpha$ are also homotopic relative to A. The corresponding homotopy

$$\gamma_t^* : X \cup Y \to X \cup_f Y, \quad 0 \leq t \leq 1,$$

relative to A has the property that

$$\gamma_t^*|_A = f$$

for any $t \in I$. Therefore, the formula

$$h_t = \gamma_t^* \circ \alpha^{-1}, \quad 0 \leq t \leq 1,$$

uniquely defines a certain homotopy

$$h_t : X \cup_f Y \to X \cup_f Y$$

(recall that by the results of §1.5 the map

$$\alpha \times 1_I : (X \cup Y) \times I \to (X \cup_f Y) \times I$$

is an identification map), which, obviously, has the property that

$$h_0 = 1_{X \cup_f Y}, \quad h_1 = \eta \circ \xi.$$

Therefore, the map

$$\eta \circ \xi : X \cup_f Y \to X \cup_f Y$$

is homotopy identical.

Since the maps f and g are completely equivalent, then, by symmetry considerations, the map

$$\xi \circ \eta : X \cup_g Y \to X \cup_g Y$$

is also homotopy identical.

Thus, the proposition formulated above is completely proven. □

Let us now prove that the homotopy type of the space $X \cup_f Y$ does not change even when the space Y is replaced by a space homotopically equivalent to it. More precisely, we will prove that

Proposition 2.30. *under the same assumptions on the pair (X, A) as above, any homotopy equivalence $f : Y \to Z$ can be extend to some homotopy equivalence*

$$H : X \cup_f Y \to X \cup_{h\circ f} Z.$$

Proof. Let

$$g : Z \to Y$$

be the homotopy equivalence inverse to h, and let

$$H = \beta^{-1} \circ \overline{h} \circ \alpha : X \cup_f Y \to X \cup_{h\circ f} Z,$$
$$G = \alpha^{'} \circ \overline{g} \circ \beta^{-1} : X \cup_{h\circ f} Z \to X \cup_{g\circ h\circ f} Y,$$

where, respectively,

$$\overline{h} = 1_X \cup h : X \cup Y \to X \cup Z,$$
$$\overline{g} = 1_X \cup g : X \cup Z \to X \cup Y,$$
$$\alpha : X \cup Y \to X \cup_f Y,$$
$$\alpha^{'} : X \cup Y \to X \cup_{g\circ h\circ f} Y,$$
$$\beta : X \cup Z \to X \cup_{h\circ f} Z,$$

are the natural projections. Since

$$\beta \circ \overline{h}|_A = \beta|_A = h \circ f, \quad \alpha^{'} \circ \overline{g}|_A = \alpha^{'}|_A = g \circ h \circ f,$$

then the maps H and G are single-valued and continuous.

Since the map

$$g \circ h : Y \to Y$$

is homotopically identical (i.e., has the form σ_1 where

$$\sigma_t : Y \to Y$$

is some deformation of the space Y), the map $f \circ h \circ f : A \to Y$ is homotopic to the map $f : A \to Y$, and therefore, according to the previous proposition, the spaces $X \cup_f Y$ and $X \cup_{g\circ h\circ f} Y$ are homotopically equivalent. Moreover, from the construction imposed in the proof of the last proposition it follows that the homotopy equivalence

$$\eta : X \cup_{g\circ h\circ f} Y \to X \cup_f Y$$

connecting the spaces $X \cup_{g\circ h\circ f} Y$ and $X \cup_f Y$, can be defined by the formula

$$\eta = \alpha_1 \circ (\alpha')^{-1},$$

where α_1 is a finite map of some homotopy

$$\alpha_t : X \cup Y \to X \cup_f Y, \quad 0 \le t \le 1,$$

for which

$$\alpha_0 = \alpha, \quad \alpha_t|_A = \sigma_t \circ f,$$

and

$$\alpha_t|_Y = 1_Y.$$

Therefore, the formula

$$k_t = \begin{cases} \alpha \circ \overline{\sigma}_{2t}, & \text{if} \quad 0 \le t \le 1/2, \\ \alpha_{2t-1} \circ \overline{g} \circ h), & \text{if} \quad 1/2 \le t \le 1. \end{cases}$$

where

$$\overline{\sigma}_t = 1_X \cup \sigma_t : X \cup Y \to X \cup Y, \quad 0 \le t \le 1,$$

defines some homotopy

$$k_t : X \cup Y \to X \cup_f Y, \quad 0 \le t \le 1,$$

connecting the map

$$k_0 = \alpha \circ \overline{\sigma}_0 = \alpha$$

with the map

$$k_1 = \alpha_1 \circ \overline{g} \circ \overline{h} = \eta \circ G \circ H \circ \alpha.$$

In this case

$$k_t|_A = \begin{cases} \alpha, & \text{if} \quad 0 \le t \le 1/2, \\ \alpha_{2t-1}, & \text{if} \quad 1/2 \le t \le 1, \end{cases}$$

i. e,

$$k_t|_A = \begin{cases} f, & \text{if} \quad 0 \le t \le 1/2, \\ \sigma_{2t-1} \circ f, & \text{if} \quad 1/2 \le t \le 1. \end{cases}$$

Therefore, assuming

$$k_t^* = k_t \circ \alpha^{-1}, \quad 0 \le t \le 1,$$

we obtain a deformation

$$k_t^* : X \cup_f Y \to X \cup_f Y, \quad 0 \le t \le 1,$$

of the space $X \cup_f Y$, connecting the identity map of this space with the map $\eta \circ G \circ H$.

Thus, the last map is homotopy identical and therefore the map H is homotopy injective, and the map $\eta \circ G$, and therefore the map G, is homotopy surjective.

Similarly (taking the maps $h \circ f$ and g as the maps f and h respectively), we obtain that the map G is homotopy injective. Consequently, the map G, and therefore the map $\eta \circ G$, is a homotopy equivalence. But then the map H will also be a homotopy equivalence. To complete the proof, it remains to note that

$$H|_Y = h.$$

□

2.4 Homotopy groups and weak homotopy equivalences

Let $n \geq 0$ and let $\boldsymbol{u}_0$ be a point $(1, 0, \ldots, 0)$ of the unit n-dimensional sphere $\mathbb{S}^n$. Let, further, x_0 be an arbitrary point of the topological space X. The subject of our study in this section will be continuous maps

$$\varphi : (\mathbb{S}^n, \boldsymbol{u}_0) \to (X, x_0)$$

of the pair $(\mathbb{S}^n, \boldsymbol{u}_0)$ into the pair (X, x_0). The set $[(\mathbb{S}^n, \boldsymbol{u}_0), (X, x_0)]$ of all homotopy classes of such maps relative to x_0 will be denoted by the symbol

$$\pi_n(X; x_0).$$

The class of the constant map $\mathbb{S}^n \to x_0$ we will denote by the symbol 0_{x_0} (or simply 0) and we will call it the *zero* of the set $\pi_n(X; x_0)$.

It is clear that the set $\pi_0(X; x_0)$ is naturally identified with the set $\pi_0(X)$ of the path-connected components of the space X. With this identification, the zero 0_{x_0} of the set $\pi_0(X; x_0)$ corresponds to the component of the space X containing the point x_0.

We will call a map of one set of the form $\pi_n(X; x_0)$ into another such set a *homomorphism* if it maps zero to zero. We will call an injective homomorphism a *monomorphism*, a surjective homomorphism an *epimorphism*, and a bijective homomorphism an *isomorphism*.

Remark 2.31. For $n > 0$, an algebraic operation can be introduced into the set $\pi_n(x; x_0)$, with respect to which this set turns out to be a group (for $n > 1$, even an Abelian group) with zero element 0_{x_0}. This group is called the *n-th homotopy group of the space X at the point x_0*. In what follows, we will also adhere to this terminology (even for $n = 0$), although we do not need the mentioned operation, and we will neither define nor consider it. All homomorphisms of homotopy groups considered below will in fact be homomorphisms in the usual group-theoretical sense. We will also ignore this fact.

According to what was said in §2.1, any continuous map

$$f : X \to Y$$

defines by the formula

$$f_*([\varphi] \operatorname{rel} \boldsymbol{u}_0) = [f \circ \varphi] \operatorname{rel} \boldsymbol{u}_0,$$

where φ is an arbitrary map $(\mathbb{S}^n, \boldsymbol{U}_0) \to (X; x_0)$, some map

$$f_* : \pi_n(X; x_0) \to \pi_n(Y; y_0), \quad y_0 = f(x_0).$$

This maps zero 0_{x_0} to zero 0_{y_0}, i.e. is a homomorphism (in the sense indicated above). In those cases where it is necessary to explicitly indicate the point x_0, we will denote this homomorphism by the symbol f_{*,x_0}.

It is clear that if $f = 1_X$, then $f_* = 1_{\pi_n(X;x_0)}$, and that for any maps $f : X \to Y$ and $g : Y \to Z$ we have

$$(g \circ f)_* = g_* \circ f_*.$$

In the language of category theory, this means that the pair $(\pi_n(X; x_0), f_*)$ is a covariant functor.

We will not present here all the numerous properties of homotopy groups - this would take us too far from the main topic. We will limit ourselves to only those properties of these groups that will be needed later. In this case, for any $n \geq 0$ we will represent the sphere $\mathbb{S}^n$ as a union of two hemispheres - the "lower" hemisphere $\mathbb{E}^n_-$ consisting of all points $\boldsymbol{u} = (u^1, \ldots, u^{n+1}) \in \mathbb{S}^n$, for which $u^{n+1} \leq 0$, and the "upper" hemisphere $\mathbb{E}^n_+$, consisting of all points $\boldsymbol{u} = (u^1, \ldots, u^{n+1}) \in \mathbb{S}^n$, for which $u^{n+1} \geq 0$. The intersection $\mathbb{E}^n_- \cap \mathbb{E}^n_+$ of these hemispheres is the *equator* $u^{n+1} = 0$ of the sphere $\mathbb{S}^n$, which we will identify with the sphere $\mathbb{S}^{n-1}$. The projection

$$\omega : \mathbb{S}^n \to \mathbb{E}^n,$$

defined by the formula

$$\omega(u^1, \ldots, u^{n+1}) = (u^1, \ldots, u^n),$$

is, obviously, a homeomorphism on each of the hemispheres $\mathbb{E}^n_-$ and $\mathbb{E}^n_+$. We will denote these homeomorphisms by the symbols ω_- and ω_+, respectively.

In what follows we will constantly use the fact that

Proposition 2.32. *pairs $(\mathbb{S}^n, \boldsymbol{u}_0)$, $(\mathbb{S}^n, \mathbb{E}^n_+)$, $(\mathbb{S}^n \times I(\mathbb{S}^n, \boldsymbol{u}_0))$ and $(\mathbb{S}^n \times I(\mathbb{S}^n, \mathbb{E}^n_+))$ satisfy the axiom of homotopy extension.*

This fact is a special case of one general statement, which we will prove in §3.4; see Remark 3.56. Therefore, we will leave it here without proof. We will not use the results of this section until §3.6.

We also note that any vector $\boldsymbol{v} \in \mathbb{E}^n$ can be "related to a point $\boldsymbol{u}_0$", i.e. represented in the form

$$\boldsymbol{v} = \boldsymbol{u}_0 + (\boldsymbol{u} - \boldsymbol{u}_0)r,$$

where $\boldsymbol{u} \in \mathbb{S}^{n-1}$ and $0 \le r \le 1$. In this case, the number r is determined by the vector $\boldsymbol{v}$ uniquely. The vector $\boldsymbol{u}$ is also determined uniquely, if only $\boldsymbol{v} \ne \boldsymbol{u}_0$. In addition, for $\boldsymbol{v} \ne \boldsymbol{u}_0$ the number r is nonzero (and for $\boldsymbol{v} = \boldsymbol{u}_0$ it is zero). To simplify the formulae, we will henceforth denote the vector $\boldsymbol{v} = \boldsymbol{u}_0 + (\boldsymbol{u} - \boldsymbol{u}_0)r \in \mathbb{E}^n$ by the symbol $[\boldsymbol{u}, r]$, and the vectors $\omega_+^{-1}(\boldsymbol{v}) \in \mathbb{E}_+^n$ and $\omega_-^{-1}(\boldsymbol{v}) \in \mathbb{E}_-^n$ by the symbols $[\boldsymbol{u}, r]_+$ and $[\boldsymbol{u}, r]_-$, respectively.

First of all, we will find out under what conditions the two maps

$$f, g : \mathbb{E}^n \to X$$

of the ball $\mathbb{E}^n$ into the space X, coinciding on its boundary $\mathbb{S}^{n-1}$, are homotopic relative to $\mathbb{S}^{n-1}$.

For this purpose, to any two such maps f and g we assign a map

$$\varphi : \mathbb{S}^n \to X$$

of the sphere $\mathbb{S}^n$ into the space X, "glued" from maps f and g, considered as maps of the hemispheres $\mathbb{E}_+^n$ and $\mathbb{E}_-^n$, i.e. defined by the equalities

$$\varphi_{\mathbb{E}_-^n} = f \circ \omega_-, \quad \varphi_{\mathbb{E}_+^n} = f \circ \omega_+.$$

By hypothesis, $f|_{\mathbb{S}^{n-1}} = g|_{\mathbb{S}^{n-1}}$ and, therefore, the map φ is uniquely defined and continuous. The element $[\varphi]$ of the group $\pi_n(X; x_0)$, where $x_0 = \varphi(\boldsymbol{u}_0)(= f(\boldsymbol{u}_0) = g(\boldsymbol{u}_0))$, defined by the map φ, we will call the element *distinguishing* the maps f and g, and denote it by the symbol $\delta(f, g)$.

It is easy to see that any element of the set $\pi_n(X; x_0)$ can serve as an element that distinguishes some map from a given one. Namely,

Proposition 2.33. *for any map*

$$g : \mathbb{E}^n \to X$$

and any element $\alpha \in \pi_n(x; x_0)$, *where* $x_0 = g(\boldsymbol{u}_0)$, *there exists a map*

$$f : \mathbb{E}^n \to X,$$

such that

$$f|_{\mathbb{S}^n} = g|_{\mathbb{S}^n}$$

and

$$\delta(f, g) = \alpha.$$

Proof. Indeed, let

$$\varphi : (\mathbb{S}^n, \boldsymbol{u}_0) \to (X, x_0)$$

be an arbitrary map of class α. Obviously, on the hemisphere $\mathbb{E}_+^n$ the map φ is homotopic relatve to $\boldsymbol{u}_0$ to the map $\boldsymbol{g} \circ \omega_+$. And the corresponding homotopy

$$\varphi_t : \mathbb{E}_+^n \to X$$

can be defined, for example, by the formula

$$\varphi_t([\boldsymbol{u}, r]_+) = \begin{cases} \varphi([\boldsymbol{u}, (1-2t)r]_+), & \text{if} \quad 0 \le t \le 1/2, \\ f([\boldsymbol{u}, (2t-1)r]), & \text{if} \quad 1/2 \le t \le 1, \end{cases}$$

where $[\boldsymbol{u}, r]_+$ is an arbitrary point of the hemisphere $\mathbb{E}^n_+$; the essence of the matter here is that the hemisphere $\mathbb{E}^n_+$ can be contracted in itself to the point $\boldsymbol{u}_0$ and therefore any map of this hemisphere is homotopic to the constant map, so that any two maps are homotopic. Since the pair $(\mathbb{S}^n, \mathbb{E}^n_+)$ satisfies, as was said, the axiom of homotopy extension, it follows that the map φ is homotopic relative to $\boldsymbol{u}_0$ to a map

$$\psi : \mathbb{S}^n \to X$$

such that $\psi|_{\mathbb{E}^n_+} = g \circ \omega_+$. But then it is clear that the map

$$f = \psi|_{\mathbb{E}^n_-} \circ \omega_-^{-1} : \mathbb{E}^n \to X$$

has all the required properties. □

The significance of distinguishing elements $\delta(f, g)$ for the problem of homotopy relative to $\mathbb{S}^{n-1}$ of maps f and g is determined by the fact that

Proposition 2.34. *if for maps coinciding on* $\mathbb{S}^{n-1}$

$$f, g : \mathbb{E}^n \to X,$$

there exists a map

$$h : \mathbb{E}^n \to X, \quad h|_{\mathbb{S}^{n-1}} = f|_{\mathbb{S}^{n-1}} = g|_{\mathbb{S}^{n-1}},$$

such that

$$\delta(f, h) = \delta(g, h),$$

then the maps f *and* g *are homotopic relative to* $\mathbb{S}^n$.

Proof. To prove this statement, we consider the maps

$$\varphi_0, \varphi_1 : \mathbb{S}^n \to X,$$

defined by the formulae

$$\varphi_0|_{\mathbb{E}^n_-} = f \circ \omega_-, \quad \varphi_1|_{\mathbb{E}^n_-} = g \circ \omega_-, \quad \varphi_0|_{\mathbb{E}^n_+} = \varphi_1|_{\mathbb{E}^n_+} = h \circ \omega_+.$$

As is easy to see, it is enough for us to prove that

$$\varphi_0 \sim \varphi_1 \operatorname{rel} \mathbb{E}^n_+. \tag{2.35}$$

Indeed, for any homotopy $\varphi_t : \varphi_0 \sim \varphi_1 \operatorname{rel} \mathbb{E}^n_+$ the family of maps

$$\varphi_t|_{\mathbb{E}^n_-} \circ \omega^{-1} : \mathbb{E}^n \to X, \quad 0 \le t \le 1,$$

will be a homotopy relatiive to $\mathbb{S}^{n-1}$, connecting the map f with the map g.

By hypothesis, $\delta(f,h) = \delta(g,h)$, so that the maps φ_0 and φ_1 are homotopic relative to $\boldsymbol{u}_0$. Having chosen some homotopy relative to $\boldsymbol{u}_0$ that connects the maps φ_0 and φ_1, we consider the map

$$\Phi : \mathbb{S}^n \times I \to X$$

of the product $\mathbb{E}^n_+ \times I$ into the space X corresponding to this homotopy. Let, in addition,

$$\Psi : \mathbb{E}^n_+ \times I \to X$$

be a map of the product $\mathbb{E}^n_+ \times I$ into the space X defined by the formula

$$\Psi(\boldsymbol{u}, t) = \varphi_0(\boldsymbol{u}) = (h \circ \omega_+)(\boldsymbol{u}), \quad (\boldsymbol{u}) \in \mathbb{E}^n_-, 0 \le t \le 1.$$

According to Proposition 2.14, it is applicable, since the pair $(\mathbb{S}^n \times I, I(\mathbb{S}^n, \mathbb{E}^n_+))$ satisfies the axiom of homotopy extension, to prove relation (2.35) it is sufficient to show that

$$\Phi_0|_{\mathbb{E}^n_+ \times I} \sim \Psi \operatorname{rel}(\mathbb{E}^n_+ \times 0 \cup \mathbb{E}^n_+ \times 1).$$

For any $\tau \in I$ and any point $([\boldsymbol{u}, r]_+, t) \in \mathbb{E}^n_+ \times I$, where $\boldsymbol{u} \in \mathbb{S}^{n-1}$, $r, t \in I$, we set (see Fig. 2.2)

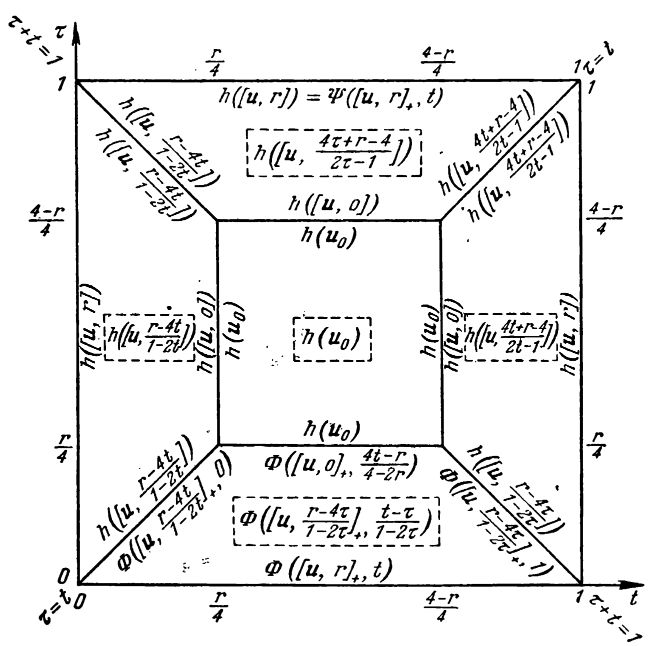


Figure 2.2:

$$\Phi_\tau([\boldsymbol{u},r]_+,t) = \begin{cases} \Phi\left([\boldsymbol{u},\frac{r-4\tau}{1-2\tau}]_+,\frac{t-\tau}{1-2\tau}\right), & \text{if } 0 \le \tau \le \frac{r}{4}, \tau \le t \le 1-\tau, \\ h\left([\boldsymbol{u},\frac{r-4t}{1-2t}]\right), & \text{if } t \le \tau \le 1-t, 0 \le t \le \frac{r}{4}, \\ h(\boldsymbol{u}_0), & \text{if } \frac{r}{4} \le \tau \le \frac{4-r}{4}, \frac{r}{4} \le t \le \frac{4-r}{4}, \\ h\left([\boldsymbol{u},\frac{4t+r-4}{2t-2}]\right), & \text{if } 1-t \le \tau \le t, \frac{4-r}{4} \le t \le 1, \\ h\left([\boldsymbol{u},\frac{4\tau+r-4}{2\tau-1}]\right), & \text{if } \frac{4-r}{4} \le \tau \le 1, 1-\tau \le t \le \tau. \end{cases}$$

It is easy to verify (Fig. 2.2) that we thereby obtain a certain homotopy

$$\Phi_\tau : \mathbb{E}^n_+ \times I \to X$$

relative to $\mathbb{E}^n_+ \times 0 \cup \mathbb{E}^n_+ \times 1$, for which

$$\Phi_0 = \Phi|_{\mathbb{E}^n_+ \times I}, \quad \Phi_1 = \Psi.$$

The statement formulated above is thus completely proven. □

Remark 2.36. For the existence of a homotopy Φ_τ it is essential that the map φ_0 is homotopic to the map φ_1 relative to $\boldsymbol{u}_0$. If φ_0 is simply homotopic to φ_1, then the homotopy Φ_τ may not exist.

Let us now consider the question of the dependence of the group $\pi_n(X; x_0)$ on the point x_0.

Let x_0 and x_1 be two arbitrary points of the space X, which can be connected in X by some path $u : I \to X$. Since the pair $(\mathbb{S}^n, \boldsymbol{u}_0)$ satisfies the axiom of homotopy extension, then for any map

$$\varphi : (\mathbb{S}^n, \boldsymbol{u}_0) \to (X, x_0)$$

there exists a homotopy

$$\varphi_t : \mathbb{S}^n \to X, \quad 0 \le t \le 1,$$

such that $\varphi_0 = \varphi$ and $\varphi_t(\boldsymbol{u}_0) = u(t)$ for each $t \in I$. We will call this homotopy a *homotopy of the map* φ *along the path* u.

We will call two paths u and v, connecting a point x_0 with a point x_1 *equivalent* if they are homotopic relative to the points 0 and 1, i.e. if there exists a homotopy

$$u_\tau : I \to X, \quad 0 \le \tau \le 1,$$

consisting of paths connecting the points x_0 and x_1 such that $u_0 = u$ and $u_1 = v$.

It turns out that

Proposition 2.37. *if the maps*

$$\varphi, \psi : (\mathbb{S}^n, \boldsymbol{u}_0) \to (X, x_0)$$

are homotopic relative to $\boldsymbol{u}_0$*, then for any two equivalent paths* u *and* v *connecting a point* x_0 *with a point* x_1*, and any two homotopies*

$$\varphi_t, \psi_t : \mathbb{S}^n \to X$$

of the maps φ *and* ψ *along, respectively, the paths* u *and* v *the pointed maps*

$$\varphi_1, \psi_1 : (\mathbb{S}^n, \boldsymbol{u}_0) \to (X, x_1)$$

are homotopic relative to $\boldsymbol{u}_0$.

Proof. Indeed, let

$$\xi_t : (\mathbb{S}^n, \boldsymbol{u}_0) \to (X, x_0), \quad 0 \le t \le 1,$$

be a homotopy relative to $\boldsymbol{u}_0$, connecting the map φ with the map ψ. It is clear that the formula

$$\omega_t = \begin{cases} \varphi_{1-3t}, & \text{if} \quad 0 \le t \le 1/3, \\ \xi_{3t-1}, & \text{if} \quad 1/3 \le t \le 2/3, \\ \psi_{3t-2}, & \text{if} \quad 2/3 \le t \le 1, \end{cases}$$

defines a certain homotopy

$$\omega_t : \mathbb{S}^n \to X, \quad 0 \le t \le 1,$$

connecting the map φ_1 with the map ψ_1. Let

$$\Omega : \mathbb{S}^n \times I \to X$$

be the map of the product $\mathbb{S}^n \times I$ into the space X corresponding to this homotopy and let

$$\Omega_1 : \boldsymbol{u}_0 \times I \to X$$

be the constant map of the segment $\boldsymbol{u}_0 \times I$ into the point $x_1 \in X$. Since the pair $(\mathbb{S}^n \times I, I(\mathbb{S}^n, \boldsymbol{u}_0))$ satisfies the axiom of homotopy extension, then, according to the proposition proved in §point 2.1, to prove the relation

$$\varphi_1 \sim \psi_1 \operatorname{rel} \boldsymbol{u}_0$$

it suffices to prove that

$$\Omega|_{\boldsymbol{u}_0 \times I} \sim \Omega_1 \operatorname{rel}(\boldsymbol{u}_0 \times 0 \cup \boldsymbol{u}_0 \times 1),$$

i.e., the path

$$w : I \to X, \quad w(0) = w(1) = x_1,$$

defined by the formula

$$w(t) = \Omega(\boldsymbol{u}_0, t) = \begin{cases} u(1-3t), & \text{if} \quad 0 \le t \le 1/3, \\ x_0, & \text{if} \quad 1/3 \le t \le 2/3, \\ v(3t-2), & \text{if} \quad 2/3 \le t \le 1, \end{cases}$$

is equivalent to the degenerate path u_{x_1}, which is a constant map of the segment I to the point x_1.

For this purpose, for any $t, \tau \in I$ we set (see Fig. 2.3)

$$w_\tau(t) = \begin{cases} u\left(\frac{3\tau-6t+2}{3\tau+2}\right), & \text{if} \quad 0 \le \tau \le 1/3, 0 \le t \le \frac{3\tau+2}{6}, \\ x_0, & \text{if} \quad 0 \le \tau \le 1/3, \frac{3\tau+2}{6} \le t \le \frac{4-3\tau}{6}, \\ v\left(\frac{3\tau+6t-4}{3\tau+2}\right), & \text{if} 0 \le \tau \le 1/3, \frac{4-3\tau}{6} \le t \le 1, \\ u(1-2t), & \text{if} \quad 1/3 \le \tau \le 2/3, 0 \le t \le 1/2, \\ u_{2-3\tau}(2t-1), & \text{if} \quad 1/3 \le \tau \le 2/3, 1/2 \le t \le 1, \\ u(1-6t(1-\tau)), & \text{if} \quad 2/3 \le \tau \le 1, 0 \le t \le 1/2, \\ u(1-6(1-t)(1-\tau)), & \text{if} \quad 2/3 \le \tau \le 1, 1/2 \le t \le 1, \end{cases}$$

where $u_\tau : I \to X$, $0 \le \tau \le 1$, is a homotopy relative to the points 0 and 1, connecting the path u with the path v. It is easy to verify that we thereby

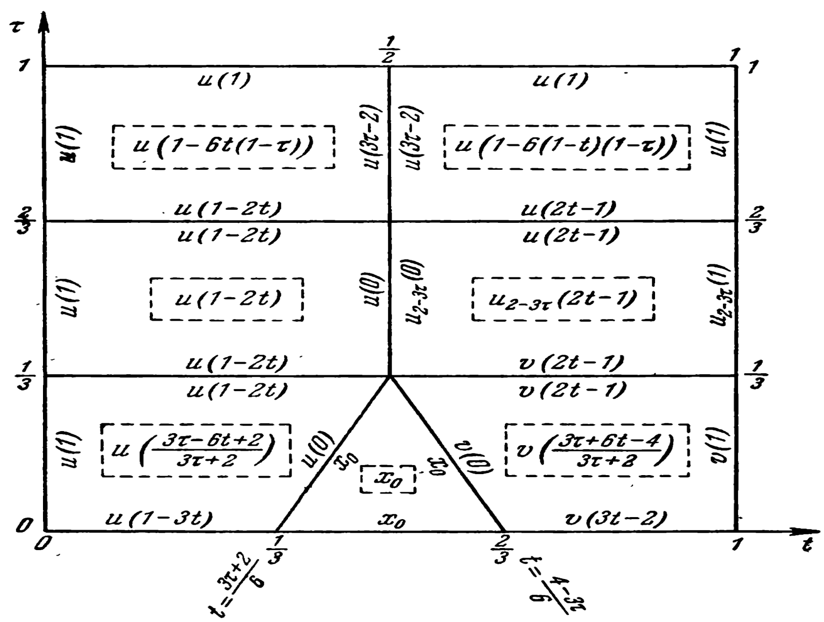


Figure 2.3:

obtain a homotopy

$$w_t : I \to X, \quad 0 \le \tau 1,$$

relative to the points 0 and 1, connecting the path w with the path u_{x_1}.

The above statement is thus completely proven. □

It follows directly from this statement that for any path u, connecting points x_0 and x_1 the formula

$$u^\#(\alpha) = [\varphi_1] \operatorname{rel} \boldsymbol{u}_0,$$

where $\alpha \in \pi_n(X;x_0)$, and φ_1 is a pointed map of an arbitrary homotopy φ_t along the path u of some map

$$\varphi : (\mathbb{S}^n, \boldsymbol{u}_0) \to (X; x_0)$$

of class α uniquely determines some map

$$u^\# : \pi_n(X;x_0) \to \pi_n(X;x_1),$$

mapping zero 0_{x_0} to zero 0_{x_1}, i.e. being a homomorphism. At the same time,

Proposition 2.38. *the homomorphism $u^\#$ depends only on the equivalence class of the path u, i.e. for any two equivalent paths u and v the homomorphisms $u^\#$ and $v^\#$ coincide.*

Now let u and v be paths of the space X such that

$$u(1) = v(0).$$

Then the formula

$$w(t) = \begin{cases} u(2t), & \text{if} \quad 0 \le t \le 1/2, \\ v(2t-1), & \text{if} \quad 1/2 \le t \le 1, \end{cases}$$

defines, as is easy to see, a certain path

$$w : I \to X,$$

connecting the point $x_0 = u(0)$ with the point $x_1 = v(1)$. We will call this path the *product* of paths u and v and will denote it by the symbol uv. It is easy to see that

Proposition 2.39. *the homomorphism*

$$(uv)^\# : \pi_n(X;x_0) \to \pi_n(X;x_2),$$

corresponding to the path uv, is a composition of $v^\# \cdot u^\#$ of homomorphisms

$$u^\# : \pi_n(X;x_0) \to \pi_n(X;x_1), \quad v^\# : \pi_n(X;x_1) \to \pi_n(X;x_2),$$

corresponding to the paths u and v.

Proof. Indeed, for any homotopy φ_t along the path u of an arbitrary map $\varphi : (\mathbb{S}^n, \boldsymbol{u}_0) \to (X;x_0)$ and any homotopy ψ_t, along the path v of an arbitrary map

$$\psi : (\mathbb{S}^n, \boldsymbol{u}_0) \to (X;x_1)$$

the formula

$$\omega_t = \begin{cases} \varphi(2t), & \text{if} \quad 0 \le t \le 1/2, \\ \psi(2t-1), & \text{if} \quad 1/2 \le t \le 1, \end{cases}$$

defines a homotopy along the path uv for which $\omega_0 = \varphi$ and $\omega_1 = \psi_1$. □

To each path $u : I \to X$ connecting a point x_0 with a point x_1 we assign a path $u^{'} : I \to X$ connecting the point x_1x with the point x_0, defined by the formula

$$u^{'}(t) = u(1-t), \quad 0 \le t \le 1.$$

It is easy to see that

Proposition 2.40. *for any path u connecting points x_0 and x_1 the paths $uu^{'}$ and $u^{'}u$ are equivalent to the corresponding degenerate paths u_{x_0} and u_{x_1}.*

Proof. Indeed, a homotopy $v_\tau : I \to X$, $0 \le \tau \le 1$, relative to the points 0 and 1, connecting, say, the path u_{x_0} with the path $uu^{'}$ can, for example, be defined by the formula

$$v_\tau(t) = \begin{cases} u(2t\tau), & \text{if} \quad 0 \le t \le 1/2, \\ u(2(1-t)\tau), & \text{if} \quad 1/2 \le t \le 1, \end{cases} \quad 0 \le \tau \le 1.$$

□

Since the homomorphisms $(u_{x_0})^{\#}$ and $(u_{x_1})^{\#}$ corresponding to the degenerate paths u_{x_0} and u_{x_1} are, as is easy to see, identity maps of the groups $\pi_n(X;x_0)$ and $\pi_n(X;x_1)$ respectively, it follows directly from the last two statements that

Proposition 2.41. *for any path $u :\to X$ connecting the point x_0 with the point x_1 the homomorphisms $u^{\#}$ and $(u^{'})^{\#}$ are mutually inverse isomorphisms between the groups $\pi_n(X;x_0)$ and $\pi_n(X;x_1)$.*

In particular,

Proposition 2.42. *for any path u connecting a point x_0 with a point x_1, the homomorphism*

$$u^{\#} : \pi_n(X;x_0) \to \pi_n(X;x_1)$$

is an isomorphism.

Thus, for any points x_0 and x_1 belonging to the same path component of the space X, the groups $\pi_n(X;x_0)$ and $\pi_n(X;x_1)$ are essentially the same. For points belonging to different path components of the space X, these groups, generally speaking, are not connected with each other in any way.

In order to clarify the geometric meaning of the isomorphisms $u^{\#}$, we will show that

Proposition 2.43. *for elements $\alpha \in pi_n(X;x_0)$ and $\beta \in \pi_n(X;x_1)$ if and only if there exists a path $u : I \to X$, $u(0) = x_0$, $u(1) = x_1$, such that*

$$\beta = u^{\#}(\alpha),$$

if and only if when the maps

$$\varphi : (\mathbb{S}^n, \boldsymbol{u}_0) \to (X;x_0), \quad \psi : (\mathbb{S}^n, \boldsymbol{u}_0) \to (X;x_1),$$

belonging to the classes α and β are homotopic to each other.

Proof. Indeed, if $\beta = u^{\#}(\alpha)$, then the map ψ is homotopic (relative to $\boldsymbol{u}_0$) to a pointed map φ_1 of some homotopy φ_t of the map φ along the path u. Thus, $\varphi_t : \varphi \sim \psi$ and, consequently, $\psi \sim \varphi$. Conversely, any homotopy $\varphi_t : \varphi \sim \psi$ can be considered as a homotopy of the map φ along the path

$$u(t) = \varphi_t(\boldsymbol{u}_0), \quad 0 \le t \le 1.$$

Therefore, if $\varphi_t : \varphi \sim \psi$, then $\beta = u^{\#}(\alpha)$. □

Since for each path u the map $u^{\#}$ is an isomorphism, the only element $\alpha \in \pi_n(X; x_0)$ for which $u^{\#}(\alpha) = 0_{x_1}$ is the element 0_{x_0}. This means that

Proposition 2.44. *the map*

$$\varphi : (\mathbb{S}^n, \boldsymbol{u}_0) \to (X; x_0)$$

is homotopic to some constant map if and only if it is homotopic relative to $\boldsymbol{u}_0$ *to the constant map to the point* x_0.

In other words (see §2.1),

Proposition 2.45. *the map*

$$\varphi : (\mathbb{S}^n, \boldsymbol{u}_0) \to (X; x_0)$$

belongs to the class 0_{x_0} *if and only if when it can be extended to some map* $\mathbb{E}^n \to X$.

Therefore,

Proposition 2.46. *the space* X *is* m*-connected if and only if for some (and therefore for any) point* $x_0 \in X$ *and any* $n \le m$ *the equality*

$$\pi_n(X; x_0) = 0$$

holds.

Let us now consider an arbitrary deformation.

$$\xi_t : X \to X$$

of the space X. For any point $x_0 \in X$ this deformation defines a path

$$u(t) = \xi_t(x_0), \quad 0 \le t \le 1,$$

connecting the point x_0 with the point $x_1 = \xi_1(x_0)$. On the other hand, it is clear that for any map

$$\varphi : (\mathbb{S}^n, \boldsymbol{u}_0) \to (X; x_0)$$

the family of maps

$$\xi_t \circ \varphi : \mathbb{S}^n \to X$$

represents a homotopy of the map φ along the path u. Therefore, for any element $\alpha \in \pi_n(X; x_0)$, the element $u^{\#}(\alpha) \in \pi_n(X; x_1)$ is the class of the map $\xi \circ \varphi$, where $\xi = \xi_1$, and φ is an arbitrary map of class α. This means that

Proposition 2.47. *the homomorphism*

$$\xi_* : \pi_n(X; x_0) \to \pi_n(X; x_1)$$

coincides with the homomorphism

$$u^\# : \pi_n(X; x_0) \to \pi_n(X; x_1)$$

Since the last homomorphism is, as we know, an isomorphism, it is thus proved that

Proposition 2.48. *for any homotopy identity map*

$$\xi : X \to X$$

and any point $x_0 \in X$, the homomorphism

$$\xi_* : \pi_n(X; x_0) \to \pi_n(X; x_1), \quad x_1 = \xi(x_0),$$

is an isomorphism.

A map

$$f : X \to Y$$

of a topological space X into a topological space Y will be called a *weak homotopy equivalence* if for any $n \geq 0$ and any point $x_0 \in X$ the homomorphism

$$f_* : \pi_n(X; x_0) \to \pi_n(Y; y_0), \quad y_0 = f(x_0),$$

is an isomorphism. From the proposition just proved it easily follows that

Proposition 2.49. *any homotopy equivalence*

$$f : X \to Y$$

is a weak homotopy equivalence.

Proof. Indeed, let

$$g : Y \to X$$

be the homotopy equivalence inverse to the equivalence f, and let x_0 be an arbitrary point in X. Setting

$$y_0 = f(x_0), \quad x_1 = g(y_0), \quad y_1 = f(x_1),$$

for any $n \geq 0$ we consider the homomorphisms

$$\begin{aligned} f_* = f_{*,x_0} &: \pi_n(X; x_0) \to \pi_n(Y; y_0), \\ f'_* = f_{*,x_1} &: \pi_n(X; x_1) \to \pi_n(Y; y_1), \\ g_* = g_{*,y_0} &: \pi_n(Y; y_0) \to \pi_n(X; x_1), \end{aligned}$$

induced by the maps f and g, as well as the homomorphsims

$$(g \circ f)_* : \pi_n(X; x_0) \to \pi_n(X; x_1),$$
$$(f^{'} \circ g)_* : \pi_n(Y; y_0) \to \pi_n(Y; y_1),$$

induced by composite maps

$$g \circ f : X \to X, \quad f^{'} \circ g : Y \to Y.$$

As we know,

$$(g \circ f)_* = g_* \circ g_*, \quad (f^{'} \circ g)_* = f^{'} \circ g_*.$$

On the other hand, since the maps $g \circ f$ and $f^{'} \circ g$ are, by hypothesis, homotopically identical, the homomorphisms $(g \circ f)_*$ and $(f^{'} \circ g)_*$ are isomorphisms. Therefore, from the equality $(g \circ f)_* = g_* \circ f*$ it follows that the map g_* is epimorphic, and from the equality $(f^{'} \circ g)_* = f^{'}_* \circ g_*$ it follows that the map g_* is monomorphic. Consequently, the map g_*, and therefore the map f_*, is an isomorphism. □

Remark 2.50. The converse is generally not true: there are weak homotopy equivalences that are not homotopy equivalences. However, as we shall see below, for sufficiently "good" spaces (namely cellular spaces), any weak homotopy equivalence is a homotopy equivalence.

It is clear that the composition of weak homotopy equivalences is also a weak homotopy equivalence. Therefore, it follows immediately from the previous statement that

Proposition 2.51. *any map that is homotopy equivalent to a weak homotopy equivalence is also a weak homotopy equivalence.*

Remark 2.52. By analogy with the relation of homotopy equivalence of spaces, one could introduce the relation of their weak homotopy equivalence. This relation is obviously reflexive and transitive, but, generally speaking, it is not symmetric. We will not consider it.

We will call a subspace A of a space X *representative* if the inclusioon map $i : A \subset X$ is a weak homotopy equivalence, i.e. if for any $n \geq 0$ and any point $x_0 \in A$ the homomorphism

$$i_* : \pi_n(A; x_0) \to \pi_n(X; x_0)$$

is an isomorphism. It follows directly from the proposition just proved that

Proposition 2.53. *the map $f : X \to Y$ is a weak homotopy equivalence if and only if the space X is a representative subspace of the cylinder Z_f of the map f.*

In connection with this proposition, it is useful to note that

Proposition 2.54. *if a subspace A of X is representative, then for any $n \geq 0$ each map*

$$f : (\mathbb{E}^n, \mathbb{S}^{n-1}) \to (X, A)$$

is homotopic relative to $\mathbb{S}^{n-1}$ to some map $g : \mathbb{E}^n \to A$ (i. e, more precisely, to the map $i \circ g$, where $i : A \subset X$).

Proof. Indeed, assuming that $n > 0$ (for $n = 0$ the reasoning is only simplified), we assign to the map f the element α' of the group $\pi_{n-1}(A; x_0$, where $x_0 = f(\boldsymbol{u}_0)$, defined by the map

$$f|_{\mathbb{S}^{n-1}} : \mathbb{S}^{n-1} \to A.$$

Since the map $i \circ f|_{\mathbb{S}^{n-1}}$ is a restriction of the map $f : \mathbb{E}^n \to X$, then, as proved above, $i_*(\alpha') = 0$, and therefore $\alpha' = 0$ (since the homomorphism i_* is, by assumption, an isomorphism). Therefore, there exists a map

$$h : \mathbb{E}^n \to A,$$

such that

$$h|_{\mathbb{S}^{n-1}} = f|_{\mathbb{S}^{n-1}}.$$

Let

$$\beta = \delta(f, i \circ h)$$

be an element of the group $\pi_n(X; x_0)$ that distinguishes the maps f and $i \circ h$. Since the subspace A is representative, in the group $\pi_n(A; x_0)$ there exists an element α such that $i*(\alpha) = \beta$. Let

$$g : \mathbb{E}^n \to A$$

be a map of the ball $\mathbb{E}^n$ into the space A that coincides on $\mathbb{S}^{n-1}$ with the map h such that

$$\delta(g, h) = \alpha.$$

Then, as is easy to see,

$$\delta(i \circ g, i \circ h) = i_*(\alpha),$$

i.e,

$$\delta(i \circ g, i \circ h) = \delta(f, i \circ h).$$

Therefore, the maps f and $i \circ g$ are homotopic relative to $\mathbb{S}^{n-1}$. □

Remark 2.55. The converse is also true: if for each $n \geq 0$ any map $(\mathbb{E}^n, \mathbb{S}^{n-1}) \to (X, A)$ is homotopic relative to $\mathbb{S}^{n-1}$ to some map $\mathbb{E}^n \to A$ then the subspace A is representative. We will not need this fact, and we will leave its proof to the reader as a simple exercise.

2.5 Homotopy limits

Let

$$X_0 \subset X_1 \subset \cdots \subset X_n \subset \cdots \tag{2.56}$$

be an increasing sequence of subspaces of a topological space X, the union of which is the entire space X. Consider the inclusion maps

$$i_m^n : X_n \to X_m, \quad i^n : X_n \to X, \quad 0 \le n \le m \le \infty,$$

and for any point $x \in X$ and any number $n \ge n_x$, where n_x is the smallest n for which $x \in X_n$, the homomorphisms induced by these maps

$$\begin{matrix} (i_m^n)_* : \pi_k(X_n;x) \to \pi_k(X_m;x), \\ (i^n)_* : \pi_k(X_n;x) \to \pi_k(X;x), \end{matrix} \quad k \ge 0,$$

of homotopy groups.

We will say that the space X is a *weak homotopy limit* of subspaces (2.56) if

1) for any number $k \ge 0$, any point $x \in X$, and any element $\alpha \in \pi_k(X;x)$, there exist a number $n \ge n_x$ and an element $\alpha_n \in \pi_k(X_n;x)$ such that

$$\alpha = (i^n)_*(\alpha_n);$$

2) for any number $k \ge 0$, any number $n \ge 0$, any point $x \in X$, and any elements $\alpha, \alpha' \in \pi_k(X;x)$, with the property that

$$(i^n)_*(\alpha_n) = (i^n)_*(\alpha_n'),$$

there exists a number $m \ge n$ such that

$$(i_m^n)_*(\alpha_n) = (i_m^n)_*(\alpha_n'),$$

It is clear that if the number $n \ge n_x$ satisfies condition 1), then any number $n' \ge n$ also satisfies this condition, i.e. there exists an element $\alpha_{n'} \in \pi_k(X_{n'};x)$ such that $(i^{n'})_*(\alpha_{n'}) = \alpha$ (at least the element $\alpha_{n'} = (i_{n'}^n)_*(\alpha n)$ has this property). Similarly, any number $m' \ge m$ satisfies condition 2) together with the number m.

Remark 2.57. The reader familiar with the concept of a spectrum of groups will immediately discover that for any number $k \ge 0$ and any point $x \in X$ the groups $\pi k(X_n;x)$, $n \ge m$, and the homomorphisms $(i_m^n)_*$, $n_x \le n \le m$ constitute a spectrum and that conditions 1) and 2) are equivalent to the fact that the homomorphism of the limit group of this spectrum into the group $\pi_k(X;x)$, induced by the homomorphisms $(i^n)_*$, is an isomorphism.

One can specify simple set-theoretic conditions under which the space X is a weak homotopy limit of the subspaces (2.56). For example,

Proposition 2.58. *if for any compact set $C \subset X$ there exists a number $n \geq 0$ such that $C \subset X_n$, then the space X is a weak homotopy limit of the subspaces X_n.*

Proof. Indeed, let $k \geq 0$, $x \in X$, and let α be an arbitrary element of the group $\pi_k(X;x)$. Consider an arbitrary map

$$f : (\mathbb{S}^k, \boldsymbol{u}_0) \to (X;x)$$

of class α. Since the sphere $\mathbb{S}^k$ is compact, the set $f(\mathbb{S}^k)$ is also compact, and, consequently, there exists a number $n \geq 0$ such that $f(\mathbb{S}^k) \subset X_n$. Therefore, we can consider the map f as a map $f : (\mathbb{S}^k, \boldsymbol{u}_0) \to (X_n;x)$. It is clear that the element $\alpha_n \in \pi_k(X_n;x)$ defined by this map has the property that $(ii^n)_*(\alpha_n) = \alpha$.

Condition 2 is verified similarly (only instead of the compactness of the sphere $\mathbb{S}^k$, we have to use the compactness of the product $\mathbb{S}^k \times I$). □

In the case of an arbitrary sequence (2.56), the space X is not, generally speaking, a weak homotopy limit of the subspaces X_n. However, it can be argued that

Proposition 2.59. *For any increasing sequence*

$$X_0 \subset X_1 \subset \cdots \subset X_n \subset \cdots \tag{2.60}$$

of subspaces of X, there exists a space X^Σ and a sequence

$$X_0^\Sigma \subset X_1^\Sigma \subset \cdots \subset X_n^\Sigma \subset \cdots$$

of subspaces of X^Σ such that

1) for any $n \geq 0$, the space X_n^Σ is homotopy equivalent to the space X_n;

2) the space X^Σ is a weak homotopy limit of the subspaces X_n^Σ.

Proof. Indeed, let X^Σ be the subspace of the product $X \times \mathbb{R}$ consisting of all points (x,t), $x \in X$, $t \in \mathbb{R}$, for which $t \geq n+1$ for $x \notin X_n$,

$$X^\Sigma = \cup_{k=0}^\infty (X_k \times [k, \infty]) = \cup_{k=0}^\infty (X_k \times [k, k+1])$$

and let X_n^Σ be the subspace of X^Σ consisting of all points $(x,t) \in X^\Sigma$ for which $t \leq n$, i.e.,

$$X_n^\Sigma = (X_0 \times [0,1]) \cup (X_1 \times [1,2]) \cup \cdots \cup (X_{n-1} \times [n-1, n]) \cup (X_n \times n).$$

Since the coordinate t of a point (x,t) of an arbitrary compact set $C \subset x^\Sigma$, being a continuous function on C, is bounded, then $C \subset X_n^\Sigma$ for some $n \geq 0$. Consequently, according to the statement proved above, the space X^Σ is a weak homotopy limit of the subspaces X_n^Σ.

Let us further consider the natural map

$$p^\Sigma : X^\Sigma \to X,$$

defined by the formula

$$p^{\Sigma}(x,t) = x, \quad (x,t) \in X^{\Sigma}.$$

It is clear that for any $n \geq 00$ the inclusion

$$p^{\Sigma}(X_n^{\Sigma}) \subset X_n,$$

holds and therefore the map p^{Σ} defines a map

$$p_n^{\Sigma} : X_n^{\Sigma} \longrightarrow X_n.$$

(This map is defined by the same formula as the map p^{Σ}, with the only difference that now $(x,t) \in X_n^{\Sigma}$.) Setting for any point $x \in X$

$$q_n^{\Sigma}(x) = (x,n),$$

we obviously obtain a continuous map

$$q_n^{\Sigma} : X_n \longrightarrow X_n^{\Sigma}.$$

for which $p_n^{\Sigma} \circ q_n^{\Sigma} = 1_{X_n}$. On the other hand, setting

$$\xi_{\tau}(x,t) = (x, t + \tau(n-t)), \quad (x,t) \in X_n^{\Sigma}, 0 \leq \tau \leq 1,$$

we obtain a deformation

$$\xi_{\tau} : X_n^{\Sigma} \longrightarrow X_n^{\Sigma}, \quad 0 \leq \tau \leq 1,$$

of the space X_n^{Σ} such that $\xi_1 = q^{\Sigma} \circ p_n^{\Sigma}$. Therefore, the maps p_n^{Σ} and q_n^{Σ} are mutually inverse homotopy equivalences.

Thus, the proposition formulated above is completely proven. □

Generally speaking, it is impossible to claim that the map p^{Σ} is a homotopy equivalence (even a weak one). Specifically, as we will now show,

Proposition 2.61. *the map*

$$p^{\Sigma} : X^{\Sigma} \to X$$

is a weak homotopy equivalence if and only if the space X is a weak homotopy limit of the subspaces X_n.

We will prove an even more general proposition, which applies to the situation where we are given an arbitrary space X, which is a weak homotopy limit of subspaces

$$X_0 \subset X_1 \subset \cdots \subset X_n \subset \cdots$$

an arbitrary space Y, which is the union of an increasing sequence of subspaces

$$Y_0 \subset Y_1 \subset \cdots \subset Y_n \subset \cdots$$

and a continuous map

$$f : X \longrightarrow Y$$

such that

$$f(X_n) \subset Y_n$$

for any $n \geq 0$. Let

$$f_n : X_n \longrightarrow Y_n, \quad n \geq 0,$$

be the map defined by the map f. This map is related to the map f by the formula

$$f \circ i^n = j^n \circ f_n,$$

where i^n is, as above, the inclusion map $X_n \subset X$, and j^n is the inclusion map $Y_n \subset Y$. Moreover, for $n \leq m$, the maps f_n and f_m are related by the formula

$$f_m \circ i^n = j^n_m \circ f_n,$$

where $i^n_m : X_n \subset X_m$ and $j^n_m : Y_n \subset Y_m$.

We will show that

Proposition 2.62. *if for any $n \geq 0$ the map f_n is a weak homotopy equivalence, then the following two statements are equivalent:*

1) the space Y is a weak homotopy limit of subspaces Y_n;

2) the map f is a weak homotopy equivalence.

Proof. Indeed, let Statement 1) be true. Consider an arbitrary number $k \geq 0$ and an arbitrary point $xinX$. Since the space Y is a weak homotopy limit of subspaces Y_n, then for any element $\beta \in \pi_k(Y; y)$, where $y = f(x)$, there exist a number $n \geq n_y n$ and an element $\beta_n \in \pi_k(Y_n; y)$ such that $\beta = (j^n)_*(\beta_n)$. In this case, without loss of generality, we can assume that $n \geq n_x$, i.e., that the group $\pi_k(X_n; x)$ is meaningful. Since the corresponding homomorphism

$$(f_n)_* : \pi_k(X_n; x) \longrightarrow \pi_k(Y_n; y)$$

is, by assumption, an isomorphism, there exists an element $\alpha_n \in \pi_k(X_n; x)$ such that $\beta_n = (f_n)_*(\alpha_n)$. Consequently, assuming $\alpha = (i^n)_*(\alpha_n)$, we get that

$$f_*(\alpha) = (f \circ i^n)_*(\alpha_n) = (j^n \circ f_n)_*\alpha_n) = (i^n)_*(\beta_n) = \beta.$$

Thus, for any number $k \geq 0$ and any point $x \in X$, the homomorphism

$$f_* : \pi_k(X; x) \longrightarrow \pi_k(Y; f(x)) \tag{2.63}$$

is an epimorphism.

Now let α and α' be elements of the group $\pi_k(X; x)$, such that $f_*(\alpha) = f_*(\alpha')$. Since the space X is a weak homotopy limit of subspaces X_n, there exist a number

$n \geq n_x$ and elements $\alpha_n, \alpha_n^{'} \in \pi_k(X; x)$ such that $\alpha = (i^n)_*(\alpha_n)$ and $\alpha^{'} = (i^n)_*(\alpha_n^{'})$. Let $\beta_n = (f_n)_*(\alpha_n)$ and $\beta_n^{'} = (f_n)_*(\alpha_n^{'})$. Since

$$\begin{aligned}(j^n)_*(\beta_n) &= (j^n \circ f_n)_*(\alpha_n) = (f \circ i^n)_*(\alpha_n)) = f_*(\alpha) \\ &= f_*(\alpha^{'}) = (f \circ i^n)_*(\alpha_n^{'}) = (j^n \circ fn)_*(\alpha_n^{'}) = (j^n)_*(\beta_n^{'}),\end{aligned}$$

then there exists a number $m \geq n$ such that

$$(j_m^n)_*(\beta_n) = (j_m^n)_*(\beta_n^{'}).$$

Therefore, putting $\alpha_m = (i_m^n)_*(\alpha_n)$ and $\alpha_m^{'} = (i_m^n)_*(\alpha_n^{'})$, we get that

$$\begin{aligned}(f_m)_*(\alpha_m) &= (f_m \circ i_m^n)_*(\alpha_n) = (j_m^n \circ f_n)_*(\alpha_n)) = (j_m^n)_*(\beta_n) \\ &= (j_m^n)_*(\beta_n^{'}) = (j_m^n \circ f_n)_*(\alpha_n^{'}) = (f_m \circ i_m^n)_*(\alpha_n^{'}) = (f_m)_*(\alpha_n^{'}).\end{aligned}$$

Since the map $(f_m)_*$ is, by assumption, isomorphic, it follows that $\alpha_m = \alpha_m^{'}$, and therefore

$$\alpha = (i^n)_*(\alpha_n) = (i^m)_*(\alpha_m) = (i^m)_*(\alpha_m^{'}) = (i^m \circ i_m^n)*(\alpha_n^{'}) = \alpha^{'}.$$

Thus, homomorphism (2.63) is also a monomorphism. Therefore, the implication 1) $\Rightarrow$ 2) is completely proved.

Now let assertion 2) be true. We will prove that condition 1) of the definition of a weak homotopy limit is satisfied for the subspaces Y_n, i.e., that for any number $k \geq 0$, any point $y \in Y$, and any element $\beta \in \pi_k(Y; y)$, there exists a number $n \geq n_y$ such that $\beta = (j^n)_*(\beta)$. Since the map f induces, by assumption, a one-to-one correspondence between the path connected components of the space X and the path connected components of the space Y, we can, without loss of generality, assume that $y = f(x)$, where x is some point of the space X. Consider the corresponding homomorphism

$$f_* : \pi_k(X; x) \longrightarrow \pi_k(Y; y).$$

Since this homomorphism is, by assumption, an isomorphism, there exists an element $\alpha \in \pi_k(X; x)$ such that $\beta = f_*(\alpha)$. Since the space X is a weak homotopy limit of subspaces X_n, there exists a number $n \geq n_x \geq n_y$ and an element $\alpha_n \in \pi_k(X; x)$ such that $\alpha = (i^n)_*(\alpha_n)$. Therefore, setting $\beta_n = (f_n)_*(\alpha_n)$, we obtain that:

$$\beta = (f \circ i^n)_*(\alpha_n) = (j^n \circ f_n)_*(\alpha_n) = (j^n)_*(\beta_n).$$

Condition 2) is checked in a completely similar way.

Thus, the above statement is fully proven. □

We will say that the space X is a *homotopy limit* of the sequence of subspaces

$$X_0 \subset X_1 \subset \cdots \subset X_n \subset \cdots \tag{2.64}$$

if the map

$$p^\Sigma : X^\Sigma \to X$$

is a homotopy equivalence. By the proposition just proved,

Proposition 2.65. *any space X that is a homotopy limit of subspaces X_n is also their weak homotopy limit.*

We will begin our study of homotopy limits by proving that

Proposition 2.66. *if*

1) *every point $x \in X$ is an interior point of some subspace X_n, i.e., the family $\{\operatorname{int} X_n; n \geq 0\}$ is an open covering of X;*

2) *the covering $\{\operatorname{int} X_n; n \geq 0\}$ can be refined into a locally finite open covering $\Gamma = \{U_\alpha; \alpha \in A\}$, for which there exists a subordinate partition of unity $\{f_\alpha; \alpha \in A\}$,*

then the space X is a homotopy limit of the subspaces X_n.

Proof. Indeed, denoting for each element $\alpha \in A$ by the symbol $b(\alpha)$ the smallest n for which $U_\alpha \in \operatorname{int} X_n$, and setting

$$f(x) = \sum_{\alpha \in A} n(\alpha) f_\alpha(x), \quad x \in X,$$

we define on the space X a certain numerical function $f(x)$. This function, as is easy to see, is continuous (since the covering Γ, by assumption, is locally finite). Furthermore, it has the property that $f(x) \geq n+1$ for $x \notin X_n$, since if $x \notin X_n$, then $f(x) = 0$ for all $\alpha \in A$ for which $n(\alpha) \leq n$, and therefore

$$f(x) \geq (n+1) \sum_{\alpha \in A} f_\alpha(x) = n+1.$$

Consequently, the formula

$$q(x) = (x, f(x)), \quad x \in X,$$

defines a continuous map

$$q : X \to X^\Sigma.$$

Clearly, $p \circ q = 1_X$. Moreover, the formula

$$\xi_\tau(x,t) = (x, (1-\tau)t + \tau f(x)), \quad (x,t) \in X^\Sigma, 0 \leq \tau \leq 1,$$

defines, as is easy to see, a deformation $\xi_\tau : X^\Sigma \to X^\Sigma$, for which $\xi_1 = q \circ p$. Thus, the map p is, as stated, a homotopy equivalence.

For the existence of a partition of unity $\{f_\alpha; \alpha \in A\}$ subordinate to a locally finite covering Γ, it is sufficient, as we know, that the space X be normal. On the other hand, for a covering Γ to exist, it is sufficient, for example, to require that the space X be paracompact or (assuming that X is normal) that for any $n \geq 0$ the inclusion $\overline{\operatorname{int} X_n} \subset \operatorname{int} X_{n+1}$ holds. (Indeed, in the latter case, the locally finite covering inscribed in the covering $\{\operatorname{int} X_n; x \geq 0\}$ is the covering $\{\operatorname{int} X_1, \operatorname{int} X_{n+1} \setminus \overline{V}_{n-1}; n \geq 0\}$, where V_{n-1} is a neighbourhood of the set $\overline{\operatorname{int} X_{n-1}}$ such that $\overline{V}_{n-1} \subset \operatorname{int} X_n$). Thus,

Proposition 2.67. *if condition 1) of the previous proposition is satisfied, and the space X is normal and, in addition, either this space is paracompact or*

$$\overline{\operatorname{int} X_n} \subset \operatorname{int} X_{n+1}$$

for any $n \geq 0$, then the space X is a homotopy limit of the subspaces X_n.

These conditions are by no means necessary.

Let us now consider (for an arbitrary space X and arbitrary subspaces X_n), the subspaces

$$X_0^\Sigma \subset X_1^\Sigma \subset \cdots \subset X_n^\Sigma \subset \cdots$$

of the space X^Σ. It is clear that $\operatorname{int} X_n^\Sigma$ is the set of all points $(x,t) \in X^\Sigma$ for which $t < n$. Therefore,

$$X^\Sigma = \sum_{n=0}^{\infty} \operatorname{int} X_n^\Sigma,$$

i.e., for the subspaces X_n^Σ, condition 1) of the proposition proved above is satisfied. Moreover, it is easy to see that the sets

$$\operatorname{int} X_1^\Sigma, \quad \operatorname{int} X_{n+1}^\Sigma \setminus \operatorname{int} X_{n-1}^\Sigma \qquad n \geq 1, \tag{2.68}$$

form a locally finite covering refining the covering $\{\operatorname{int} X_n^\Sigma; n \geq 0\}$, and the functions f_{n+1}, $n \geq 0$, defined by the formula

$$f_{n+1}(x,t) = \begin{cases} 0, & \text{if} \quad t \leq n - \frac{3}{4} \text{ or } t \geq n + \frac{3}{4}, \\ \frac{4t-4n+3}{2}, & \text{if} \quad n - \frac{3}{4} \leq t \leq n - \frac{1}{4}, \\ 1, & \text{if} \quad n - \frac{1}{4} \leq t \leq n + \frac{1}{4}, \\ \frac{3+4n-4t}{2}, & \text{if} \quad n - +\frac{1}{4} \leq t \leq n + \frac{3}{4}, \end{cases}$$

constitute a decomposition of the unit subordinate to covering (2.68). Hence,

Proposition 2.69. *for any space X and any of its subspaces*

$$X_0 \subset X_1 \subset \cdots \subset X_n \subset \cdots$$

the space X^Σ is the homotopy limit of the subspaces

$$X_n^\Sigma = (X_0 \times [0,1]) \cup \cdots \cup (X_{n-1} \times [n-1,n]) \cup (X_n \times n).$$

This proposition is a strengthening of the property of subspaces X_n^Σ proved above. Using it, we can extend the basic properties of weak homotopy limits to the case of homotopy limits.

Let, for example, X be an arbitrary space that is a homotopy limit of the subspaces

$$X_0 \subset X_1 \subset \cdots \subset X_n \subset \cdots$$

Y be an arbitrary space in which an increasing sequence of subspaces

$$Y_0 \subset Y_1 \subset \cdots \subset Y_n \subset \cdots$$

is defined, and f be a continuous map

$$f : X \to Y$$

such that

$$f(X_n) \subset Y_n$$

for any $n \geq 0$. Then

Proposition 2.70. *if for each $n \geq 0$ the map f defined by*

$$f_n : X_n \to Y_n$$

is a homotopy equivalence, then the following two conditions are equivalent:

1) the space Y is a homotopy limit of the subspaces Y_n;

2) the map f is a homotopy equivalence.

Proof. Indeed, let, as above,

$$X^{\Sigma} = \cup_{k=0}^{\infty}(X_k \times [k, k+1])$$

and let similarly

$$Y^{\Sigma} = \cup_{k=0}^{\infty}(Y_k \times [k, k+1]).$$

Since $f(X_k) \subset Y_k$, the formula

$$f^{\Sigma}(x, t) = (f(x), t), \quad (x, t) \in X^{\Sigma},$$

defines a continuous map

$$f^{\Sigma} : X^{\Sigma} \to Y^{\Sigma}$$

The map f is obviously related to the map f^{Σ} by the formula

$$f \circ p^{\Sigma} = \overline{p}^{\Sigma} \circ f^{\Sigma},$$

where

$$p^{\Sigma} : X^{\Sigma} \to X, \quad \overline{p}^{\Sigma} : Y^{\Sigma} \to Y$$

are natural maps. Therefore, if statement 1) is true, then the map f^{Σ} is homotopically equivalent to the map f, and if statement 2) is true, then the composition of the maps f^{Σ} and $\overline{p}^{\Sigma}$ is a homotopy equivalence. Therefore, to prove the equivalence of statements 1) and 2), it suffices to prove that

Proposition 2.71. *the map f^{Σ} is a homotopy equivalence.*

Let us first consider the case when $X = Y$ and $X_n = Y_n$ for all $n \geq 0$. Let, in addition, each map

$$f_n : X_n \to X_n, \quad n \geq 0,$$

be homotopically identical and let

$$\xi_{n,t} : X_n \to X_n$$

be a deformation of the space X_n that connects the identity map 1_{X_n} with the map f_n. For any point $x \in X_n$, $n \geq 0$, and any $t, \tau \in I$, we set

$$h_\tau(x, n+t) = \begin{cases} (f(x), n + t(2\tau + 1)), & \text{if } 0 \leq \tau \leq 1/2,\ 0 \leq t \leq 1/2, \\ (f(x), n + 2(1-t)\tau + t), & \text{if } 0 \leq \tau \leq 1/2,\ 1/2 \leq t \leq 1, \\ (\xi_{n,2-2\tau}(x), n + 2t), & \text{if } 1/2 \leq \tau \leq 1,\ 0 \leq t \leq 1/2, \\ (\xi_{n,1-(3-4t)(2\tau-1)}(x), n+1), & \text{if } 1/2 \leq \tau \leq 1,\ 1/2 \leq t \leq 3/4, \\ (\xi_{n+1,1-(4t-3)(2\tau-1)}(x), n+1), & \text{if } 1/2 \leq \tau \leq 1,\ 3/4 \leq t \leq 1. \end{cases}$$

It is easy to see (Fig. 2.4) that we thereby obtain a certain homotopy

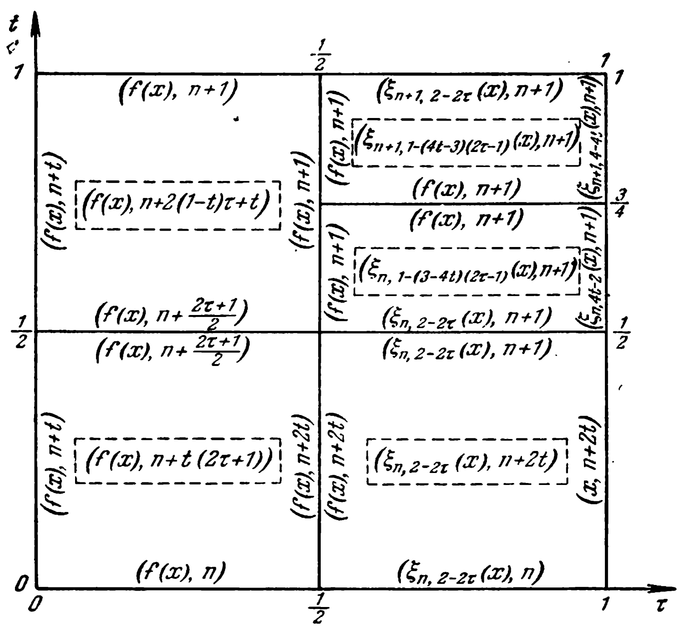


Figure 2.4:

$$h_\tau : X^\Sigma \to X^\Sigma$$

connecting the map f^Σ with the map h defined by the formula

$$h(x, n+t) = \begin{cases} (x, n + 2t), & \text{if } 0 \leq t \leq 1/2, \\ (\xi_{n,4t-2}(x), n+1), & \text{if } 1/2 \leq t \leq 3/4, \\ (\xi_{n+1,4-4t}(x), n+1), & \text{if } 3/4 \leq t \leq 1. \end{cases}$$

Therefore, it is sufficient for us to prove that the map h is a homotopy equivalence. With this in mind, we note that by the equality

$$h\left(x, n + \frac{1}{2}\right) = h(x, n+1) = h(x, n+1)$$

the formula

$$g(x,n+t)=\begin{cases}(x,n+2t), & \text{if} \quad 0\le t\le 1/2,\\ h\left(x,n+\frac{3-2t}{2}\right), & \text{if} \quad 1/2\le t\le 1.\end{cases}$$

uniquely determines some continuous map

$$g : X^{\Sigma}\longrightarrow X^{\Sigma}.$$

Let

$$\eta_\tau(x,n+t)=\begin{cases}(x,n+(1+6\tau)t), & \text{if} \quad 0\le\tau\le 1/2,\ 0\le t\le 1/4,\\ (x,n+2(1-t)\tau+t), & \text{if} \quad 0\le\tau\le 1/2,\ 1/4\le t\le 1,\\ (h\circ g)(x,n+t), & \text{if} \quad 1/2\le\tau\le 1,\ 0\le t\le \tau/2,\\ h(x,n+\tau), & \text{if} \quad 1/2\le\tau\le 1,\ \tau/2\le t\le \frac{3-2\tau}{2},\\ (h\circ g)(x,n+t), & \text{if} \quad 1/2\le\tau\le 1,\ \frac{3-2\tau}{2}\le t\le 1.\end{cases}$$

Since

$$(h\circ g)\left(x,n+\frac{\tau}{2}\right)=h(x,n+\tau),\quad 0\le\tau\le 1,$$

we have

$$\begin{aligned}(h\circ g)\left(x,n+\frac{\tau}{2}\right)&=(h\circ h)\left(x,n+\frac{3-2\frac{3-2\tau}{2}}{2}\right)\\ &=(h\circ h)(n+\tau)=\begin{cases}h(\xi_{n,\tau-2}(x),n+1) & =(\xi_{n,\tau-2}(x),n+1)\\ h(\xi_{n+1,4-4\tau}(x),n+1) & =(\xi_{n+1,4-4\tau}(x),n+1)\end{cases}\\ &=h(x,n+\tau),\quad 1/2\le\tau\le 1,\end{aligned}$$

and

$$(h\circ g)(x,n+t)=(x,n+4t),\quad 0\le t\le 1/4,$$

this formula defines (Fig. 2.5) a certain homotopy,

$$\eta_\tau : X^{\Sigma}\to X^{\Sigma},\quad 0\le\tau\le 1,$$

connecting the identity map $1_{X^{\Sigma}}$ with the map

$$h\circ g : X^{\Sigma}\longrightarrow X^{\Sigma}.$$

Similarly, in view of the equalities

$$(g\circ h)(x,n+\frac{\tau}{2})=g(x,n+\tau),\quad 0\le\tau\le 1,$$

we have

$$\begin{aligned}(g\circ h)\left(x,n+\frac{3-2\tau}{2}\right)&=h\left(x,n+\frac{3-2\tau}{2}\right)\\ &=g(n+\tau),\quad 1/2\le\tau\le 1,\end{aligned}$$

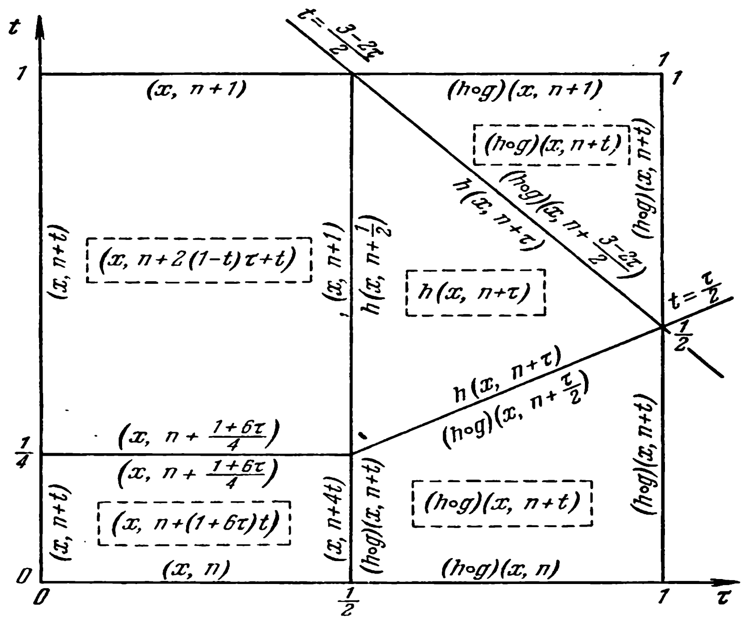


Figure 2.5:

and

$$(g \circ h)(x, n + t) = (x, n + 4t), \quad 0 \leq t \leq 1/4,$$

the formula

$$\overline{\eta}_\tau(x, n+t) = \begin{cases} (x, n + (1 + 6\tau)t), & \text{if } \ 0 \leq \tau \leq 1/2, \ 0 \leq t \leq 1/4, \\ (x, n + (1 - t)\tau + t), & \text{if } \ 0 \leq \tau \leq 1/2, \ 1/4 \leq t \leq 1, \\ (g \circ h)(x, n + t), & \text{if } \ 1/2 \leq \tau \leq 1, \ 0 \leq t \leq \tau/2, \\ g(x, n + \tau), & \text{if } \ 1/2 \leq \tau \leq 1, \ \tau/2 \leq t \leq \frac{3-2\tau}{2}, \\ (g \circ h)(x, n + t), & \text{if } \ 1/2 \leq \tau \leq 1, \ \frac{3-2\tau}{2} \leq t \leq 1. \end{cases}$$

defines some homotopy

$$\overline{\eta}_\tau : X^\Sigma \to X^\Sigma, \quad 0 \leq \tau \leq 1,$$

connecting the identity map 1_{X^Σ} with the map

$$g \circ h : X^\Sigma \to X^\Sigma.$$

Thus, in the particular case under consideration, the map f^Σ is indeed a homotopy equivalence.

Let us now consider the general case.

Let

$$g_n : Y_n \to X_n, \quad n \geq 0,$$

be homotopy equivalences inverse to the homotopy equivalences of f_n, and let

$$i_n : X_n \subset X_{n+1}, \quad j_n : Y_n \subset Y_{n+1}$$

be inclusion maps. Since $f_{n+1} \circ i_n = j_n \circ f_n$, then

$$i_n \circ g_n \sim g_{n+1} \circ f_{n+1} \circ i_n \circ g_n = g_{n+1} \circ j_n \circ f_n \circ g_n \sim g_{n+1} \circ j_n.$$

Let

$$h_{n,\tau} : Y_n \to X_{n+1}, \quad 0 \le \tau \le 1,$$

be an arbitrary homotopy connecting the map $i_n \circ g_n$ with the map $g_{n+1} \circ j_n$. It is easy to see that the formula

$$h(y, n+t) = \begin{cases} (g_n(y), n+2t), & \text{if} \quad 0 \le t \le 1/2, \\ (h_{n,2t-1}(y), n+1), & \text{if} \quad 1/2 \le t \le 1. \end{cases}$$

where $y \in Y_n$, $n \ge t$ and $0 \le t \le 1$, uniquely determines some continuous map

$$h : Y^\Sigma \to X^\Sigma.$$

Consider the map

$$h \circ f^\Sigma : X^\Sigma \to X^\Sigma.$$

It is clear that

$$(h \circ f^\Sigma)(X_n^\Sigma) \subset X_n^\Sigma, \quad n \ge 0,$$

and therefore for any $n \ge 0$ the map $f \circ f^\Sigma$ defines a certain map

$$(h \circ f^\Sigma)_n : X_n^\Sigma \to X_n^\Sigma.$$

Moreover, for any point $x \in X$ the equality

$$(h \circ f^\Sigma)_n(x, n) = ((g_n \circ f_n)(x), n).$$

holds. In other words,

$$(h \circ f^\Sigma)_n \circ q_n^\Sigma = q_n^\Sigma \circ (g_n \circ f_n).$$

where q_n^Σ is the homotopy equivalence $q_n^\Sigma : X_n \to X_n^\Sigma$ constructed above, which is inverse to the homotopy equivalence $p_n^\Sigma : X_n^\Sigma \to X_n$. Therefore,

$$(h \circ f^\Sigma)_n \sim (h \circ f^\Sigma)_n \circ (q_n^\Sigma \circ p_n^\Sigma) = q_n^\Sigma \circ ((g_n \circ f_n) \circ p_n^\Sigma \sim q_n^\Sigma \sim q_n^\Sigma \circ p_n^\Sigma \sim 1_{X_n^\Sigma}.$$

Thus, the map $h \circ f^\Sigma$ has (with respect to the subspaces X_n^Σ) the property that we required of the map f in the first part of the proof. Therefore, by what has already been proved, the map $h \circ f^\Sigma$ is a homotopy equivalence. Consequently, the map f^Σ is homotopy injective.

Similarly, the map

$$f^\Sigma \circ h : Y^\Sigma \to Y^\Sigma$$

has the property that

$$(f^{\Sigma} \circ h)(Y_n^{\Sigma}) \subset Y_n^{\Sigma}, \quad n \geq 0,$$

and therefore defines maps

$$(f^{\Sigma} \circ h)_n : (Y_n^{\Sigma}) \longrightarrow Y_n^{\Sigma}, \quad n \geq 0,$$

satisfying the relation

$$(f^{\Sigma} \circ h)_n \circ \overline{q}_n^{\Sigma} = \overline{q}_n^{\Sigma} \circ (f_n \circ g_n).$$

where $\overline{q}_n^{\Sigma}$ is the map q_n^{Σ} constructed for the spaces Y_n. From the last relation it follows that the maps $(f^{\Sigma} \circ h)_n$ are homotopy identities. Therefore, the map $f^{\Sigma} \circ h$ is also a homotopy equivalence, and therefore the map f^{Σ} is homotopy sujective. □

Being both homotopically injective and homotopically surjective, the map f^{Σ} is a homotopy equivalence. □

Chapter 3

Cellular decompositions

General cellular decompositions seem to provide us with the most natural stock of objects for constructing a homotopy theory. They have all the basic geometric properties of classical simplicial decompositions (triangulations), and at the same time their theory compares favourably with the theory of simplicial decompositions in its generality, structure, and internal integrity. However, until now no one has apparently attempted to give a coherent and independent exposition of the basic properties of cellular decompositions. This chapter is the first attempt in this direction.

The definition and simplest properties of cellular decompositions are presented in §3.1 and §3.2.

In §3.3 it is proved that any cellularly decomposed space is paracompact (and, therefore, normal). Here it is also proved that the topological product of two cellular decompositions is a cellular decomposition if at least one of the given decompositions is locally finite or if both of them are locally countable.

In §3.4, after a number of simple remarks on continuous maps of cellularly decomposed spaces, it is proved that any cellular pair satisfies the axiom of homotopy extension, and the main theorem on cellular maps is formulated (along with some corollaries). §3.5 is devoted to the proof of this last theorem.

In the final §3.6, a remarkable theorem of Whitehead is proved that any weak homotopy equivalence connecting cellular spaces is a homotopy equivalence. In this section, several simple remarks are placed on quasi-polyhedra (i.e., spaces homotopy equivalent to cellularly decomposed spaces).

3.1 Cellular pre-decompositions

Let X be an arbitrary Hausdorff space. We call a subset e of X an *(open) cell* if there exists a continuous map

$$\chi : \mathbb{E}^n \to X \tag{3.1}$$

of the unit Euclidean ball $\mathbb{E}^n$ into X that homeomorphically maps the open ball $\dot{\mathbb{E}}^n$ onto the set e and has the property that $\chi(\mathbb{S}^{n-1}) \cap e = \varnothing$ where $\mathbb{S}^{n-1} = \mathbb{E}^n \setminus \dot{\mathbb{E}}^n$. We will call the dimension n of the ball $\mathbb{E}^n$ the *dimension* $\dim e$ of the cell e. Clearly, it is uniquely determined by the cell e (since the balls $\mathbb{E}^n$ are not homeomorphic for different n). When it is necessary to specifically indicate the dimension n of the cell e, we will denote this cell by the symbol e^n.

The image $\chi(\mathbb{E}^n)$ of the ball $\mathbb{E}^n$ under map (3.1) is compact and therefore a closed subset of the Hausdorff space X. Consequently, the closure $\overline{e} = \overline{\chi(\mathbb{E}^n)}$ of the cell e is contained in the set $\chi(\mathbb{E}^n)$ and therefore coincides with it (since the continuity of the map χ implies the inclusion $\chi(\mathbb{E}^n) \subset \overline{\chi(\dot{E}^n)}$ (see the properties of continuous maps in §1.5). Thus,

Proposition 3.2. *the image of the ball $\mathbb{E}^n$ under the map (3.1) is the closure $\overline{e}$ of the cell e.*

The cells of the space X need not be open in X. For example, for $n = 0$, the ball $\mathbb{E}^0 = \dot{E}^n$ is a one-point space, so that any point of the space X is its zero-dimensional cell.

For $n > 0$, the points of the ball $\mathbb{E}^n$ have the form $v\boldsymbol{u}$, where $0 \le v \le 1$, and $\boldsymbol{u}$ is the unit vector, i.e., a point of the unit sphere $\mathbb{S}^{n-1}$. Thus, any point of the set $\overline{e}$ (after the map (3.1) is chosen) has the form $\chi(v\boldsymbol{u})$, where $0 \le v \le 1$, and $|\boldsymbol{u}| = 1$. Moreover, the number v is uniquely determined by a given point. The vector $\boldsymbol{u}$ is also uniquely determined if only $0 < v < 1$.

We will call map (3.1) *characteristic* for the cell e. Clearly, every map of the form $\chi \circ \alpha$ is also characteristic, where $\alpha : \mathbb{E}^n \to \mathbb{E}^n$ is an arbitrary homeomorphism of the ball $\mathbb{E}^n$ onto itself. However, there may exist characteristic maps $\mathbb{E}^n \to X$ for the cell e that are different from maps of the form $\chi \circ \alpha$.

Thus, for example, for a disk $\mathbb{E}^2$ in a plane, along with all possible homeomorphisms $\mathbb{E}^2 \to \mathbb{E}^2$, its characteristic maps are identification maps, under which some arc of the boundary circle $\mathbb{S}^1$ is contracted to a point.

Sometimes it will be convenient for us to imagine the set $\overline{e}$ as a continuous image not of the unit ball $\mathbb{E}^n$, but of some space $\mathbb{E}^n_e$ homeomorphic to the ball $\mathbb{E}^n$ (for example, the unit cube or the product $\mathbb{E}^p \times \mathbb{E}^q$, where $p+q = n$). In this case, the maps $\mathbb{E}^n_e \to X$, which are the composition of some homeomorphism $\mathbb{E}^n_e \to \mathbb{E}^n$ and an arbitrary characteristic map $\mathbb{E}^n \to X$, we will also call *characteristic maps*.

Since the space X is, by assumption, Hausdorff and the ball $\mathbb{E}^n$ is compact, the characteristic map χ (considered as a map of the ball $\mathbb{E}^n$ onto the closure $\overline{e}$ of the cell e) is an identification map and therefore

Proposition 3.3. *the topology of the set $\overline{e}$ is the identification topology.*

In particular, we see that the topology of the set $\overline{e}$ does not depend on the topology of the space X in the sense that any other topology in X in which the set e is a cell (with characteristic map χ) induces the same identification topology on $\overline{e}$.

Note that, being a continuous image of a compact and path-connected set $\mathbb{E}^n$,

Proposition 3.4. *the set $\overline{e}$ is compact and path-connected.*

The set

$$\dot{e} = \overline{e} \setminus e$$

will be called the *boundary* of the cell e. For any choice of the characteristic map $\chi : \mathbb{E}^n \to X$, it coincides with the set $\chi(\mathbb{S}^{n-1})$. Therefore, for $n > 1$, the set $\dot{e}$ is path-connected, and for $n = 1$, it consists of at most two points. For $n = 0$, the set $\dot{e}$ is empty.

The main object of our study will be certain families K of pairwise disjoint cells $e \subset X$. For each such family, we will denote by $|K|$ its *body*, i.e., the subspace of X that is the union of all cells $e \in K$, and by K^n, where n is some non-negative integer, its n-th skeleton, i.e., the subfamily of the family K consisting of all cells $e \in K$ whose dimension $\dim e$ does not exceed n. It is convenient to add to the number of skeletons the empty subfamily of the family K, considering it the (-1)-th skeleton of K^{-1}, as well as the family K itself, considering it the ∞-th skeleton of K^{∞}.

By choosing characteristic maps $\chi_e : \mathbb{E}_e \to X$ for all cells e of the family K such that $\mathbb{E}_{e_1} \cap \mathbb{E}_{e_2} = \varnothing$ for $e_1 \neq e_2$, we can define a topological sum

$$P_K = \cup_{e \in K} \mathbb{E}_e$$

of the spaces $\mathbb{E}_e$ and a map

$$\chi : P_k \to X,$$

coinciding on each of the summands of $\mathbb{E}_e$ with the corresponding map χ_e. We will call the (obviously continuous) map $\chi : P_k \to X$ constructed in this way *characteristic* for the family K.

A family K of pairwise disjoint cells $e \subset K$ will be called a *cellular pre-decomposition* of the space X if $|K| = X$ and for any $n > 0$ the boundary $\dot{e}^n$ of each n-dimensional cell $e^n \in K$ belongs to the subspace $|K^n|$ (the body of the $n-1$-th skeleton).

Remark 3.5. Usually, the condition $\chi(\mathbb{S}^{n-1} \cap e) = \varnothing$ is not included in the definition of a cell, since only cells that make up pre-decompositions are of interest, and for such cells this condition is satisfied automatically.

It is clear that

Proposition 3.6. *any cellular sub-decomposition K contains at least one zero-dimensional cell, i.e., its 0-th skeleton K^0 is not empty (unless, of course, the space X is empty).*

In what follows, we will call the zero-dimensional cells of a cellular pre-decomposition K its *vertices.*

Since

$$K^m \subset K^n$$

for $m \leq n$ and, in particular, $K^0 \subset K^n$ for $n \geq 0$, then

Proposition 3.7. *any skeleton K^n of an arbitrary cellular pre-decomposition K is not empty.*

A pre-decomposition K is called *finite-dimensional* if there exists $n \geq 0$ such that $K = K^n$. The smallest of these n is called the *dimension* $\dim K$ of the pre-decomposition K. Clearly, $K^m = K$ for any $m \geq \dim K$. Moreover, $\dim K^m \leq m$ for any $m \geq 0$ (the strict inequality $\dim K^m < m$ holds if the decomposition K has no cells of dimension m).

Remark 3.8. Thus, for unequal n and m, the equality $K^n = K^m$ is possible (even for n and m less than $\dim K$). This is possible when the pre-decomposition has no cells of intermediate dimensions.

Example 3.9. (of pre-decompositions.)

1) Every Hausdorff space X can be "scattered" into a family K consisting of all its points. This family will be a cellular pre-decomposition of X containing only zero-dimensional cells.

2) For any $n \geq 0$, the family $K = \{e^0, e^n\}$ consisting of the point $e^0 = \boldsymbol{u}_0 = (1, 0, \ldots, 0)$ of the sphere $\mathbb{S}^n$ and its complement $e^n = \mathbb{S}^n \setminus \boldsymbol{u}_0$ is a cellular pre-decomposition of the sphere $\mathbb{S}^n$. Here $\dot{e}^n = e^0$ (if $n > 0$), $k^m = \{e^0\}$ for $0 \leq m < n$, and $K^m = K$ for $m \geq n$.

3) For any $n \geq 0$, the family $k = \{e^0, e^n, e^{n+1}\}$ consisting of the cell $e^{n+1} = \dot{\mathbb{E}}^n$ and the pre-decomposition $\{e^0, e^n\}$ of the sphere $\mathbb{S}^n$ constructed in Example 2) is a cellular pre-decomposition of the ball $\mathbb{E}^{n+1}$. Similarly, the family consisting of two points 0 and 1 and the open interval $(0, 1)$ is a cellular pre-decomposition of the segment $I = [0, 1]$. Taking liberties, we will denote this pre-decomposition by the same symbol I.

4) For any $n \geq 0$, the family K consisting of the cell $e^{n+1} = \dot{\mathbb{E}}^n$ and all points of the sphere $\mathbb{S}^n$ is also a cellular pre-decomposition of the ball $\mathbb{E}^{n+1}$. Its 0-th skeleton K^0 contains (for $n > 0$) an uncountable number of cells and is a pre-decomposition of 1), constructed for the space $X = \mathbb{S}^n$.

Remark 3.10. A reader familiar with the concept of a (finitely) triangulated space will immediately recognise that any triangulation of such a space is a cellular pre-decomposition of it. In the following discussion, cellular pre-decomposition of this kind are not used.

We will call a subfamily L of a cellular pre-decomposition K its *sub-pre-decomposition* if $\overline{e} \subset |L|$ for any cell $e \in L$. It is clear that

1) any sub-pre-decomposition L is a cellular pre-decomposition of the subspace $|L|$;

2) any subfamily L of a cellular pre-decomposition K, for which the subspace $|L|$ is closed, is a sub-pre-decomposition;

3) for any $n \geq 0$, the skeleton K^n is a sub-pre-decomposition of the pre-decomposition K;

4) the union and intersection of any (finite or infinite) system $\{L_\alpha\}$ of sub-pre-decomposition of the pre-decomposition K is also a sub-pre-decomposition, and
$$|\cap L_\alpha| = \cap|L_\alpha|, \quad |\cup L_\alpha| = \cup|L_\alpha|.$$

In connection with assertion 2), we note that there may well exist sub-pre-decompositions of L for which the subspace $|L|$ is not closed. (For example, any subset of the space X is a sub-pre-decomposition of the pre-decomposition of this space specified in Example 3.9 - 1)). However,

Proposition 3.11. *for any finite (i.e., consisting of a finite number of cells) sub-pre-decomposition of L, the subspace $|L|$ is closed and even compact.*

Proof. Indeed, in this case the subspace $|L|$ is the union of a finite number of compact sets $\overline{e}$, $e \in L$. □

From property 4) it follows, in particular, that

Proposition 3.12. *for any subset $A \subset X$ there exists a smallest sub-pre-decomposition $L \subset K$ for which $|L| \supset A$.*

Such a sub-pre-decomposition is the intersection of all sub-pre-decompositions of $L \subset K$ for which $|L| \supset A$. We will denote it by the symbol $K(A)$.

It is clear that

1) for any point $x \in X$, the equality
$$K(x) = K(e) = K(\overline{e})$$
holds, where e is a cell of the pre-decomposition K containing the point x;

2) for any cell $e \in K$, the sub-pre-decomposition $K(e)$ consists of the sub-pre-decomposition $K(\dot{e})$ and the cell e:
$$K(e) = \{K(\dot{e}), e\};$$

3) for any cell e belonging to a sub-pre-decomposition L of a pre-decomposition K, the equality
$$L(e) = K(e)$$
holds.

We will call a cellular pre-decomposition K of a space X *point finite* if any point $x \in X$ belongs to the body $|L|$ of some finite sub-pre-decomposition $L \subset K$ or, in other words, if for any point $x \in X$ the sub-pre-decomposition $K(x)$ is finite. Similarly, if any point $x \in X$ is an interior point of the body $|L|$ of some finite (resp. countable) sub-pre-decomposition $L \subset K$, then we will call the sub-pre-decomposition K *locally finite* (respectively, *locally countable*). It is clear that

Proposition 3.13. *any locally finite pre-decomposition of K is point finite.*

Moreover,

Proposition 3.14. *any sub-pre-decomposition L of a point finite (resp. locally finite or locally countable) pre-decomposition K is point finite (respectively, locally finite and locally countable).*

Further, it is easy to see that

Proposition 3.15. *for any locally finite (resp. locally countable) pre-decomposition K of X, the characteristic map*

$$\chi : P_K \to X$$

is locally compact (respectively, locally countably compact).

Proof. Indeed, let p be an arbitrary point in the space P_K. By the condition, the point $\chi(p) \in X$ has in the space X a neighbourhood U contained in the body $|L|$ of some finite (resp. countable) sub-pre-decomposition $L \subset K$. Let $S = \chi^{-1}(U)$ and $C = \chi^{-1}(|L|)$. It is clear that the set S is a saturated (with respect to the map χ) neighbourhood of the point p in the space P_K, and the set C is the union of a finite (respectively, countable) number of terms $\mathbb{E}_e$ of the topological sum P_K and is therefore closed and compact (respectively, countably compact). Moreover, the set C is saturated with respect to the map χ and contains the closure $\overline{S}$ of the neighbourhood S. Thus, each point in the space P_K has a saturated neighbourhood, the closure of which is contained in a saturated, closed, and compact (resp. countably compact) set. But this, by definition, means that the map χ is locally compact (resp. locally countably compact). □

Remark 3.16. It is easy to see that if a pre-decomposition K is point finite, then the converse is also true, i.e., any point finite pre-decomposition K for which the characteristic map $P_K \to X$ is locally compact (reps. locally countably compact) is a locally finite (resp. locally countable) pre-decomposition. We will not need this fact.

For any closed set $A \subset X$, all sets of the form $A \cap \overline{e}$, $e \in K$, are, of course, closed (in X). If the converse is true, i.e. if every set $A \subset X$ for which the family $\{A \cap \overline{e}; e \in K\}$ consists of closed sets is itself closed (in X), then we will say that the topology of the space X is a *weak topology* with respect to the pre-decomposition K. In other words, the topology of a space X is a weak topology with respect to the pre-decomposition K if this space is a free union of subspaces $\overline{e} \subset X$, $e \in K$. Clearly, this is the case if and only if the characteristic map

$$\chi : P_K \to X$$

is an identification map.

Similarly, for any closed set $A \subset X$ and any finite sub-pre-decomposition $L \subset K$ the intersection $A \cap |L|$ is closed in X. It is easy to see that

Proposition 3.17. *if the converse is true, i.e., if a set $A \subset X$ is closed in X when it has the property that for any finite sub-pre-decomposition L of a pre-decomposition K the intersection $A \cap |L|$ is closed in X, then the topology of the space X is a weak topology with respect to the pre-decomposition K.*

Proof. Indeed, if for any cell $e \in K$ the intersection $A \cap \overline{e}$ is closed, then for any finite sub-pre-decomposition $L \subset K$ the intersection $A \cap |L|$ is also closed (since the space $|L|$ is the union of a finite number of sets of the form $\overline{e}, e \in L$), and therefore, according to the condition, the set A is closed. □

However, in the space X, whose topology is a weak topology with respect to the pre-decomposition K, there may exist non-closed subsets A possessing the property that for any finite sub-pre-decomposition $L \subset K$, the intersection $A \cap |L|$ is closed (this property is possessed, for example, by any subset A of the sphere $\mathbb{S}^{n-1}$ in Example 3.9 - 4)).

It follows directly from the proven statement, in particular, that

Proposition 3.18. *for any finite pre-decomposition K of the space X (when such a pre-decomposition exists), the topology of the space X is a weak topology with respect to the pre-decomposition K.*

Moreover, it is easy to see that

Proposition 3.19. *the last statement holds not only for finite but also for any locally finite pre-decompositions.*

Proof. Indeed, for any locally finite pre-decomposition K of the space X, the family $\{\overline{e} : e \in K\}$ is obviously a locally finite family of closed subspaces of the space X, and therefore this space is their free union (see §1.1). □

Let K and L be arbitrary cellular pre-decompositions of Hausdorff spaces X and Y, respectively. It is clear that in the case when the spaces X and Y do not intersect, the union $K \cup L$ of the cellular pre-decompositions K and L is a cellular pre-decomposition of the topological sum $X \cup Y$. We will call the pre-decomposition $K \cup L$ the *sum* of the pre-decompositions K and L. According to what has just been said,

$$|K \cup L| = |K| \cup |L|.$$

Similar statements hold for any number of summands.

Furthermore, since for any $n \geq 0$ and $m \geq 0$, the ball $\mathbb{E}^{n+m}$ is homeomorphic to the product $\mathbb{E}^n \times \mathbb{E}^n$, then for any two cells $e^n \in K$ and $e^m \in L$ (with characteristic maps $\chi_1 : \mathbb{E}^n \to X$ and $\chi_2 : \mathbb{E}^n \to Y$), the product $e^n \times e^m$ represents an $(n+m)$-dimensional cell of the space $X \times Y$ (for the corresponding characteristic map, one can take, for example, the map $\chi_1 \times \chi_2 : \mathbb{E}^n \times \mathbb{E}^m \to X \times Y$), where

$$(e^n \times e^m)^{\cdot} = \dot{e}^n \times e^m \cup e^n \times \dot{e}^m.$$

Consequently, the family $K \times L$ of all cells of the form $e^1 \times e^2$, where $e^1 \in K$ and $e^2 \in L$, constitutes a cellular pre-decomposition of the space $X \times Y$, denoted as $|K \times L| = |K| \times |L|$. We will call the pre-decomposition $K \times L$ the *product* of the pre-decompositions K and L. The product of any finite number of pre-decompositions is defined similarly.

It is clear that

Proposition 3.20. *the product of point finite (respectively, locally finite and locally countable) pre-decompositions is point finite (respectively, locally finite and locally countable).*

In the special case where the pre-decomposition L is the pre-decomposition $I = \{0, 1, (0, 1)\}$ of the interval $I = [0, 1]$, we obtain a pre-decomposition $\times I$ of the product $X \times I$. The cells of this pre-decomposition are of the form $e \times 0$, $e \times 1$ and $e \times (0, 1)$, where $e \in K$. In this case, the dimension of the cells $e \times 0$ and $e \times 1$ is equal to the dimension n of the cell e, and the dimension of the cell $e \times (0, 1)$ is equal to $n + 1$.

Let

$$\chi_K : P_K \to X, \quad \chi_L : P_L \to Y$$

be the characteristic maps for the pre-decompositions K and L respectively. It is clear that in the case when the spaces P_K and P_L are chosen to be disjoint, we can take their topological sum $P_K \cup P_L$ as the space $P_{K \cup L}$, and the corresponding characteristic map $P_{K \cup L} \to X \cup Y$ can then be considered as the map

$$\chi_K \cup \chi_L : P_K \cup P_Y \to X \cup Y.$$

Similarly, we can consider the space $P_K \times P_L$ as the space $P_{X \times Y}$ and the map

$$\chi_K \times \chi_L : P_K \times P_Y \to X \times Y.$$

as the characteristic map of the pre-decomposition $K \times L$.

3.2 Cellular decompositions

We will call a cellular pre-decomposition K of a space X a *cellular decomposition* if

- it is point finite, and
- the topology of the space X is a weak topology with respect to the pre-decomposition K.

It follows directly from the results of the previous section that

Proposition 3.21. *any locally finite (in particular, any finite) pre-decomposition of K is a decomposition.*

Thus, the pre-decompositions indicated in examples 2) and 3) of Example 3.9 in §3.1 are decompositions. Conversely, the pre-decomposition in example 1) is a decomposition only when the space X is discrete (because otherwise its topology will not be a weak topology). In particular, when $X = \mathbb{S}^n$, $n > 0$, this pre-decomposition is not a decomposition. Similarly, the pre-decomposition in example 4) of Example 3.9 in §3.1 is not a decomposition either, because it is not point finite (note that the condition of weakness of the topology is satisfied in this example).

The last pre-decomposition also possesses sub-pre-decompositions that are not decompositions (these are arbitrary subsets of the sphere $\mathbb{S}^{n-1}$). At the same time,

Proposition 3.22. *any sub-pre-decomposition of an arbitrary cell decomposition K is itself a decomposition.*

Before proving the last statement, we note that the sufficient condition for the weakness of the topology of the space X indicated in §3.1 is also necessary in the case of decompositions, so that

Proposition 3.23. *A subset A of a space $X = |K|$ is closed if for any finite sub-decomposition L of the decomposition K the intersection $A \cap |L|$ is closed.*

Proof. Indeed, since for any cell $e \in K$ the intersection $A \cap |K(e)|$ is closed (since the sub-decomposition $K(e)$ is finite) and since $\overline{e} \subset |K(e)|$, the intersection $A \cap \overline{e} = (A \cap |K(e)|) \cap \overline{e}$ is closed. Therefore, the subset A is also closed. □

It follows from this that, in contrast to the case of arbitrary sub-decompositions,

Proposition 3.24. *the body $|L|$ of each sub-pre-decomposition L of the decomposition K is closed in the space $X = |K|$.*

Proof. Indeed, for any finite sub-decomposition of $N \subset K$ the intersection $L \cap N$ is also finite and therefore the intersection $|L| \cap |N| = |L \cap N|$ is closed. □

Now let A be an arbitrary subset of $|L|$ such that, for any finite sub-decomposition $N \subset L$, the intersection $A \cap |N|$ is closed (in $|L|$, and therefore, by what has been proved, in X). Then, for any finite sub-decomposition $N \subset K$, the intersection

$$A \cap |N| = A \cap |L \cap N|$$

is also closed, and, consequently, the set A is closed (in X, and therefore in $|L|$). This proves that the topology of the space $|L|$ is a weak topology with respect to the pre-decomposition L. Since the pre-decomposition L is, moreover, point finite (being a sub-pre-decomposition of the point finite decomposition K), this pre-decomosition is, as stated above, a decomposition.

In what follows, we will call sub-decompositions L of an arbitrary cellular decomposition K its sub-decompositions.

When studying the properties of cellular decompositions, it is useful to keep in mind that

Proposition 3.25. *for any compact subset C of X and any cellular decomposition K of X, the sub-decomposition $K(C)$ is finite.*

Proof. To prove this statement, it is sufficient to prove that the subset C intersects only a finite number of cells $e \in K$, since then it will be contained in a finite sub-decomposition that is the union of sub-decompositions $K(e)$ for which $e \cap C \neq \varnothing$.

Let there exist an infinite system $\{e_i\}$ of different cells $e_i \in K$ for which $e_i \cap C \neq \varnothing$ for each i, and let $x_i \in e_i \cap C$. Consider an arbitrary finite sub-decomposition L of the decomposition K. It contains only a finite number of cells e_i and $e_i \cap |L| = \varnothing$ if $e_i \notin L$. Therefore, the subspace $|L|$ contains only finitely many points x_i, so the intersection $\{x_i\} \cap |L|$ is closed. Consequently, the set Xi is closed. By similar arguments, any subset of the set $\{x_i\}$ is also closed, so the set $\{x_i\}$ is discrete. But this is impossible, since any discrete subset of a compact set must be finite. Consequently, the set C actually intersects only a finite number of cells of the decomposition K. □

It easily follows from the proven statement, for example, that

Proposition 3.26. *any pre-decomposition K, for which, for each $n = 0, 1, \ldots, \infty$, the topology of the space K^n is a weak topology with respect to the pre-decomposition K^n, constitutes a decomposition.*

Proof. Indeed, since $K = K^\infty$, we only need to prove that the pre-decomposition K is point finite, for which it is sufficient in turn to prove that for any finite n, the pre-decomposition K^n is point finite (because if $\dim e = n$, then $K(e) = K^n(e)$). We will prove this by induction on n, considering that the pre-decomposition K^0 is obviously point finite.

Let it already be proven for some $n > 0$ that the pre-decomposition K^{n-1} is point inite. Since, by assumption, the topology of the space $|K^{n-1}|$ is a weak topology with respect to this pre-decomposition, the pre-decomposition K^{n-1} is a decomposition, and therefore, according to the statement just proven, for any compact subset $C \subset |K^{n-1}|$ the sub-pre-decomposition $K^{n-1}(C) = K(C)$ is finite. In particular, for any n-dimensional cell $e^n \in K$, the sub-pre-decomposition $K(\dot{e}^{n-1})$ is finite. But, as we know, $K(e^n) = \{K(e^n), e^n\}$. Therefore, the sub-pre-decomposition $K(e^n)$ is also finite. □

In the study of cellular decompositions, an important role is also played by the fact that

Proposition 3.27. *for any increasing sequence*

$$K_0 \subset K_1 \subset \cdots \subset K_n \subset \cdots$$

of sub-decompositions of the decomposition K, whose union is the entire decomposition K, the space $X = |K|$ is a free union of subspaces

$$|K_0| \subset |K_1| \subset \cdots \subset |K_n| \subset \cdots$$

Proof. We must show that any subset A of the space X for which the intersection $A \cap |K^n|$ is closed for any n is itself closed. Since the space X is a free union of the sets $\overline{e}$, $e \in K$, it suffices to prove that for any cell $e \in K$, the intersection $A \cap e$ is closed.

Let n be a number such that $e \in K^n$. Then

$$A \cap \overline{e} = (A \cap \{K^n\}) \cap \overline{e},$$

and therefore, this intersection is closed. □

In particular,

Proposition 3.28. *for any decomposition K of the space X, this space is a free union of subspaces*

$$|K^0| \subset |K^1| \subset \cdots \subset |K^n| \subset \cdots$$

Note that the converse of the proposition proven above also holds, i.e.,

Proposition 3.29. *if a cellular pre-decomposition K of the space X is the union of an increasing sequence*

$$K_0 \subset K_1 \subset \cdots \subset K_n \subset \cdots$$

of sub-decompositions K^n, each of which is a cellular decomposition (of the subspace K^n), and if the space X is a free union of subspaces

$$|K_0| \subset |K_1| \subset \cdots \subset |K_n| \subset \cdots$$

then the pre-decomposition K is a decomposition.

Proof. Indeed, since for any cell $e \in K^n$ the equality $K(e) = K_n(e)$ holds, the pre-decomposition K is point finite. Therefore, we only need to prove that the topology of the space X is a weak topology with respect to the pre-decomposition K. Let A be an arbitrary subset of the space X for which the family $\{A \cap \overline{e}; e \in K\}$ consists of closed sets. Since

$$(A \cap |K^n|) \cap \overline{e} = A \cap \overline{e}$$

for $e \in K^n$ and since the topology of the space $|K^n|$ is a weak topology with respect to the decomposition K^n, for any $n \geq 0$ the set $A \cap |K^n|$ is closed in $|K^n|$. Therefore, the set A is closed in X. Consequently, the topology of the space X is indeed a weak topology with respect to the pre-decomposition K. □

The existence of a cellular decomposition K for a space X imposes rather strong restrictions on the topology of that space. For example, it is clear that

Proposition 3.30. *Any space X that admits a cellular decomposition K is a canonical space.*

Furthermore, since the closure $\overline{e}$ of each cell e of an arbitrary cell decomposition K is connected, any connected component A of the space $X = |K|$ either contains such a closure entirely, $A \cap \overline{e} = \overline{e}$, or does not intersect with it, $A \cap \overline{e} = \emptyset$. Since in both cases the intersection $A \cap \overline{e}$ is closed and open in $\overline{e}$, the connected component A is closed and open in X. This means that

Proposition 3.31. *any space X that admits a cellular decomposition K is a topological sum of its connected components.*

Furthermore, for each connected component $A \subset X$, the set L of all cells $e \in K$ whose closures are contained in A (i.e., for which $A \cap \overline{e} = \overline{e}$) constitutes a sub-decomposition of the decomposition K, and the body of this sub-decomposition coincides with $A : |L| = A$.

Thus,

Proposition 3.32. *any connected component A of the space X, which admits a cell decomposition K, serves as the body of some sub-decomposition $L \subset K$.*

On the other hand, since the sets $\overline{e}$, $e \in K$, are not only connected but even path-connected, the path-connected components A of the space X also possess the property that for any cell $e \in K$, either $A \cap \overline{e} = \varnothing$ or $A \cap \overline{e} = \overline{e}$, and therefore each path-connected component A is simultaneously closed and open in X. Consequently,

Proposition 3.33. *every connected space X that admits a cellular decomposition K is path-connected.*

Furthermore, it is easy to see that

Proposition 3.34. *for every space X admitting a cell decomposition K, any of its connected components is a path-connected component, and, conversely, any of its path-connected components is a connected component.*

Proof. Indeed, any connected component of the space X, being the body of some sub-decomposition of the decomposition K, is path-connected and therefore is a path-connected component. Conversely, since any path-connected component is both closed and open, it coincides with the connected component containing it. □

Since each connected component of the space X, being the body of some sub-decomposition, contains at least one vertex of the decomposition K, and since his component is path-connected, it follows that

Proposition 3.35. *any point in the space $X = |K|$ can be connected by a path to at least one of the vertices of the decomposition.*

Therefore,

Proposition 3.36. *the space $X = |K|$ is connected if its subspace K^1 is connected.*

It turns out that the converse statement is also true, i.e.,

Proposition 3.37. *if the space $X = |K|$ is connected, then its subspace K^1 is also connected.*

Proof. Indeed, let A be an arbitrary connected component of the space $|K^1|$. As we know, $A = |L|$, where L is some sub-decomposition of the decomposition of the skeleton K^1. We will show that there exists a sequence of sub-decompositions $L1 \subset L_2 \subset \cdots \subset L_n \subset \cdots$ of the form $L_n \subset K^n$, such that

1) $L_1 = L$;

2) $(L_{n+1})^n = L_n$;

3) For any cell $e \in K^n$, either $\overline{e} \cap |L_n| = \varnothing$ or $\overline{e} \cap |L_n| = \overline{e}$.

Let a sub-decomposition L_n be already constructed for some $n > 0$. From condition 3) it immediately follows that the set $|L_n|$ of this sub-decomposition is both closed and open in the space $|K^n|$. Therefore, for any $(n+1)$-dimensional cell $e^{n+1} \in K$, either $\dot{e}^{n+1} \cap |L_n| = \varnothing$ or $\dot{e}^{n+1} \cap |L_n| = \dot{e}^{n+1}$ (recall that the set $\dot{e}^{n+1}$ is connected). Therefore, the collection L_{n+1} of all cells $e \in K^{n+1}$ for which $\overline{e} \cap |L_n| \neq \varnothing$ is a sub-decomposition of the decomposition K^{n+1}. It is clear that this sub-decomposition satisfies conditions 1) - 3).

Now let us consider the sub-decomposition

$$L_\infty = \cup_{n=1}^{\infty} L_n$$

of the partition K. It is clear that its body $|L_\infty| \subset X$ has the property that for any cell $e \in K$, either $\overline{e} \cap |L_\infty| = \varnothing$ or $\overline{e} \cap |L_\infty| = \overline{e}$. Therefore, the subspace $|L_\infty|$ is closed and simultaneously open in X. Consequently, $|L_\infty| = X$, i.e., $L_\infty = K$. But it is clear that $(L_\infty)^1 = L_1 = L$. Therefore, $L = K^1$ and, consequently, $A = |L| = |K^1|$. Therefore, the subspace $|K^1|$ of the space X is connected. □

Since $(K^n)^1 = K^1$ for any $n \geq 1$, it immediately follows from the last two statements that

Proposition 3.38. *if the space $X = |K|$ is connected, then for any $n \geq 1$ its subspace $|K^n|$ is also connected; if at least for one $n \geq 1$ the subspace $|K^n|$ is connected, then the entire space X is also connected.*

As for the subspace $|K^0|$, it is easy to see that

Proposition 3.39. *for any cell decomposition K of the space X, the subspace $|K^0|$ is discrete.*

Therefore, the subspace $|K^0|$ is connected if and only if the decomposition K contains only one zero-dimensional cell (such decompositions are called *single-vertex* decompositions).

Remark 3.40 (Terminology Convention). Further, for the sake of brevity, we will not distinguish between cellular decompositions of K and the corresponding spaces $|K|$. Accordingly, we will allow expressions such as

“connected cellular decomposition of K,”

“continuous map of a decomposition of K into a decomposition of L,” and so on, meaning, respectively,

"a cellular decomposition of K for which the space $|K|$ is connected,"

"a continuous map of the space $|K|$ into the space $|L|$,"

and so on. In particular, the formula $e \in K$ will, as above, mean that the cell e belongs to the cellular decomposition of K, and the formula $e \subset K$, means that the cell e lies in the space $|K|$.

In cases where this convention could lead to misunderstandings, we will, naturally, continue to distinguish between decompositions of K and spaces $|K|$. In this case, spaces of the form $|K|$, i.e., spaces that admit cellular decompositions, will be called *cellular polyhedra.*

3.3 Theorem on paracompactness

In this section, we prove that any cellular decomposition (i.e., the space $|K|$; see the terminology convention at the end of the previous section) is paracompact, i.e., any open covering Γ of it can be refined into a locally finite covering Δ. We will construct the covering structure Δ by "ascending" step by step along the skeletons K^n of the decomposition K. To describe this construction, it is convenient for us to introduce the following general definition:

Let $X \subset Y$ and let $\{U_\alpha; \alpha \in A\}$ and $\{V_\beta; \beta \in B\}$ be families of open sets in spaces X and Y respectively. The family $\{V_\beta; \beta \in B\}$ will be called an *extension of* the family $\{U_\alpha; \alpha \in A\}$ if $A \subset B$ and

$$U_\alpha = V_\alpha \cap X$$

for every $\alpha \in A$.

In order to construct an open covering Δ of a decomposition K, it is obviously sufficient to construct a sequence of open coverings

$$\Delta^n = \{V_\beta^n; \beta \in B^n\}$$

of the skeletons K^n, having the property that for any $n > 0$ the covering Δ^n" is an extension of the covering Δ^{n-1}. Indeed, then the family Δ of sets

$$V_\beta = \cup_n V_\beta^n$$

where the summation is extended to all n for which $\beta \in B^n$, will, as is easily seen, be an open covering of the decomposition K (with the index set B being the union of the sets $B^0 \subset B^1 \subset \cdots \subset B^n \subset \cdots$).

For this covering to be refined in the given covering

$$\Gamma = \{U_\alpha; \alpha \in A\}$$

of the decomposition K, it is necessary to ensure that there exists a map α

$$\alpha : B \to A,$$

such that for any $n > 0$ the inclusion holds

$$V_\beta^n \subset U_{\alpha(\beta)} \cap K^n.$$

To carry out the proof, it will be more convenient for us to require more, namely, that the inclusion take place

$$\overline{V}_\beta^n \subset U_{\alpha(\beta)} \cap K^n.$$

Let us now consider the question of conditions that ensure local finiteness of the covering

$$\Delta = \{V_\beta; \beta \in B\}$$

Let x be an arbitrary point in the space K. To construct a neighbourhood of a point x in the space K, it is sufficient for any x for which $x \in K^n$ to define in the space K^n a neighbourhood W_x^n of the point x with the property that for $x \in K^n$

$$W_x^n \cap K^{n-1} = W_x^{n-1}.$$

Indeed, then the formula

$$W_x = \cup_n W_x^n$$

(the summation is extended to all n for which $x \in K^n$) will define for us some neighbourhood of the point x in the entire space K. This method can obviously be used to obtain any neighbourhood W_x (it suffices to set $W_x^n = W_x \cap K^n$).

Note that for $W_x \cap V_\beta \neq \varnothing$ there exists a number $n \geq 0$ such that $x \in K^n$ and

$$W_x^n \cap V_\beta^n \neq \varnothing.$$

Now suppose that the neighbourhoods W_x^n have the following two properties:

i) the intersection $W_x^n \cap V_\beta^n$ for $x \in K^m$, where $m \leq n$, is non-empty if and only if $\beta \in B^m$ and the intersection $W_x^m \cap V_\beta^m$ is non-empty;

ii) for $x \in K^m \setminus K^{n-1}$ the intersection $W_x^n \cap V_\beta^n$ is non-empty only for a finite number of indices β.

From what has just been said, it follows directly that, if conditions i) and ii) are satisfied, the intersection $W_x \cap V_\beta$ will be non-empty only for a finite number of indices β. In other words, the existence of neighborhoods W_x^n satisfying conditions i) and ii) ensures the local finiteness of the covering Δ.

Summarising all that has been said, we see that to construct a locally finite open covering Δ refined in a given open covering Γ, it is sufficient to construct

a) a sequence of open coverings $\Delta^n = \{V_\beta^n; \beta \in B^n\}$, $n = 0, 1, 2, \ldots$, each of which is an extension of the previous covering;

b) a map

$$\alpha : B \to A,$$

where $B = \cup_{n=0}^{\infty} B^n$, having the property that for any $n \geq 0$ and any $\beta \in B^n$, the inclusion

$$\overline{V}_\beta^n \subset U_{\alpha(\beta)} \cap K^n$$

holds;

c) for any $n \geq 0$ and any point $x \in K^n$ in the neighbourhood W_x^n, with the above properties i) and ii).

We will construct these objects by induction on the number n.

For $n = 0$, we take the 0-th skeleton K^0 of the decomposition K for the set B^0. We define the map α on the set B^0 by choosing, for each point $\beta \in B^0 = K^0$, the element of the covering Γ that contains it and taking $\alpha(\beta)$ as the index of this element. For the set V_β^0 corresponding to the point $\beta \in B^0$, we take this point itself. Similarly, for the set W_x^0 corresponding to the point $x \in K^0$, we also take this point itself. Since the space K^0 is discrete, it is easy to see that all our conditions are satisfied.

Let for some $n \geq 0$ we have already constructed a set B, a map α of this set into a set A, open sets $V_\beta^n \subset K^n$ and neighbourhoods $W_x^n \subset K^n$, $x \in K^n$, satisfying conditions a), b), c).

In the trivial case $K^{n+1} =^n$ we set

$$B^{n+1} = B^n, \quad V_\beta^{n+1} = V_\beta^n, \quad W_x^{n+1} = W_x^n.$$

It is clear that the index sets B^{n+1}, open sets V_β^{n+1} and neighbourhoods W_x^{n+1} constructed in this way still satisfy conditions a), b), c).

Thus, we need to consider only the case when $K^{n+1} \neq K^n$, i.e., when the decomposition K contains $n+1$-dimensional cells $^{n+1}$. For each such cell, we, having chosen some characteristic map

$$\chi : \mathbb{E}^{n+1} \to X,$$

wee put

$$U'_\alpha = \chi^{-1}(U_\alpha) = \chi^{-1}(U_\alpha \cap \bar{e}^{n+1}), \quad \alpha \in A,$$
$$V'_\beta = \chi^{-1}(V_\beta) = \chi^{-1}(V_\beta^n \cap \dot{e}^{n+1}), \quad \beta \in B.$$

The sets U'_α (many of these sets are, generally speaking, empty) constitute an open covering Γ' of the ball $\mathbb{E}^{n+1}$, and the sets V'_β (among which there may also be empty sets) constitute an open covering of its boundary sphere $\mathbb{S}^n$. Moreover, in view of condition b), for any element $\beta \in B^n$, the inclusion holds

$$\overline{V}'_\beta \subset U'_{\alpha(\beta)} \cap \mathbb{S}^n$$

(from which, in particular, it follows that the covering $\{V'_\beta\}$ is refined in the covering $\Gamma' \cap \mathbb{S}^n = \{U'_{\alpha(\beta)} \cap \mathbb{S}^n; \alpha \in A\}$). Moreover, in view of condition c) (by which the covering $\Delta^n = \{V_\beta^n; \beta \in B^n\}$ of the space K^n is locally finite), the covering $\{V'_\beta; \beta \in B^n\}$ of the sphere $\mathbb{S}^n$ is locally finite. (However, it is easy to see by induction that in fact only finitely many elements of this covering are non-empty.)

Next, for each point $x \in K^n$ we set

$$W'_x = \chi^{-1}(W_x^n) = \chi^{-1}(W_x^n \cap \dot{e}^{n+1}).$$

Clearly, the (possibly empty) sets W_x' are open in the sphere $\mathbb{S}^n$ and have the property that for each point x there are only finitely many indices β for which $W_x' \cap V_\beta' \neq \varnothing$.

For any set $G \subset \mathbb{S}^n$ and any positive $\varepsilon < 1$, we will denote by G_ε the "radial extension to ε" of the set G into the ball, i.e., the subset of the ball $\mathbb{E}^{n+1}$ consisting of all points $\boldsymbol{v} = v\boldsymbol{u} \in \mathbb{E}^{n+1}$, $0 < v < 1$, $|\boldsymbol{u}| = 1$, for which

$$1 - \varepsilon < v \leq 1, \quad \boldsymbol{u} \in G.$$

It is clear that if G is open in the sphere $\mathbb{S}^n$, then the set G_ε is open in the ball $\mathbb{E}^{n+1}$.

The sets G_ε have the property that $G_\varepsilon \cap \mathbb{S}^n = G$. Moreover, for any $\varepsilon < \varepsilon'$, the inclusion

$$G_{\varepsilon'} \subset G_\varepsilon$$

holds. Moreover, for any sets $G, H \subset \mathbb{S}^n$ and any positive $\varepsilon < 1$ and $\eta < 1$, the intersection $G_\varepsilon \cap H_\eta$ of their "radial extensions" is non-empty if and only if the intersection of the sets $G \cap H$ is non-empty. In particular, the set G_ε is non-empty if and only if the set G is non-empty.

Now, choosing for each element $\beta \in B^n$ some positive number $\varepsilon_\beta < 1$, we set

$$V_\beta^* = (V_\beta')_{\varepsilon_\beta}.$$

Since $\overline{V}_\beta' \subset U_{\alpha(\beta)}' \cap \mathbb{S}^n$, for sufficiently small ε_β the inclusion

$$V_\beta^* \subset U_{\alpha(\beta)}'$$

(and even the inclusion $\overline{V}_n^* \subset U_{\alpha^0(\beta)}'$) holds, so that all sets V_β^*, $\beta \in B^n$, form an open (and obviously locally finite) covering "adjacent" to the sphere $\mathbb{S}^n$ of the set

$$V^* = \cup_{\beta \in B^n} V_\beta^*,$$

refined in the covering $\Gamma' \cap V^* = \{U_\alpha' \cap V^*\}$.

Since the set $\mathbb{E}_{n+1} \setminus V^*$ is closed in the ball $\mathbb{E}^{n+1}$, it is compact, and therefore it can be covered by a finite number of non-empty sets of the covering Γ'. Let $U_{\alpha_1}', \ldots, U_{\alpha_s}'$ be these sets. Since V^* is open and contains the sphere $\mathbb{S}^n$, it is possible to refine the open covering $U_{\alpha_1}', \ldots, U_{\alpha_s}'$ of the set $\mathbb{E}^{n+1} \setminus V^*$ into an open covering $T_1, \ldots, T_s$ such that

$$\overline{T}_i \subset U_{\alpha_i}', \quad \overline{T}_i \cap \mathbb{S}^n = \varnothing.$$

For consistency of notation, we will write $V_{\alpha_i}^*$ instead of T_i. Here, the superscript $*$, unlike in the previous case, of course, does not mean that $V_{\alpha_i}^*$ is obtained by radial extension. Furthermore, this notation implies that the sets B^n and A do not intersect; clearly, this last assumption does not restrict generality. It is obvious that the sets V_β^*, $\beta \in B^n$, and $V_{\alpha_i}^*$, $i = 1, \ldots, s$, provide an open locally finite (essentially, even finite) covering of Δ^* of the ball $\mathbb{E}^{n+1}$, refined in the covering Γ.

Since the set $\cup_{i=1}^{s} \overline{V}^*_{\alpha_i}$ is closed and does not intersect the sphere $\mathbb{S}^n$, then for any point $x \in K^n$ there exists a positive $\varepsilon_x^{'} < 1$ such that the set

$$W_x^* = (W_x^{'})_{\varepsilon_x^{'}}$$

does not intersect with the set $\cup_{i=1}^{s} \overline{V}^*_{\alpha_i}$. The sets W_x^* are open in the ball $\mathbb{E}^{n+1}$ and have the property that

$$W_x^* \cap \mathbb{S}^n = W_x^{'}.$$

Furthermore, each of the sets W_x^* intersects with only a finite number of sets from the covering Δ^*.

Let v be an arbitrary interior point of the ball $\mathbb{E}^{n+1}$. Since the covering Δ^* is locally finite, this point has neighbourhoods that consist entirely of interior points of the ball $\mathbb{E}^{n+1}$ and intersect with only a finite number of sets of this covering. Choosing one of these neighbourhoods, we denote it by W_x^*, where $x \in e^{n+1}$ is the image of the point v under the characteristic map χ. Here, the asterisk again does not denote radial extension, but is introduced only for uniformity of notation.

All the constructed objects will have to be considered below for all cells $e^{n+1} \in K$ simultaneously. Therefore, we will introduce an additional index e^{n+1} into all notations. Thus, we will consider maps $\chi_{e^{n+1}}$, numbers $s_{e^{n+1}}$, indices $\alpha_{i,e^{n+1}}$, sets $V^*_{e^{n+1}}$, etc.

Now we have everything ready to construct the objects we need for $n+1$. We obtain the set B^{n+1} by adding to the set B^n all possible pairs of the form (i, e^{n+1}), where $e^{n+1} \in K$, $i = 1, \ldots, s_{e^{n+1}}$. We define the map α on the set B^{n+1} by the formula

$$\alpha(\beta) = \begin{cases} \alpha(\beta) & \text{if} \quad \beta \in B^n, \\ \alpha_{i,e^{n+1}} & \text{if} \quad \beta = (i, e^{n+1}), \end{cases}$$

of the set $V_\beta^{n+1} \subset K^{n+1}$, $\beta \in B^{n+1}$, by the formula

$$V_\beta^{n+1} = \begin{cases} V_\beta^n \cup_{e^{n+1} \in K} \chi_{e^{n+1}}(V^*_{\beta,e^{n+1}}), & \text{if} \quad \beta \in B^n, \\ \chi_{e^{n+1}}(V_{\alpha_i,e^{n+1}}), & \text{if} \quad \beta = (i, e^{n+1}), \end{cases}$$

and, finally, on the neighbourhood W_x^{n+1}, $x \in K^{n+1}$, by the formula

$$W_x^{n+1} = \begin{cases} W_x^n \cup_{e^{n+1} \in K} \chi_{e^{n+1}}(W^*_{x,e^{n+1}}), & \text{if} \quad x \in K^n, \\ \chi_{e^{n+1}}(W^{n+1}_{x,e^{n+1}}), & \text{if} \quad x \in e^{n+1}. \end{cases}$$

Since for any cell $e^{n+1} \in K$ the characteristic map

$$\chi_{e^{n+1}} : \mathbb{E}^{n+1} \to X$$

is homeomorphic on the open ball $\dot{\mathbb{E}}^{n+1} = \mathbb{E}^{n+1} \setminus \mathbb{S}^n$, and $V^*_{\alpha_{i,e^{n+1}}} \subset \dot{\mathbb{E}}^{n+1}$ for each $i = 1, \ldots, s_{e^{n+1}}$, then for any $\beta = (i, e^{n+1})$ the set $V_\beta^{n+1} = \chi_{e^{n+1}}(V^*_{\alpha,e^{n+1}})$

is open in $\overline{e}^{n+1}$. Consequently, this set is also open in K^{n+1}, because for any cell $e \in K^{n+1}$ different from the cell e^{n+1}, the intersection $V_\beta^{n+1} \cap \overline{e}$ is obviously empty (and therefore open in $\overline{e}$).

Furthermore, since for any $\beta \in B^n$ and any cell $e^{n+1} \in K$ the following equality holds:

$$V^*_{\beta,e^{n+1}} \cap \mathbb{S}^n = V'_{\beta,e^{n+1}} = \quad \chi^{-1}_{e^{n+1}}(V^n_\beta),$$

so the set $V^*_{e^{n+1}}$ is saturated with respect to the map $\chi_{e^{n+1}}$ (considered as a map from the ball $\mathbb{E}^{n+1}$ to the closure of the cell $\overline{e}^{n+1}$) and therefore the set $\chi_{e^{n+1}}(V^*_{e^{n+1}})$ is open in $\overline{e}^{n+1}$. On the other hand, for any cell $e \in Kn+1$ either $V_\beta n+1 \cap \overline{e} = V^n_\beta \cap \overline{e}$ (if $\dim e \leq n$) or $V^{n+1}_\beta \cap \overline{e} = \chi_{e^{n+1}}(V^*_{\beta,e^{n+1}}$ (if $e =^{n+1}$). Since in both cases the intersection $V_\beta n+1 \cap \overline{e}$ is open in $\overline{e}$, it is thus proved that the set V^{n+1}_β is open in the space K^{n+1} even for $\beta inn B^n$.

It is similarly proved that all sets of the form W^{n+1}_x are also open in the space K^{n+1}.

Let us now check conditions i) and ii) for these sets. First, we will consider condition ii).

Let $x \in K^{n+1} \setminus K^n$ and $\beta \in B^{n+1}$. Consider a cell e^{n+1} of the decomposition K containing the point x. By definition,

$$W^{n+1}_x = \chi_{e^{n+1}}(W^*_{x,e^{n+1}}),$$

and therefore the intersection $W^{n+1}_x \cap V^{n+1}_\beta$ can be non-empty, only when $\beta \in B^n$ or when $\beta = (i, e^{n+1})$, where $i = 1, \ldots, s_{e^{n+1}}$. Since the number of indices β of the form (i, e^{n+1}) is finite, it suffices to prove that the number of those $\beta \in B^n$ for which this intersection is non-empty is also finite. But it is clear that for $\beta \in B^n$

$$W^{n+1}_x \cap V^{n+1}_\beta = \chi_{e^{n+1}}(W^*_{x,e^{n+1}} \cap V^*_{\beta,e^{n+1}}),$$

and therefore, according to the choice of the set $W^*_{x,e^{n+1}}$, the intersection $W^{n+1}_x \cap V^{n+1}_\beta$ is non-empty only for a finite number of indices $\beta \in B^n$.

Let us now check condition i). Let $m < n+1$, $\beta \in B^{n+1}$ and $x \in K^m$. Since $m < n+1$, then

$$W^{n+1}_x = W^n_x \cup \cup_{e^{n+1} \in K} \chi_{e^{n+1}}(W^*_{x,e^{n+1}})$$

Therefore if $\beta = (i, e^{n+1})$, then

$$W^{n+1}_x \cap V^{n+1}_\beta = \chi_{e^{n+1}}(W^*_{x,e^{n+1}} \cap V^*_{\beta,e^{n+1}}) = \emptyset.$$

If $\beta \in B^n$, then due to inclusions

$$\chi_{e^{n+1}}(W^*_{x,e^{n+1}}) \cap K^n \subset W^n_x,$$
$$\chi_{e^{n+1}}(V^*_{\beta,e^{n+1}}) \cap K^n \subset V^n_\beta$$

the equality will hold

$$W^{n+1}_x \cap V^{n+1}_\beta = (W^n_x \cap V^n_\beta) \cup \cup_{e^{n+1} \in K} \chi_{e^{n+1}}(W^*_{x,e^{n+1}} \cap V^*_{\beta,e^{n+1}}).$$

Let for some cell $e^{n+1} \in K$ the intersection

$$W^*_{x,e^{n+1}} \cap V^*_{\beta,e^{n+1}}$$

is non-empty. Then the intersection is non-empty

$$W'_{x,e^{n+1}} \cap V'_{\beta,e^{n+1}} = \chi^{-1}_{e^{n+1}}(W^n_x \cap V^n_\beta),$$

and the intersection as well

$$W^n_x \cap V^n_\beta.$$

It follows from this that the intersection $W^{n+1}_x \cap V^{n+1}_\beta$, $\beta \in B^n$, is non-empty if and only if the intersection $W^n_x \cap V^n_\beta$ is non-empty, i.e., by the induction hypothesis, when $\beta \in B^m$ and the intersection $W^m_x \cap V^m_\beta$ is non-empty. Thus, the construction of coverings Δ^n and neighbourhoods W^n_x by induction is accomplished for all n. According to the above, it is thus proved that

Proposition 3.41. *any open covering Γ of a cellular decomposition K can be refined into a locally finite open covering Δ.*

In other words, we have proved that

Proposition 3.42. *any cellular decomposition of K is a paracompact space.*

Since the space K is, by assumption, Hausdorff, it follows directly from this statement that

Proposition 3.43. *any cellular decomposition of K is a normal (and even stably normal) space.*

Let us now consider the characteristic map $\chi_{K\times L} : P_{K\times L} \to K \times L$ for the pre-decomposition $K \times L$ which is the product of two cellular decompositions K and L. By what has just been proved, the spaces K and L are regular. Moreover, as we know, if the decomposition K is locally finite (resp. locally countable), then the characteristic map $\chi_K : P_K \to K$ is locally compact (resp. locally countable). On the other hand, as noted at the end of §3.1, we can assume that $P_{K\times L} = P_K \times P_L$ and $\chi_{K\times L} = \chi_K \times \chi_L$. Therefore, by the theorem proved in §1.5, the map $\chi_{K\times L}$ is an identification map if at least one of the decompositions K and L is locally finite or if both these decompositions are locally countable. Since the pre-decomposition $K\times L$ is obviously point-wise finite, it is thus proved that

Proposition 3.44. *The product $K \times L$ of cellular decompositions K and L is a cellular decomposition if at least one of the decompositions K and L is locally finite or if both these decompositions are locally countable.*

In particular,

Proposition 3.45. *for any cellular decomposition K, the product $K \times I$ is also a cellular decomposition.*

Remark 3.46. As Dowker showed, there exist cellular decompositions K and L such that their product $K \times L$ is not a cellular decomposition.

3.4 Continuous maps of cellular decomposition

Let K be an arbitrary cellular decomposition and Y be an arbitrary topological space.

Since the decomposition K is a free union of sets $\overline{e}$, $e \in K$, then

Proposition 3.47. *the single-valued map*

$$f : K \to Y$$

is continuous if and only if for any cell $e \in K$ the map

$$f|_{\overline{e}} : \overline{e} \to Y$$

is continuous.

Similarly, since a cellular decomposition K is a free union of its skeletons K^n, then

Proposition 3.48. *the map*

$$f : K \to Y$$

is continuous if and only if for any finite $n \geq 0$ the map

$$f|_{K^n} : K^n \to Y.$$

is continuous.

In what follows, we will repeatedly construct maps

$$f : K \to Y.$$

by constructing, for any $n \geq 0$, "consistent" continuous maps f^n, i.e., continuous maps

$$f^n : K^n \to Y.$$

such that

$$f^{n+1}|_{K^n} = f^n$$

for any $n \geq 0$, and setting for any point $x \in K$

$$f(x) = f^n(x), \quad \text{if} \quad x \in K^n.$$

According to the previous assertion,

Proposition 3.49. *the map*

$$f : K \to Y.$$

constructed in this way is continuous.

All these statements are applicable, in particular, to the decomposition $K \times I$ and to continuous maps $K \times I \to Y$, i.e., to homotopies $f_t : K \to Y$, $0 \leq t \leq 1$. Since the decomposition $K \times I$ is obviously a free union of sets $\overline{e} \times I$, we obtain from this that

Proposition 3.50. *the family of maps*

$$f_t : K \to Y, \quad 0 \le t \le 1,$$

is a homotopy if and only if for any cell $e \in K$ the homotopy is the family

$$f_t|_{\overline{e}} : \overline{e} \to Y, \quad 0 \le t \le 1,$$

and also if and only if for any $n \ge 0$ the homotopy is the family

$$f_t|_{K^n} : K^n \to Y.$$

Moreover,

Proposition 3.51. *if for all $n \ge 0$ we are given "consistent" skeleton homotopies, i.e., homotopies*

$$fn_t : K^n \to Y, \quad 0 \le t \le 1,$$

such that

$$f_t^{n+1}|_{K^n} : \overline{e} \to Y, \quad 0 \le t \le 1,$$

for any $n \ge 0$, then, setting for each point $x \in K$ and any $t \in I$

$$f_t(x) = f_t^n(x), \quad \text{if} \quad x \in K^n,$$

we obtain some homotopy

$$f_t : K \to Y, \quad 0 \le t \le 1.$$

Sometimes we will consider not the sequence $K^0, K^1, \dots, K^n, \dots$ of skeletons, but an arbitrary increasing sequence

$$K_0 \subset K_1 \subset \cdots \subset K_n \subset \cdots$$

of sub-decompositions of the decomposition K, the union of which is the entire decomposition K. Since the decomposition K is the free union of any such sequence, then

Proposition 3.52. *All the assertions proved above remain valid even if the sequence of skeletons is replaced by an arbitrary increasing sequence of decompositions of the decomposition K, the union of which is this entire decomposition.*

On the other hand, since any closed (or open) subspace X of the decomposition K is a free union of sets $X \cap \overline{e}$, $e \in K$, then everything said above is true (with obvious modifications) for any such space. Thus, for example,

Proposition 3.53. *the map*

$$f : X \to Y$$

is continuous if and only if for any cell $e \in K$ the map

$$f|_{X \cap \overline{e}} : X \cap \overline{e} \to Y$$

is continuous, and also if and only if for any $n \geq 0$ the map

$$f|_{X \cap K^n} : X \cap K^n \to Y$$

is continuous.

A pair (K, L) consisting of a cellular decomposition K and an arbitrary sub-decomposition L of it will be called a *cellular pair*. Using the above remarks, one can easily prove that the extension property of maps of spheres, which underlies the definition of m-connected spaces (see §2.1), holds for any cellular pairs (K, L) for which $\dim(K \setminus L) \leq m+1$, i.e., for which $\dim e \leq m+1$ for each cell $e \in K \setminus L$. Namely,

Proposition 3.54. *if $\dim(K \setminus L) \leq m + 1$, then any map $f : L \to Y$ of a sub-decomposition L into an m-connected ($m \geq 0$) space Y can be extended to some map $g : K \to Y$.*

Proof. Indeed, let

$$K_n = K^n \cup L.$$

For any $n = 0, 1, \ldots, m + 1$, we construct a continuous map

$$g_n : K_n \to Y,$$

such that $g_n|_L = f$, $g_{n+1}|_{K_n} = g_n$. Then the map g_{m+1} will be the desired map g (since, according to the condition, $K_{m+1} = K$).

We will construct the map g_0 by arbitrarily defining it on the vertices of the decomposition K that do not belong to the sub-decomposition L (on L, it must, of course, coincide with f). Clearly, this map is continuous.

Let g_n already be constructed for some non-negative $n \leq m$. If $K_{n+} = K_n$, then we set ${}_{n+1} = g_n$. Let ${}_{n+1} \neq K_n$ and let e^{n+1} be an arbitrary $n+1$-dimensional cell of the decomposition K that does not belong to the sub-decomposition L. Having chosen a characteristic map for each such cell,

$$\chi_{e^{n+1}} : \mathbb{E}^{n+1} \to K,$$

we consider the composition

$$g_n \circ \chi_{e^{n+1}}|_{\mathbb{S}^n} : \mathbb{S}^n \to Y,$$

of the restriction $\chi_{e^{n+1}}|_{\mathbb{S}^n}$ of the map $\chi_{e^{n+1}}$ to the sphere $\mathbb{S}^n$ and the ma g_n. Since the space Y is, by assumption, m-connected, and $n \leq m$, we can extend this composition to some map

$$\chi_{e^{n+1}} : \mathbb{E}^{n+1} \to Y.$$

For any point $x \in K_{n+1}$, we now set

$$g_{n+1}(x) = \begin{cases} g_n(x), & \text{if} \quad x \in K_n, \\ \chi_{e^{n+1}}(\chi_{e^{n+1}}^{-1}(x)) & \text{if} \quad x \notin K_n, \end{cases}$$

where e^{n+1} is the cell of the decomposition K containing the point x. It is clear that we thereby obtain a single-valued continuous map

$$g_{n+1} : K_{n+1} \longrightarrow Y,$$

possessing all the required properties. □

Let us now show that

Proposition 3.55. *any cellular pair* (K, L) *satisfies the axiom of homotopy extension.*

Proof. Indeed, let Y be an arbitrary space, $f : K \longrightarrow Y$ an arbitrary map, and $g_t : L \longrightarrow Y$, $0 \leq t \leq 1$, be a homotopy such that $g_0 = f|_L$. We must construct a homotopy $f_t : K \longrightarrow Y$ for which $f_0 = f$ and $f_t|_L = g_t$ for any $t \in I$. Let $K_n = Kn \cup L$. According to the above, to construct a homotopy f_t, it is sufficient to construct for any $n \geq 0$ a homotopy

$$f_t^n : K_n \longrightarrow Y,$$

such that

$$f_0^n = f|_{K_n}, \quad f_t^n|_L = g_t, \quad \text{and} \quad f_t^{n+1}|_{K_n} = f_t^n.$$

For any point $x \in K_0$ we define a homotopy f_t^0 by the formula

$$f_t^0 = \begin{cases} g_t(x), & \text{if} \quad x \in L, \\ f(x), & \text{if} \quad x \in K^0. \end{cases}$$

It is clear that this formula indeed defines a homotopy, for which $f_0^0 = |_{K_0}$ and $f_t^0 = g_t$ for any $t \in I$.

Let for $n \geq 0$ we have already constructed a homotopy f_t^n. If $K_{n+1} = K_n$, then we set $f_t^{n+1} = f_t^n$, $0 \leq t \leq 1$. Let $K_{n+1} \neq K_n$ and let e^{n+1} be an arbitrary $n+1$-dimensional cell of K that does not belong to L. Having chosen for each such cell e^{n+1} a characteristic map

$$\chi_{e^{n+1}} : \mathbb{E}^n \longrightarrow K,$$

For each point $x = \chi_{e^{n+1}}(v\boldsymbol{u})$, $0 \leq v \leq 1$, $|\boldsymbol{u}|$, of the set $\overline{e}^{n+1}$ and any number $t \in I$ we set

$$g_{e^{n+1},t}(x) = \begin{cases} f\left(\chi_e^{n+1}\left(\frac{2v}{2-t}\boldsymbol{u}\right)\right), & \text{if} \quad 0 \leq v \leq \frac{2-t}{2}, \\ f_{2v+t-2}^n(\chi_e^{n+1}(\boldsymbol{u})), & \text{if} \quad \frac{2-t}{2} \leq v \leq 1. \end{cases}$$

It is easy to see (Fig. 3.1) that this formula uniquely defines a certain map

$$g_{e^{n+1},t} : \overline{e}^{n+1} \longrightarrow Y$$

(depending on the choice of the map χ). Moreover, since the family of maps

$$g_{e^{n+1},t} \circ \chi_{e^{n+1}} : \mathbb{E}^{n+1} \longrightarrow Y$$

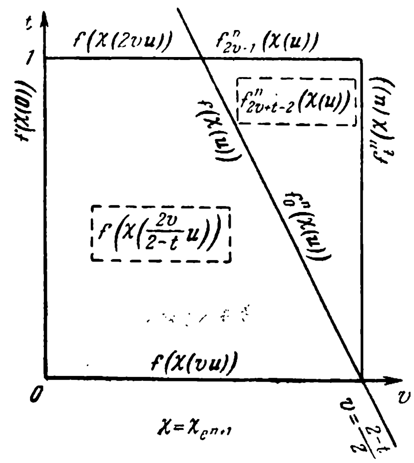


Figure 3.1:

is obviously a homotopy, the family of maps $g_{e^{n+1},t}$ is also a homotopy (since the characteristic map $\chi_{e^{n+1}}$ considered as the map $\mathbb{E}^{n+1} \to \overline{e}^{n+1}$ is an identity map, and the segment I is compact, so the map $\chi_{e^{n+1}} \times 1 : \mathbb{E}^{n+1} \times I \to \overline{e}^{n+1} \times I$ is also an identity map). This homotopy has the property that

$$g_{e^{n+1},0} = f|_{\overline{e}^{n+1}} \quad \text{and} \quad g_{e^{n+1},t}|_{\dot{e}^{n+1}} = f_t^n|_{\dot{e}^{n+1}}$$

for any $t \in$. Therefore, the formula

$$f_t^{n+1} = \begin{cases} f_t^n(x), & \text{if} \quad x \in K_n, \\ g_{e^{n+1},t}(x) & \text{if} \quad x \in e^{n+1} \in K_{n+1} \setminus K_n, \end{cases}$$

defines a certain homotopy $f_t^{n+1} : K_{n+} \to Y$, $0 \leq t \leq 1$, which obviously has all the required properties.

Thus, the homotopies f_t^n are constructed, by induction, for all $n \geq 0$. The proposition formulated above is thus completely proven. □

Remark 3.56. A special case of the proved proposition is the statement formulated in §2.4, since all pairs listed in this statement, as is easy to see, are cellular.

We see, therefore, that all the results of §2.3 are applicable to arbitrary cellular pairs (K, L), since, together with the pair (K, L), the pair $(K \times I, I(K, L))$ is also cellular. Thus, firstly,

Proposition 3.57. *for any cellular pair* (K, L) *of any space* Y *and any two homotopic maps*

$$f, g : L \to Y$$

the spaces $K \cup_f Y$ and $K \cup_g Y$ are homotopically equivalent moreover, the corresponding homotopic equivalence can be chosen so that it is the identity map on K,

and, secondly,

Proposition 3.58. *For any cell pair (K, L), any space Y, any continuous map*

$$f : L \to Y$$

and any space Z homotopically equivalent to Y, every homotopy equivalence

$$h : Y \to Z$$

can be extended to some homotopy equivalence

$$H : K \cup_f Y \to K \cup_{h \circ f} Z.$$

Generally speaking, the space $K \cup_f Y$ is not a cellular decomposition even in the case when the space Y is a cellular decomposition. One can only say that

Proposition 3.59. *for any cellular pair (K, L), any cellular decomposition Q and any continuous map*

$$f : L \to Q$$

the space $K \cup_f Q$ is homotopy equivalent to some cellular decomposition $\widetilde{Q}$ that contains the decomposition Q as a sub-decomposition, and the homotopy equivalence $K \cup_f Q \to \widetilde{Q}$ can be chosen such that it is the identity map on Q.

To prove this statement, we introduce continuous maps

$$f : K \to Q$$

of a cellular decomposition K into a cellular decomposition Q, with the property that

$$f(K^n) \subset Q^n$$

for any $n \geq 0$. We will call such maps of cellular decompositions *cellular*. The statement formulated above will obviously be proven if we show, first, that

Proposition 3.60. *any continuous map $K \to Q$ is homotopic to some cellular map $K \to Q$*

and, secondly, that

Proposition 3.61. *For any cellular pair (K, L), any cellular decomposition Q, and any cellular map,*

$$f : L \to Q$$

the space $K \cup_f Q$ is a cellular decomposition containing the decomposition Q as a sub–decomposition.

Let us first consider the second statement. Since the cellular decompositions K and Q are, as we know, normal spaces, the space $K \cup_f Q$ is Hausdorff (§1.6). Moreover, since the natural projection $\alpha : K \cup Q \to K \cup_f Q$ onto $K \setminus L$ and Q is a homeomorphism, then for any cell $e \in (K \setminus L) \cup Q$ (with characteristic map χ) the set $\alpha(e)$ is a cell of the space $K \cup_f Q$ (with characteristic map $\alpha \circ \chi$). Moreover, given that the map f is cellular, all cells $\alpha(e)$, $e \in (K \setminus L) \cup Q$ constitute a cellular pre-decomposition of this space.

Let us show that this subdivision is point-finite, i.e., that any cell $\alpha(e)$ of it belongs to some finite subdivision. For $e \in Q$, this is obvious, since the map $\alpha|_Q$ is homeomorphic, and therefore the image $\alpha(Q(e))$ of the sub-decomposition $Q(e)$ under the map α is a finite sub-decomposition of the pre-decomposition $K \cup_f Q$, containing the cell $\alpha(e)$. Let $e \in K \setminus L$. Since the sub-decomposition $K(e) \cap L$ is finite, it is compact, and therefore its image $f(K(e) \cap L) \subset Q$ under the map f is contained in some finite sub-decomposition Q_e of the decomposition Q. Consider the set $\alpha(K(e) \cup Q_e)$. It is obviously compact. Furthermore, it is the union of cells of the form $\alpha(e)$, $e \in (K(e) \setminus L) \cup Q_e$. Therefore, this set is a finite sub-decomposition of the pre-decomposition $K \cup_f Q$, containing the cell $\alpha(e)$.

Thus, to prove the statement under consideration, we only need to prove that the topology of the space $K \cup_f Q$ is a weak topology. As we know, for this it is sufficient to prove (see §3.1), that the set $A \subset K \cup_f Q$ is closed if its intersection $a \cap P$ with any finite sub-decomposition $P \subset K \cup_f Q$ is closed. At the same time, since the natural projection α is an identification map, and the topological sum $K \cup Q$ is a cellular decomposition, the set $A \subset KK \cup_f Q$ is closed if and only if for any cell $e \in K \cup Q$ the intersection $\alpha^1(A) \cap \overline{e}$ is closed. Thus, we need to prove that if the intersection $A \cap P$ of some set $A \subset K \cup_f Q$ with any finite sub-decomposition $P \subset K \cup_f Q$ is closed, then for any cell $e \in K \cup Q$ the intersection $\alpha^{-1}(A) \cap \overline{e}$ is closed. It is easy to see that this fact will be proved if we show that for any cell $e \in K \cup Q$ the set $\alpha(e)$ is contained in some finite sub-decomposition P_e of the pre-decomposition $K \cup_f Q$.

Indeed, since the subdivision P_e, being finite, is closed, and $\alpha(e) \subset P_e$, then $\alpha(\overline{e}) \subset \overline{\alpha(e)} \subset P_e$, and therefore $\overline{e} \subset \alpha^{-1}(P_e)$. Therefore,

$$\alpha^{-1}(A) \cap \overline{e} \subset \alpha^{-1}(A) \cap \alpha^{-1}(P_e) \cap \overline{e} \subset \alpha^{-1}(A \cap P_e) \subset \alpha^{-1}(A) \cap \overline{e},$$

and therefore

$$\alpha^{-1}(A) \cap \overline{e} = \alpha^{-1}(A \cap P_e) \cap \overline{e}.$$

By assumption, the right-hand side of this equality is closed. Therefore, its left-hand side is also closed.

Thus, we only need to prove that for any cell $e \in K \cup Q$ there exists a finite sub-decomposition $P_e \subset K \cup_f Q$ such that $\alpha(e) \subset P_e$. But if $e \notin L$, then the set $\alpha(e)$ is a cell of the pre-decomposition $K \cup_f Q$ and therefore, as has already been proved, is contained in some finite sub-pre-decomposition P_e of this pre-decomposition. Let $e \in L$. Since the set $f(\overline{e}) \subset Q$ is compact, it is contained in some finite sub-decomposition of the decomposition Q. Therefore, in this case, the set $\alpha(e) = f(e) \subset f(\overline{e})$ (generally speaking, no longer a cell) is contained in some finite sub-decomposition P_e of the pre-decomposition $K \cup_f Q$.

Thus, the statement that the pre-decomposition $K \cup_f Q$ is a decomposition is completely proven.

In particular, we see that

Proposition 3.62. *for any cellular pair (K, L) the space K/L is a cellular decomposition.*

Similarly,

Proposition 3.63. *the cylinder Z_f of any cellular map $f : K \to L$ is a cellular decomposition (with cells of the form $e \in K$, $e \in L$ and $e \times (0, 1)$, $e \in K$).*

Remark 3.64. For the above statement to be true, the requirement that the map f be cellular is not necessary. It is sufficient that this map have the property that $f(L \cap \dot{e} \subset Q^{n-1}$ for any cell e $K \setminus L$, where $n = \dim e$.

As for the first of the above statements, it can even be strengthened somewhat. Namely, it turns out that

Proposition 3.65. *any continuous map $K \to Q$ that is a cellular map on some sub-decomposition $L \subset K$ is homotopic to the cellular map $K \to Q$ $\operatorname{rel} L$.*

We will prove this fundamental theorem about cellular maps in the next section. We will dedicate the end of this section to deducing one important consequence from it.

A homotopy

$$f_t : K \to Q, \quad 0 \le t \le 1,$$

will be called *cellular* if the maps f_0 and f_1 are cellular and if

$$f_t(K^n) \subset Q^{n+1} \quad \text{for any} \quad n \ge 0.$$

Clearly, a homotopy $f_t : K \to Q$, $0 \le t \le 1$, is cellular if and only if the corresponding map $K \times I \to Q$ is a cellular map. Cellular maps associated by a cellular homotopy will be called *cellularly homotopic*.

It follows from the theorem on cellular maps that

Proposition 3.66. *if cellular maps*

$$f, g : K \to Q$$

are homotopic relative to some (possibly empty) sub-decomposition $L \subset K$, then they are also cellularly homotopic rel L.

Proof. Indeed, the statement that the maps f and g are homotopic rel L means that there exists a map

$$F : K \times I \to Q,$$

such that

$$F(x, 0) = f(x), \quad F(x, 1) = g(x)$$

for any point $x \in K$ and

$$F(x,t) = f(x) = g(x)$$

for any points $x \in L$ and $t \in I$. It is clear that the map F is cellular on the sub-decomposition

$$I(K,L) = L \times \cup K \times I \cup L \times 1$$

of the decomposition $K \times I$ and therefore, according to the theorem on cellular maps, it is homotopic relative to this sub-decomposition to some cellular map

$$G : K \times I \to Q.$$

The family of maps corresponding to G

$$g_t : K \to Q, \quad 0 \le t \le 1,$$

obviously represents a cellular homotopy rel L, connecting the maps f and g. □

Thus, when studying the homotopy properties of cellular decompositions, we can, without loss of generality, restrict ourselves to considering only cellular maps and their cellular homotopies.

Let us apply, for example, this remark to the study of homotopy groups $\pi_n(X; x_0)$ in the case where the space X is a cellular decomposition of K. Here we can obviously assume that the point x_0 is some vertex e^0 of the decomposition K.

Having agreed to regard the sphere $\mathbb{S}^n$ as a cellular decomposition with a zero-dimensional cell $\boldsymbol{u}_0$ and an n-dimensional cell $\mathbb{S}^n \setminus \boldsymbol{u}_0$, we apply to an arbitrary map

$$\varphi : (\mathbb{S}^n, \boldsymbol{u}_0) \to (K, e^0)$$

the theorem on cellular maps. Since the points $\boldsymbol{u}_0$ and e^0 are vertices, this map is cellular on $\boldsymbol{u}_0$. Therefore, according to the theorem on cellular maps, it is homotopic rel $\boldsymbol{u}_0$ to some cellular map

$$\psi : (\mathbb{S}^n, \boldsymbol{u}_0) \to (K, e^0).$$

But the last map, being cellular, is a map of the sphere $\mathbb{S}^n$ into the n-dimensional skeleton K^n of the decomposition K. Therefore, any map $(\mathbb{S}^n, \boldsymbol{u}_0) \to (K, e^0)$ is homotopic rel $\boldsymbol{u}_0$ to a map of the form $i \circ \psi$, where $i : K^n \to K$ is an inclusion, and ψ is some map $(\mathbb{S}^n, \boldsymbol{u}_0) \to (K^n, e^0)$. This means that

Proposition 3.67. *the inclusion $i : K^n \to K$ defines an epimorphism*

$$i_* : \pi_n(K^n; e^0) \to \pi_n(K; e^0)$$

of the group $\pi_n(K^n; e^0)$ onto the group $\pi_n(K; e^0)$.

Similarly, using the fact that homotopic cellular maps are cellularly homotopic, we obtain that

Proposition 3.68. *the group $\pi_n(K; e^0)$ is isomorphic to the group $\pi_n(K^{n+1}; e^0)$ (and, in general, to any group $\pi_n(K^m; e^0)$ with $m \geq n + 1$).*

In particular, it directly follows from this general statement that

Proposition 3.69. *for any $m < n$, the following equality holds:*

$$\pi_n(\mathbb{S}^m, \boldsymbol{u}_0) = 0,$$

i.e., that the sphere $\mathbb{S}^m$ is an $(m-1)$-connected space.

We emphasise, however, that we obtained this statement by relying on the theorem about cellular maps. On the other hand, in the next section, when proving the last theorem, we will make substantial use of this statement. Therefore, to avoid a vicious circle, we will be forced to provide an independent proof of it there.

3.5 Proof of the theorem on cellular maps

Let Q be an arbitrary cellular decomposition and

$$\varnothing = Q_1 \subset Q_0 \subset \cdots \subset Q_n \subset \cdots$$

be an arbitrary increasing sequence of its sub-decompositions (the union of these sub-decompositions may be the decomposition Q or may not). A map

$$f : K \to Q$$

of some cellular decomposition K into a decomposition Q will be called *subordinate to* the sequence $\{Q_n\}$ if

$$f(K^n) \subset Q_n$$

for any $n \geq 0$. Accordingly, we will say that for a sequence $\{Q_n\}$ *the theorem on subordinate maps holds* if for any cellular pair (K, L) each map $f : K \to Q$ whose restriction $f|_L$ is subordinate to $\{Q_n\}$ is homotopic rel L to the subordinate $\{Q_n\}$ map $g : K \to Q$.

In the case where the sequence $\{Q_n\}$ consists of skeletons Q^n of the decomposition Q, the subordination of the map f means its cellularity.

Thus, the theorem on cellular maps formulated in the previous section means that

Proposition 3.70. *for the sequence $\{Q_n\}$ of skeletons of the decomposition Q, the theorem on subordinate maps holds.*

The basis of the proof of the theorem on cellular maps is the fact that

Proposition 3.71. *if the sequence $\{Q_n\}$ is such that for any n^0 every map $(\mathbb{E}^n, \mathbb{S}^{n-1}) \to (Q, Q_{n-1})$ is homotopic* rel $\mathbb{S}^{n-1}$ *to some map $(\mathbb{E}^n, \mathbb{S}^{n-1}) \to (Q_n, Q_{n-1})$, then the theorem on subordinate maps holds for the sequence $\{Q_n\}$.*

In other words, if the sequence $\{Q_n\}$ has the indicated property, then every continuous map $f : K \to Q$ subordinated on L to the sequence $\{Q_n\}$ is homotopic rel L to some map $g : K \to Q$ subordinated to the sequence $\{Q_n\}$.

Proof. According to the remarks made at the beginning of the previous section, to construct a homotopy

$$f_t : K \to Q \ \text{rel} \ L, \quad 0 \le t \le 1,$$

connecting the map f with the map g, it is sufficient for us to construct for each $n \ge 0$ a homotopy

$$f_t^n : K \to Q, \quad 0 \le t \le 1,$$

such that

$$f_0^n = f|_{K^n}, \quad f_1^n(K^n) \subset Q_n, \quad f_t^n|_{L^n} = f|_{L^n}, \quad f_t^{n+1}|_{K^n} = f_t^n$$

for a $n \ge 0$ and $t \in I$.

By hypothesis, every map $(\mathbb{E}^n, \mathbb{S}^{n-1}) \to (Q, Q_{n-1})$ is a homotopy rel $\mathbb{S}^{n-1}l$ to some map $(\mathbb{E}^n, \mathbb{S}^{n-1}) \to (Q_n, Q_{n-1})$. For $n = 0$, this means that for any point $y \in Q$, there exists a path $u_y : I \to Q$ such that $u_y(0) = y$ and $u_y(1) \in Q_0$. For every point $x \in K^0$ and any number $t \in I$ we put

$$f_t^0(x) = \begin{cases} u_{f(x)}, & \text{if} \quad x \notin L, \\ f(x), & \text{if} \quad x \in L. \end{cases}$$

It is clear that this will give us a homotopy

$$f_t^0 : K^0 \to Q, \quad 0 \le t \le 1,$$

for which

$$f_0^0 = f|_{L^0}, \quad f_1^0(K^0) \subset Q_0, \quad f_t^0|_{L^0} = f|_{L^0}.$$

Let for $n \ge 0$ a homotopy f_t^n has already been constructed. If $K^{n+1} = K^n$, then we set $f_t^{n+1} = f_t^n$. Let $K^{n+1} \ne K^n$. Since the pair (K^{n+1}, K^n), being cellular, satisfies the axiom of homotopy extension, there exists a homotopy

$$g_t : K^{n+1} \to Q, \quad 0 \le t \le 1,$$

such that

$$g_0 = f|_{K^{n+1}}, \quad \text{and} \quad g_t|_{K^n} = f_t^n$$

for any $t \in I$.

Let's consider the map

$$g_1 : K^{n+1} \to Q.$$

Since $g_1|_{K^n} = f_t^1$, then

$$g_1(K^n) \subset Q_n.$$

Let e^{n+1} be an arbitrary $n+1$-dimensional cell of the decomposition K and let

$$\chi : \mathbb{E}^{n+1} \to K$$

be an arbitrary characteristic map of this cell. Since $\dot{e}^{n+1} \subset K$ and $g_1(K^n) \subset Q_n$, then

$$g_1 \circ \chi : (\mathbb{E}^{n+1}, \mathbb{S}^n) \to (Q, Q_n)$$

Therefore, according to the condition, there exists a homotopy

$$\xi_t : \mathbb{E}^{n+1} \to Q \;\; \text{rel} \;\; \mathbb{S}^n,$$

such that

$$\xi_0 = g_1 \circ \chi, \quad \xi_1(\mathbb{E}^{n+1}) \subset Q_{n+1}.$$

For any point $\boldsymbol{x} = \chi(v\boldsymbol{u})$, $0 \le v \le 1$, $|\boldsymbol{u}| = 1$, any cell $e^{n+1}1$, and any number $t \in I$, we set

$$h_{e^{n+1},t}(x) = \begin{cases} g_{\frac{2t}{1+v}}(\chi(v\boldsymbol{u})), & \text{if} \quad 0 \le t \le \frac{1+v}{2}, \\ \xi_{\frac{2t-v-1}{1-v}}(v\boldsymbol{u}), & \text{if} \quad \frac{1+v}{2} \le t \le 1. \end{cases}$$

It is easy to verify (Fig. 3.2) that we thereby obtain a certain uniquely defined

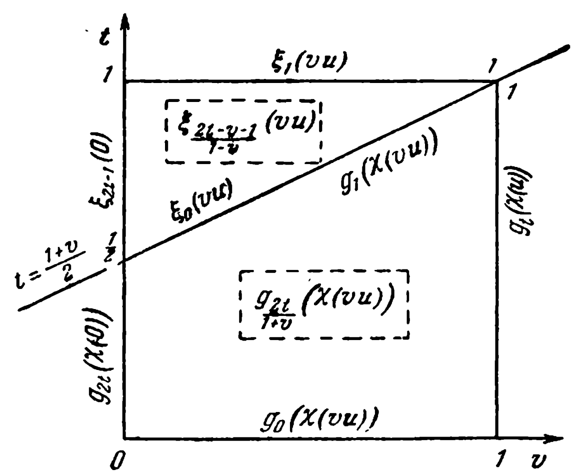


Figure 3.2:

homotopy

$$h_{e^{n+1},t} : \overline{e}^{n+1} \to L, \quad 0 \le t \le 1.$$

To construct a homotopy f_t^{n+1}, we now consider an arbitrary point $x \in K^{n+1}$. In the case where $x \notin K^n$, the point x belongs to some uniquely determined $n+1$-dimensional cell $e_x^{n+1} \in K$. For any number $t \in I$

$$f_t^{n+1}(x) = \begin{cases} f_t^n(x), & \text{if} \quad x \in K^n, \\ f(x), & \text{if} \quad x \notin K^n \quad \text{and} \quad e_x^{n+1} \in L, \\ h_{e_x^{n+1}}(x), & \text{if} \quad x \notin K^n \quad \text{and} \quad e_x^{n+1} \notin L. \end{cases}$$

It is easy to see that this gives us a homotopy

$$f_t^{n+1} : K^{n+1} \to Q,$$

such that

$$f_0^{n+1} = f|_{K^{n+1}}, \quad f_1^{n+1}(K^{n+1}) =\subset Q_{n+1}$$

and

$$f_t^{n+1}|_{L^{n+1}} = f_{K^{n+1}}, \quad f_t^{n+1}|_{K^n} = f_t^n \text{ for any } t \in I.$$

Thus, the homotopies f_t^n are constructed for all $n \geq 0$.

The proposition formulated above is completely proved. □

Now consider the following statement:

Proposition 3.72 (A_n)**.** *Any map*

$$(\mathbb{E}^n, \mathbb{S}^{n-1}) \to (Q, Q^{n-1})$$

is homotopic rel $\mathbb{S}^{n-1}$) *to some map*

$$(\mathbb{E}^n, \mathbb{S}^{n-1}) \to (Q^n, Q^{n-1}).$$

According to the proposition just proved,

Proposition 3.73. *To prove the theorem on cellular maps, it suffices to prove statement (A_n) for any $n \geq 0$.*

Remark 3.74. For $n = 0$, the statement (A_n) is true (since in this case it simply states that any point of the partition Q can be connected by a path to some of its vertices). Thus, this statement requires proof only for $n > 0$.

A pair (X, X_0) consisting of some Hausdorff space X and its closed subspace X_0, which has the property that the complement $X \setminus X_0$ is an m-dimensional $(m \geq 0)$ cell e^m, we will call an *m-dimensional relative cell.*

Consider the following statement:

Proposition 3.75 (B_n)**.** *For any $m > n$ and any m-dimensional relative cell (X, X_0), every continuous map f of the ball $\mathbb{E}^n$ into the space X that maps the sphere $\mathbb{S}^{n-1}$ into the subspace X_0, i.e., a map of the pair $(\mathbb{E}^n, \mathbb{S}^{n-1})$ into the pair (X, X_0), is homotopic* rel $\mathbb{S}^{n-1}$ *to some map $\mathbb{E}^n \to X_0$.*

It is easy to see that

Proposition 3.76. *if for some $n \geq 0$ statement B_n is true, then statement A_n is also true (for any cellular decomposition Q and any map $f : (\mathbb{E}^n, \mathbb{S}^{n-1}) \to (Q, Q^{n-1})$).*

Proof. Indeed, let assertion (B_n) be true and let f be an arbitrary map of the ball $\mathbb{E}^n$ into some cellular decomposition Q that takes the sphere $\mathbb{S}^{n-1}$ to the $(n-1)$-th skeleton Q^{n-1}. Since the ball $\mathbb{E}^n$ is compact, its image $f(\mathbb{E}^n)$

under the map f is also compact and therefore is contained in some finite subdecomposition of the decomposition Q. Therefore, without loss of generality, we can assume from the outset that Q is finite.

Using the finiteness of the decomposition Q, we will prove the statement (A_n) for it by induction on its dimension $m = \dim Q$ and on the number $a_m(Q)$ of its m-dimensional cells. If $m = n$, then the assertion (A_n) is trivially true. Assuming that this assertion has already been proven for all finite decompositions of Q for which either $\dim Q < m$, where $m > n$, or the number $a_m(Q)$ is less than some positive integer k, consider an arbitrary decomposition of Q for which $\dim Q = m$ and $a_m(Q) = k$. By choosing an arbitrary m-dimensional cell $e^m \in Q$ and setting $Q_0 = Q \setminus e^m$, we obviously obtain an m-dimensional relative cell (Q, Q_0). Since any map $(\mathbb{E}^n, \mathbb{S}^{n-1}) \to (Q, Q^{n-1})$ is automatically a map $(\mathbb{E}^n, \mathbb{S}^{n-1}) \to (Q, Q_0)$, then, according to the statement (B_n), each such map is homotopic rel $\mathbb{S}^{n-1})$ to some map $g : \mathbb{E}^n \to Q_0$.

On the other hand, it is clear that the subspace Q_0 is a cellular decomposition such that either $\dim Q_0 < m$ or $a_m(Q_0) < k$. Therefore, by the induction hypothesis, the map g is homotopic with rel $\mathbb{S}^{n-1})$ to some map $h : \mathbb{E}^n \to Q_0^n$. Consequently, the map f is also homotopic rel $\mathbb{S}^{n-1})$ to the map h. Thus, assertion (A_n) is completely proved. □

Thus, to prove the theorem on cellular maps, we only need to prove the statement (B_n) for any $n \geq 0$.

Remark 3.77. Like the statement (A_n), the statement (B_n) is trivially true when $n = 0$ (since the set $\overline{e}^m$ is path-connected).

Now consider the following statement:

Proposition 3.78 (C_n)**.** *For any $m > n$ and any m-dimensional relative cell (X, X_0), every map $f : (\mathbb{E}^n, \mathbb{S}^{n-1}) \to (X, X_0)$ is homotopic* rel $\mathbb{S}^{n-1}$ *to some map $g : \mathbb{E}^n \to X$ with the property that the image $g(\mathbb{E}^n)$ of the ball $\mathbb{E}^n$ under g does not contain the entire cell $e^m = X \setminus X_0$.*

It is easy to see that

Proposition 3.79. *statement (C_n) implies statement (B_n).*

Proof. Indeed, let the statement (C_n) be true and let x_0 be an arbitrary point of the cell e^m that does not belong to the set $g(\mathbb{E}^n)$. Let us choose a map

$$\chi : \mathbb{E}^m \to \overline{e}^m,$$

characteristic of the cell e^m such that $\chi(\mathbf{0}) = x_0$ (it is clear that such a map can always be found). It is easy to see that, by setting

$$g(x) = \chi : ((v(x), \boldsymbol{u}(x)), \quad x \in G$$

(where g is the map $\mathbb{E}^n \to X$ specified by the statement (C_n)), we uniquely define two continuous maps on the open subset $G = g^{-1}(e^m)$ of the ball $\mathbb{E}^n$:

$$v : G \to I, \quad \boldsymbol{u} : G \to \mathbb{S}^{m-1}$$

(since, by hypothesis, $\chi(\mathbf{0}) \notin g(\mathbb{E}^n)$, then $v(x) \neq 0$ for each point $x \in G$ and therefore the map $\boldsymbol{u} : G \to \mathbb{S}^{m-1}$ is uniquely defined). For any point $x \in \mathbb{E}^n$ and any number $t \in I$, we now set

$$g_t(x) = \begin{cases} \chi(((1-t)v(x) + t)\boldsymbol{u}(x)), & \text{if } x \in G, \\ g(x), & \text{if } x \notin G \end{cases}$$

It is clear that we thereby obtain a homotopy $g_t : \mathbb{E}^n \to X$ rel $\mathbb{S}^{n-1}$, for which $g_0 =$ and $g_1(\mathbb{E}^n) \subset X_0$. Thus, the map g, and hence the map f, is homotopic rel $\mathbb{S}^{n-1}$ to some map $\mathbb{E}^n \to X_0$. Assertion (B_n) is thus completely proved. □

Thus, all that remains for us to prove is statement (C_n) for all $n > 0$.
To this end, we consider the following statement:

Proposition 3.80 (D_n). *For any $m > n$, any n-dimensional relative cell (X, X_0), and any map $f : (\mathbb{E}^n, \mathbb{S}^{n-1}) \to (X, X_0)$, there exists a map*

$$g : (\mathbb{E}^n, \mathbb{S}^{n-1}) \to (X, X_0),$$

such that

1) *the image $g(\mathbb{E}^n)$ of the ball $\mathbb{E}^n$ under the map g does not contain the entire cell $e^n = X \setminus X_0$;*

2) *the map g coincides with the map f outside some open set $U \subset \mathbb{E}^n$;*

3) *the image of the set U under each of the maps f and g is contained in the cell e^m.*

It is easy to see that

Proposition 3.81. *the map g provided by statement (D_n) is homotopic* rel $\mathbb{S}^{n-1}$ *to the map f.*

Proof. Indeed, let

$$\chi : \mathbb{E}^m \to X$$

be an arbitrary map characteristic of the cell e^m. Since, by hypothesis, $f(U) \subset e^m$ and $g(U) \subset e^m$, then single-valued continuous maps are defined on the set U

$$f^{'} \chi^{-1} \circ f : U \to \mathbb{E}^m, \quad g^{'} \chi^{-1} \circ g : U \to \mathbb{E}^m$$

For any point $\boldsymbol{v} \in \mathbb{E}^n$ and any number $t \in I$ we put

$$g_t(\boldsymbol{v}) = \begin{cases} \chi(tf^{'}(\boldsymbol{v}) + (1-t)g^{'}(\boldsymbol{v})), & \text{if } \boldsymbol{v} \in U, \\ f(\boldsymbol{v}), & \text{if } \boldsymbol{v} \notin U. \end{cases}$$

Clearly, we thereby obtain a homotopy $g_t : \mathbb{E}^n \to X$ rel $\mathbb{S}^{n-1}$, connecting the map f with the map g. □

By condition 1) of assertion (D_n), the proven proposition means that

Proposition 3.82. *the truth of assertion (D_n) implies the truth of assertion (C_n).*

Thus, to prove the theorem on cellular maps, we only need to prove statement (D_n) for any $n > 0$. We will carry out this proof by induction on the number n. First, we will consider the case $n = 1$.

Proof. (the case $n = 1$) Let x_0 be an arbitrary interior point of the cell e^m and let F be its complete preimage $f^{-1}(x_0)$ under the map $f : \mathbb{E}^1 \to X$. The set F is a closed subset of the segment $\mathbb{E}^1 = [-1, 1]$ that does not contain its endpoints. Since the map f is continuous and the cell e^m is open in the space X, on the segment $\mathbb{E}^1$ there exists a finite system of open intervals (a_i, b_i), $i = 1, \dots, k$ such that their union U contains the set F and goes under the map f into the cell e^m. Since the set $e^m \setminus x_0$ is linearly connected (since $m > 1$), then for any $i = 1, \dots, k$ there exists a map g_i of the closed segment $[a_i, b_i]$ into the set $e^m \setminus x_0$ such that $g_i(a_i) = f(a_i)$ and $g_i(b_i) = f(b_i)$. But then it is clear that the map $g : \mathbb{E}^1 \to X$ defined by the formula

$$g(\boldsymbol{v}) = \begin{cases} g_i(\boldsymbol{v}), & \text{if} \quad \boldsymbol{v} \in (a_i, b_i), \\ f(\boldsymbol{v}), & \text{if} \quad \boldsymbol{v} \notin U, \end{cases}$$

is continuous and satisfies all the conditions of assertion (D_1). Therefore, assertion (D_1) is completely proved. □

Now suppose that for some $n > 1$ the assertions $(D_1, \dots, D_{n-1})$ have already been proved, and we prove the assertion (D_n).

To this end, we first note that

Proposition 3.83. *the validity of the assertions $(D_1, \dots, D_{n-1})$ implies that for any $m > n$ the sphere $\mathbb{S}^m$ is an $n-1$-connected space.*

Since the statement (D_k), $k = 1, \dots, n-1$, implies the statement (B_k), to prove this proposition it suffices to prove that

Proposition 3.84. *from the validity of the statement (B_k) it follows that for any $m > k$ the sphere $\mathbb{S}^m$ is a k-aspherical space.*

Proof. Let g be an arbitrary map $\mathbb{S}^k \to \mathbb{S}^m$. Having chosen some map

$$\chi : \mathbb{E}^k \to \mathbb{S}^k,$$

that maps the sphere $\mathbb{S}^{k-1} \subset \mathbb{E}^k$ to the point $\boldsymbol{u}_0 = (1, 0, \dots, 0) \in \mathbb{S}^k$ and homeomorphically maps the open ball $\dot{\mathbb{E}}^k = \mathbb{E}^k \setminus \mathbb{S}^{k-1}$ to the set $\mathbb{S}^k \setminus \boldsymbol{u}_0$ (i.e., a characteristic ma of the cell $\mathbb{S}^k \setminus \boldsymbol{u}_0$), we consider the map

$$f = g \circ \chi : (\mathbb{E}^k, \mathbb{S}^{k-1}) \to (\mathbb{S}^m, \boldsymbol{w}_0),$$

where $\boldsymbol{w}_0 = g(\boldsymbol{u}_0) \in \mathbb{S}^m$. Since the pair $(\mathbb{S}^m, \boldsymbol{w}_0)$ is obviously an m-dimensional relative cell and $m > k$, assertion (B_k) applies to this map. Therefore, it is homotopic rel $\mathbb{S}^{k-1}$ to the map of the ball $\mathbb{E}^k$ to the point $\boldsymbol{w}_0$. Let

$$f_t : \mathbb{E}^k \to \mathbb{S}^m$$

be the corresponding homotopy. Clearly, for any $t \in I$, the map

$$g_t = f_t \circ \chi^{-1} : \mathbb{S}^k \to \mathbb{S}^m$$

is uniquely defined and that all these maps form a homotopy connecting the map g with the constant map $\mathbb{S}^k \to \boldsymbol{w}_0$. Therefore, the sphere $\mathbb{S}^m$ is k-aspherical. □

Using the statements proved in §2.1, we immediately obtain from this that

Proposition 3.85. *from the validity of the statements (D_1), …, (D_n) it follows that for any $m > n$ the product $\mathbb{S}^{m-1} \times (0,1)$ is an $n-1$-connected space.*

After these preliminary remarks, we can now proceed directly to the proof of the assertion (D_n) (assuming that the assertions (D_1), …, (D_{n-1}) have already been proven).

Proof. (of the assertion (D_n)) Let $m > n$ and let

$$f : (\mathbb{E}^n, \mathbb{S}^{n-1}) \to (X, X_0)$$

be an arbitrary map of the pair $(\mathbb{E}^n, \mathbb{S}^{n-1})$ into some m-dimensional relative cell (X, X_0). Let, in addition, x_0 be an arbitrary interior point of the cell $e^m = X \setminus X_0$ and let

$$F = f^{-1}(x_0)$$

be its complete preimage under the map f. Since, by hypothesis, the boundary $\mathbb{S}^{n-1}$ of the ball $\mathbb{E}^n$ is mapped by f into the subspace X_0, the set F (obviously closed) is contained in the open ball $\dot{\mathbb{E}}^n = \mathbb{E}^n \setminus \mathbb{S}^{n-1}$. By similar considerations, the set $G = f^{-1}(e^m)$ (obviously open) is also contained in the ball $\dot{\mathbb{E}}^n$. Since $F \subset G$, there exists a cubillage of the Euclidean space $\mathbb{R}^n$ (i.e., a partition of the space $\mathbb{R}^n$ into cubes by orthogonal systems of parallel hyperplanes) so small that the set F has a neighbourhood U contained in the open set G, which is the union of some open cubes of this cubillage. Let P be the closure of the set U and Q be the boundary of this closure, i.e., its intersection with the union of all closed cubes of the cubillage under consideration that do not belong to it. It is clear that the set P is a cellular decomposition of dimension n (whose cells are the open cubes of the cubillage under consideration and their open faces), and the subset Q is a cellular subdivision of it. On the other hand, it is clear that $F \cap Q = \varnothing$ and therefore

$$f(Q) \subset e^m \setminus x_0,$$

so that we can view the map $f|_Q$ as the map

$$f|_Q : Q \to e^m \setminus x_0,$$

Finally, it is obvious that the set $e^m \setminus x_0$ is homeomorphic to the product $\mathbb{S}^{n-1} \times (0,1)$ and therefore, by the induction hypothesis and the remarks made above, it is $n-1$-connected.

Thus, the map $f|_Q$ is a continuous map of the sub-decomposition Q of P into the $n-1$-connected space $e^m \setminus x_0$, where $\dim(P \setminus Q) \leq n$. Therefore, by the theorem proved in the previous section, there exists some extension

$$h : P \to e^m \setminus x_0.$$

for this map. We define the map

$$g : \mathbb{E}^n \to X$$

by putting

$$g(\boldsymbol{v}) = \begin{cases} h(\boldsymbol{v}), & \text{if} \quad \boldsymbol{v} \in P, \\ f(\boldsymbol{v}), & \text{if} \quad \boldsymbol{v} \notin P \end{cases}$$

for any point $\boldsymbol{v} \in \mathbb{E}^n$. Clearly, this map is continuous and satisfies all the conditions of assertion (D_n). □

Thus, the assertions (D_n) are proved for all $n > 0$. Along with them, the theorem on cellular maps is also completely proven.

Remark 3.86. Incidentally, we have proven that for any $n > 0$, the sphere $\mathbb{S}^n$ is an $(n-1)$-connected space (see the end of §3.4).

3.6 Whitehead's theorem. Quasi-polyhedra

Let us return to the proposition proved at the beginning of the previous section. Setting in this proposition $Q = K$ and $Q_0 = Q_1 = \cdots = Q_n = \cdots = K_0$, where K_0 is some sub-decomposition of the cellular decomposition K, we immediately obtain that

Proposition 3.87. *if for every $n \geq 0$ any map $(\mathbb{E}^n, \mathbb{S}^{n-1}) \to (K, K_0)$ is homotopic* rel $S\mathbb{S}^{n-1}$ *to some map $\mathbb{E}^n \to K_0$, then every map*

$$f : (K, K_0) \to (K, K_0)$$

is homotopic rel K_0 *to some map*

$$g : K \to K_0.$$

According to the assertion proved at the end of §2.4, the conditions of this proposition are satisfied if the sub-decomposition K_0 is a representative subspace of the space K. On the other hand, in the case where the map f is the identity map 1_K of the decomposition K, the map g provided by this proposition is obviously a retraction $K \supset K_0$. Thus,

Proposition 3.88. *any sub-decomposition K_0 of the decomposition K, which is its representative subspace, is a deformation retract of the space K.*

In other words,

Proposition 3.89. *for sub-decompositions of cellular decompositions, the property of being a representative subspace is equivalent to the property of being a deformation retract.*

It easily follows from this proposition that

Proposition 3.90. *for any cellular decompositions K and L, every weak homotopy equivalence*

$$f : K \to L$$

is a homotopy equivalence.

In other words,

Proposition 3.91. *For cell decomposition maps, the property of "being a weak homotopy equivalence" is equivalent to the property of "being a homotopy equivalence".*

Proof. Indeed, according to the theorem on cellular maps, we can assume without loss of generality that the map f is cellular, and therefore its cylinder Z_f is a cellular decomposition. On the other hand, the fact that the map f is a weak homotopy equivalence means, as we know, that the subspace K of the cylinder Z_f is representative. Therefore, since this subspace is clearly a sub-decomposition of the cylinder Z_f, it is a deformation retract of it, and therefore the map f is a homotopy equivalence. □

The proved proposition is known as *Whitehead's theorem*. It is one of the fundamental tools for studying the homotopy properties of cellular decompositions. For example, this theorem almost immediately implies that

Proposition 3.92. *any cellular decomposition K is the homotopy limit of every increasing sequence*

$$K_0 \subset K_1 \subset \cdots \subset K_n \subset \cdots$$

of its sub-decomposition, the union of which coincides with the entire decomposition K.

Proof. Indeed, the space K^{Σ} corresponding to the sequence under consideration is, obviously, a cellular decomposition, and the natural map

$$p^{\Sigma} : K^{\Sigma} \to K$$

is weak homotopy equivalence. Therefore, according to Whitehead's theorem, this map is a homotopy equivalence. □

We will call a topological space X a *quasi-polyhedron* if it is homotopically equivalent to some cellular decomposition. Clearly, Whitehead's theorem remains valid for quasi-polyhedra as well, i.e.,

Proposition 3.93. *any continuous map*

$$f : X \to Y$$

of a quasipolyhedron X into a quasipolyhedron Y that is a weak homotopy equivalence is a homotopy equivalence.

Moreover, from what was said at the end of §3.4 it immediately follows that

Proposition 3.94. *for any cell pair (K, L) and any continuous map $f : L \to Y$ of a sub-decomposition of L into an arbitrary quasi-polyhedron Y, the space $K \cup_f Y$ is a quasi-polyhedron.*

Remark 3.95. Quasi-polyhedra constitute a remarkable class of topological spaces, distinguished by the property that practically any "reasonable" operations on topological spaces do not lead outside this class. For example, for any quasi-polyhedra X and Y, the space Y^X of all continuous maps $X \to Y$, considered in the so-called "compact-open topology" (see §9.1), is also a quasi-polyhedron, and every topological space X for which there exists a homotopically injective map $A \to K$ into some quasi-polyhedron K is itself necessarily a quasi-polyhedron. On the other hand, in all homotopy questions one can restrict oneself only to quasi-polyhedra, since for any space X there exists a continuous map from some quasi-polyhedron Z to the space X that is a weak homotopy equivalence. These properties of quasi-polyhedra will not be needed by us, and we will leave them without proof.

Chapter 4

Smooth Manifolds. I

This chapter is mainly devoted to the construction of tensor calculus on arbitrary smooth manifolds.

In the preparatory §4.1, we introduce the concept of a smooth premanifold as a Hausdorff topological space on which a certain sheaf of germs of real-valued functions is defined (to use the currently fashionable terminology).

In §4.2, which also has a preparatory character, we prove the classical theorem on differentiable maps with non-zero Jacobian for Euclidean spaces; in doing so, we specifically emphasise some details that are essential for what follows, related to estimating the diameter of the domain in which the map is diffeomorphic, which are usually left without consideration.

In §4.3, smooth manifolds are defined as smooth premanifolds that are locally diffeomorphic to Euclidean spaces. Here, the concept of a product of smooth manifolds is also introduced.

In §4.4, we prove analogues of Urysohn's lemma and Tietze's theorem for smooth functions. Here, we also prove that any convex open subset of Euclidean space is diffeomorphic to an open ball.

In §4.5, we introduce vector fields as derivations of the algebra of smooth functions and show that in any coordinate neighbourhood, each vector field is a linear combination of partial derivations with respect to local coordinates.

In §4.6, we introduce the concept of a vector at a point, define the manifold M_* of all vectors at all possible points of a given manifold M, and show that the vector fields introduced in §4.5 can be interpreted as smooth maps $M \to M_*$. In this section, we also introduce the concept of a differential of a smooth map and, in connection with this, the concept of a regular map.

In §4.7, linear differential forms are considered in a similar way, and in §4.8, tensor fields of arbitrary type.

In §4.9, an algebra of tensors and tensor fields is constructed. In particular, the operation of convolution is considered in detail here.

In the final §4.10, the concept of a Riemannian space is defined and, based on the results of §4.4, it is proved that on any smooth separable (i.e., with a countable base) manifold there exists a metric tensor field, i.e., that each such manifold can be defined as a Riemannian space.

4.1 Smooth premanifolds

Let M be an arbitrary set and let $f, f^1, \dots, f^r$ be some real functions defined on the space M. We will say that a function f *depends smoothly* on functions $f^1, \dots, f^r$ if there exists an infinitely differentiable function $u(t^1, \dots, t^r)$ of real variables $t^1, \dots, t^r$, defined (and infinitely differentiable) for all values of these variables, such that $f = u(f^1, \dots, f^r)$ on M, i.e. such that

$$f(p) = u(f^1(p), \dots, f^r(p)) \tag{4.1}$$

for any point $p \in M$. If equality (4.1) holds only for points p of some set $U \subset M$, then we will say that the function f depends smoothly on the functions $f^1, \dots, f^r$ on the set U.

Remark 4.2. Some authors require that the function u, which establishes the smooth dependence (4.1), be defined (and infinitely differentiable) only in some open set of the arithmetic space $\mathbb{R}^r$, containing all points of the form $(f^1(p), \dots, f^r(p))$, $p \in M$. It is easy to show that this (formally more general) definition essentially coincides with our definition. On the other hand, it is often required not that the function u be infinitely differentiable, but only that it have a finite number of derivatives (up to some fixed order N). It can easily be verified that all the theory developed below remains valid with this definition of smooth dependence, provided that the number N is sufficiently large.

Now let M be a topological space. We will say that a function f defined on M depends smoothly on the functions $f^1, \dots, f^r$ *near* a point $p_0 \in M$ if equality (4.1) holds for all points p in some neighbourhood of the point p_0. Similarly, we will say that two functions *coincide near a point* $p_0 \in M$ if they take the same values in some neighbourhood of this point.

Definition 4.3. By *smoothness* on a topological space M we will define a non-empty set $\mathcal{O}$ of real functions defined on the space M such that

1) any function that depends smoothly on functions from $\mathcal{O}$ itself belongs to $\mathcal{O}$;

2) any function that coincides near each point $p \in M$ with some function from $\mathcal{O}$ (generally speaking, depending on the point p) itself belongs to $\mathcal{O}$.

Condition 1) implies that all constant functions belong to $\mathcal{O}$ and that the sum and product of any functions from $\mathcal{O}$ also belong to $\mathcal{O}$. Consequently,

Proposition 4.4. *every smoothness $\mathcal{O}$ is an algebra over the field $\mathbb{R}$ of real numbers.*

This algebra is commutative, associative, and has a unit.

A Hausdorff topological space M on which some smoothness $\mathcal{O} = \mathcal{O}(M)$ is defined will be called a *smooth premanifold*, and functions $f \in \mathcal{O}$ will be called *smooth functions on M*.

Note that smoothness is very weakly related to topology (only through condition 2). In particular, smooth functions, generally speaking, may not be continuous. A close interaction between smoothness and topology occurs only for manifolds (see below, section 4.3).

We obtain an important example of a smooth premanifold by considering an arbitrary n-dimensional linear space L. By choosing a basis in this space, we can consider every function on L as a function of m real variables - the coordinates of vectors from L in this basis. Clearly, the set $\mathcal{O}(L)$ of all functions on L that are infinitely differentiable functions of coordinates is independent of the choice of basis and satisfies conditions 1 and 2) of the definition of smoothness. Furthermore, when speaking of the linear space L as a smooth premanifold, we will always imply that the functions in $\mathcal{O}(L)$ are considered smooth on L.

In particular, the arithmetic space $\mathbb{R}^m$ is a smooth premanifold. Smooth functions on $\mathbb{R}^m$ are simply infinitely differentiable functions of the coordinates $t^1, \ldots, t^m$ of its points.

Let N be an arbitrary subspace of some smooth premanifold M. A real function f, defined on N, will be called *smooth* (*on* N) if, near any point $\in N$, it coincides with some smooth function on M, i.e., if for any point $p \in N$ there exists in M a neighbourhood U and a smooth function g on M such that

$$f|_{U \cap N} = g|_{U \cap N}.$$

It is easy to see that

Proposition 4.5. *the set $\mathcal{O}(N)$ of all smooth functions on N is a smoothness (on N).*

Proof. Indeed, let a function f on N depend smoothly on smooth functions $f^1, \ldots, f^r$ on N, i.e., let

$$f = u(f^1, \ldots, f^r) \quad \text{on} \quad N,$$

where u is some infinitely differentiable function of r real variables. By definition, for any point $p \in N$ and any $i = 1, \ldots, r$ there exists in M a neighbourhood U_i of the point p and a smooth function g^i on M such that

$$f^i|_{U_i \cap N} = g^i|_{U_i \cap N}, \quad i = 1, \ldots, r.$$

Let $U = \cap_{i=1}^r U_i$ and let $g = u(g^1, \ldots, g^r)$. Then $g \in \mathcal{O}(M)$ and

$$f|_{U \cap N} = g|_{U \cap N}.$$

Thus, for any point $p \in N$, there exists a neighbourhood U in M and a function g smooth on M such that $f = g$ on $U \cap N$. Consequently, the function f is smooth on N. Thus, for the set $\mathcal{O}(N)$, condition 1) of the definition of smoothness is completely verified.

Let us now check condition 2). Let a function f defined on N coincide near any point $p \in N$ with some smooth function on N, i.e., let for any point $p \in N$

there exist in N a neighbourhood V of p and a smooth function h on N such that $h = f$ on V. By definition, in M there exist a neighbourhood W of p and a smooth function g on M such that $h = g$ on $N \cap W$. Let U be a neighbourhood of p in M such that $U \cap N \subset W \cap V$. Then $f = g$ on $N \cap U$. Thus, near any point $p \in N$ the function f coincides with some smooth function g on M. Therefore, the function F is a smooth function on N. Condition 2) is thus also verified. □

The subspace $N \subset M$ considered together with the smoothness $\mathcal{O}(N)$ will be called a *sub-premanifold* of the premanifold M. According to the proved assertion, any sub-premanifold is a smooth premanifold.

A function f that is smooth in some neighbourhood of a point $p \in M$ (and defined in this or a larger neighbourhood) will be called a function *smooth at* p. Clearly, if we agree to consider two functions smooth at a point p to be the same when they coincide near this point, then the set $\mathcal{O}(p) = \mathcal{O}_M(p)$ of all such functions will also be an algebra over the field $\mathbb{R}$.

Remark 4.6. In the terminology of sheaf theory, the algebra $\mathcal{O}(p)$ is nothing more than the algebra of germs of smooth functions in p, and the algebra $\mathcal{O}(M)$ is the algebra of sections of the sheaf of these germs. We will not use this terminology, since we essentially do not need to employ the apparatus of sheaf theory.

Let N and M be arbitrary smooth premanifolds. A continuous map

$$\Phi : N \to M$$

from a smooth premanifold N to a smooth premanifold M will be called *smooth at a point* $P \in N$ if, for any function f that is smooth at a point $\Phi(p) \in M$, the composite function $f \circ \Phi$ is smooth at p. When a map Φ is smooth at all points $p \in N$, we say that this map is *smooth on* N. An example of a smooth map is the inclusoin $i : N \subset M$ of an arbitrary sub premanifold N of the premanifold M. Another example is an arbitrary smooth function f on M, considered as a map of the premanifold M onto the real axis $\mathbb{R}$.

A homeomorphic map

$$\Phi : N \to M$$

of a premanifold N onto a premanifold M will be called a *diffeomorphism* (or *diffeomorphic map*) if it is smooth on N, and its inverse map $\Phi^{-1} : M \to N$ is smooth on M. Premanifolds N and M for which at least one diffeomorphism $N \to M$ exists are called *diffeomorphic*. As a rule, we will consider diffeomorphic premanifolds to be the same.

We will call a map $\Phi : N \to M$ *diffeomorphic at a point* $p \in N$ if it is a diffeomorphic map of some neighbourhood of p onto some neighbourhood of $\Phi(p)$. A map that is diffeomorphic at all points $p \in N$ is called *locally diffeomorphic*. Clearly,

Proposition 4.7. *a map* $\Phi : N \to M$ *is diffeomorphic if and only if it is bijective and locally diffeomorphic.*

4.2 Inverse function theorem

In the special case where the premanifolds N and M are open subsets U and V of the space $\mathbb{R}^m$, each smooth map

$$\Phi : U \to V$$

is defined by a system of m smooth (i.e., infinitely differentiable) functions

$$u^1(t^1, \ldots, t^m), \ldots u^m(t^1, \ldots, t^m)$$

expressing the coordinates $u^1, \ldots, u^m$ of the point $\varphi(\boldsymbol{t}) \in V$ in terms of the coordinates $t^1, \ldots, t^m$ of the point $\boldsymbol{t} \in U$. The Jacobian

$$\frac{D(u^1, \ldots, u^m)}{D(t^1, \ldots, t^m)} = \det \left\| \frac{\partial u^i}{\partial t^j} \right\|_{i,j=1, \ldots, m}$$

of this system of functions we will call the *Jacobian of the map* φ and will denote it by the symbol D_φ. This Jacobian is a smooth function on an open set U. From the rule for differentiating composite functions, it follows directly that the Jacobian $D_{\psi\circ\varphi}$ of the composition $\psi \circ \varphi : U \to W$ of two smooth maps $\varphi : U \to V$ and $\psi : V \to W$, where U, V, and W are open subsets of the space $\mathbb{R}^m$, is expressed in terms of the Jacobians D_ψ and D_φ of the maps ψ and φ according to the formula

$$D_{\psi\circ\varphi} = (D_\psi \circ \varphi) \cdot D_\varphi. \tag{4.8}$$

In particular, if the map φ is diffeomorphic, then

$$D_{\varphi^{-1}\circ\varphi} = D_\varphi^{-1}.$$

Consequently,

Proposition 4.9. *the Jacobian D_φ of the diffeomorphic map*

$$\varphi : U \to V$$

does not vanish at any point of the set U.

Furthermore, since for any open set $U_0 \subset U$ the Jacobian D_ι of the inclusin $\iota : U_0 \subset U$ is obviously equal to one, we, by applying relation (4.8) to the map $\varphi \circ \iota$, immediately obtain that

Proposition 4.10. *the Jacobian D_{φ_0} of the restriction $\varphi_0 = \varphi|_{U_0}$ of an arbitrary smooth map $\varphi : U \to V$ on the set $U_0 \subset U$ is the restriction $D_\varphi|_{U_0}$ of the Jacobian D_φ of the map φ:*

$$D_{\varphi_0} = D_\varphi|_{U_0}.$$

Therefore,

Proposition 4.11. *if the map $\varphi : (U \to V)$ is diffeomorphic at a point $t_0 \in U$, then its Jacobian D_φ is non-zero at this point.*

In particular,

Proposition 4.12. *the Jacobian D_φ of an arbitrary locally diffeomorphic map*

$$\varphi : U \to V$$

is nonzero everywhere on the set U.

It turns out that the converse is also true, i.e.,

Proposition 4.13. *if the Jacobian D_φ of a smooth map*

$$\varphi : U \to V$$

is nonzero at a point $\boldsymbol{t}_0 \in U$, then the map φ is diffeomorphic at this point.

This statement is known as the *inverse function theorem.* Although its proof can be found in any sufficiently comprehensive course on analysis, we will present it here because we will need some aspects of this theorem that are usually left aside.

It is clear that in proving the theorem on inverse functions, we can, without loss of generality, assume that $\boldsymbol{t}_0 = \mathbf{0}$, $\varphi(\boldsymbol{t}_0) = \mathbf{0}$, and that the set U is a ball $|\boldsymbol{t}_0| < R$ of some radius $R > 0$ with centre at the point $\mathbf{0}$. Since the map φ is smooth by assumption, for any two points $\boldsymbol{t}$, $\boldsymbol{t} + \boldsymbol{h} \in U$ the equality holds:

$$\varphi(\boldsymbol{t} + \boldsymbol{h}) - \varphi(\boldsymbol{t}) = \varphi^{'}(\boldsymbol{t})\boldsymbol{h} + |\boldsymbol{h}|\boldsymbol{\varepsilon}(\boldsymbol{h}, \boldsymbol{t}),$$

where $\varphi^{'}(\boldsymbol{t})$ is a linear operator whose matrix is the Jacobian matrix of the map φ, and $\boldsymbol{\varepsilon}(\boldsymbol{h}, \boldsymbol{t})$ is a vector whose length $|\boldsymbol{\varepsilon}(\boldsymbol{h}, \boldsymbol{t})|$ tends to zero as $|\boldsymbol{h}| \to 0$. The operator $\varphi^{'}(\boldsymbol{t})$ depends smoothly on the point $\boldsymbol{t} = (t^1, \ldots, t^m)$ in the sense that the elements $\frac{\partial u^i}{t^f}$ of its matrix are infinitely differentiable functions of the variables $t^1, \ldots, t^m$. Therefore, this operator satisfies the Lipschitz condition at the point $\mathbf{0}$, i.e., there exists a constant number $M > 0$ such that for any points]$\boldsymbol{t} \in U$, $\boldsymbol{h} \in \mathbb{R}^m$ the following equality holds:

$$|\varphi^{'}(\boldsymbol{t})\boldsymbol{h} - \varphi^{'}(\mathbf{0})\boldsymbol{h}| \le M \cdot |\boldsymbol{t}| \cdot |\boldsymbol{h}|.$$

Let

$$m = \inf_{|\boldsymbol{h}|=1} |\varphi^{'}(\mathbf{0})\boldsymbol{h}|.$$

Since, by hypothesis,

$$D_\varphi(\mathbf{0}) = \det \varphi^{'}(\mathbf{0}) \neq 0,$$

then $m > 0$ and therefore $r_0 = \min\left(R, \frac{m}{2M}\right) > 0$.

It is easy to see that for $|\boldsymbol{t}| < r_0$ the operator $\varphi^{'}(\boldsymbol{t})$ is invertible. Indeed, if $\varphi^{'}(\boldsymbol{t})\boldsymbol{h} = 0$, then

$$|\varphi^{'}(\mathbf{0})\boldsymbol{h}| = |\varphi^{'}(\boldsymbol{t})\boldsymbol{h} - \varphi^{'}(\mathbf{0})\boldsymbol{h}| \le M \cdot |\boldsymbol{t}| \cdot |\boldsymbol{h}| < M r_0 \le m,$$

which is impossible for $\boldsymbol{h} \neq \mathbf{0}$.

Now let $|\boldsymbol{t}| < r_0$ and $|\boldsymbol{t}+\boldsymbol{h}| < r_0$. To each unit vector $\boldsymbol{e} \in \mathbb{R}^m$ we assign a numerical function $f(\lambda)$ of a real variable $\lambda \in I$, equal to the scalar product of the vectors $\varphi(\boldsymbol{t}+\lambda\boldsymbol{h})$ and $\boldsymbol{e}$:

$$f(\lambda) = (\varphi(\boldsymbol{t}+\lambda\boldsymbol{h}), \boldsymbol{e}).$$

Since

$$f^{'}(\lambda) = (\varphi^{'}(\boldsymbol{t}+\lambda\boldsymbol{h}), \boldsymbol{e}),$$

then, applying Lagrange's theorem to the function $f(\lambda)$, we obtain that

$$\varphi(\boldsymbol{t}+\boldsymbol{h}, \boldsymbol{e}) - (\varphi(\boldsymbol{t}, \boldsymbol{e}) = (\varphi^{'}(\boldsymbol{t}+\lambda_0\boldsymbol{h}), \boldsymbol{e}),$$

where λ_0 is some number of the segment I (depending on the vector $\boldsymbol{e}$). Hence,

$$\begin{aligned} |(\varphi(\boldsymbol{t}+\boldsymbol{h}) - \varphi(\boldsymbol{t})| &\geq |\varphi(\boldsymbol{t}+\boldsymbol{h}) - \varphi(\boldsymbol{t}), \boldsymbol{e})| = |\varphi^{'}(\boldsymbol{t}+\lambda_0\boldsymbol{h})\boldsymbol{h}, \boldsymbol{e})| \\ &\geq |(\varphi^{'}(\mathbf{0})\boldsymbol{h}, \boldsymbol{e})| - |(\varphi^{'}(\boldsymbol{t}+\lambda_0\boldsymbol{h})\boldsymbol{h} - \varphi^{'}(\mathbf{0})\boldsymbol{h}, \boldsymbol{e})|. \end{aligned}$$

On the other hand,

$$\begin{aligned} |(\varphi^{'}(\boldsymbol{t}+\lambda_0\boldsymbol{h})\boldsymbol{h} - \varphi^{'}(\mathbf{0})\boldsymbol{h}, \boldsymbol{e})| &\leq |\varphi^{'}(\boldsymbol{t}+\lambda_0\boldsymbol{h})\boldsymbol{h} - \varphi^{'}(\mathbf{0})\boldsymbol{h}| \\ &\leq M \cdot |\boldsymbol{t}+\lambda_0\boldsymbol{h}| \cdot |\boldsymbol{h}| \leq M r_0 |\boldsymbol{h}| \leq \frac{m}{2}|\boldsymbol{h}|. \end{aligned}$$

Also, if $\boldsymbol{e} = \frac{\varphi^{'}(\mathbf{0})\boldsymbol{h}}{|\varphi^{'}(\mathbf{0})\boldsymbol{h}|}$, then

$$|\varphi^{'}(\mathbf{0})\boldsymbol{h}, \boldsymbol{e})| = |\varphi^{'}(\mathbf{0})\boldsymbol{h}| \geq m|\boldsymbol{h}|.$$

Hence

$$|\varphi(\boldsymbol{t}+\boldsymbol{h}) - \varphi(\boldsymbol{t})| \geq \frac{m}{2}|\boldsymbol{h}|. \tag{4.14}$$

Moreover, as is easy to see, the equality is possible here only if $\boldsymbol{h} = \mathbf{0}$. Therefore, if $\boldsymbol{h} \neq \mathbf{0}$, then

$$\varphi(\boldsymbol{t}+\boldsymbol{h}) \neq \varphi(\boldsymbol{t}).$$

This means that on the ball $U_0 : |\boldsymbol{t}| < r_0$ the map φ is injective, and therefore on the set

$$V_0 = \varphi(U_0)$$

the inverse map

$$\varphi^{-1} : V_0 \to U_0$$

is defined. By relation (4.14), the map φ^{-1} is continuous.

Let us show that the set V_0 is open in the space $\mathbb{R}_m$, i.e., that any of its points $\varphi(\boldsymbol{t}_0)$, $|\boldsymbol{t}_0| < r_0$ is its interior point. Let $|\boldsymbol{t}_0| < r < r_0$ and let

$$\delta = \frac{m}{4}(r - |\boldsymbol{t}_0|) > 0.$$

We will show that an open ball of radius δ with centre at the point $\varphi(\boldsymbol{t}_0)$ is entirely contained in the set V_0, i.e., that for any point $\boldsymbol{c} \notin V_0$ (if such a point exists) the inequality

$$|\varphi(\boldsymbol{t}_0) - \boldsymbol{c}| \geq \delta.$$

holds.

For this purpose, we consider the number

$$\eta = \inf_{|\boldsymbol{t}| \leq r} |\varphi(\boldsymbol{t}) - \boldsymbol{c}| > 0.$$

By definition, in the ball $|\boldsymbol{t}_0| \leq r$ there exists a sequence of points t_n such that

$$\lim_{n\to\infty} |\varphi(\boldsymbol{t}_n) - \boldsymbol{c}| = \eta.$$

Since the ball $|\boldsymbol{t}| \leq r$ is compact, we can assume (passing to a subsequence if necessary) that the sequence t_n converges to some point $\boldsymbol{a}$. It is easy to see that

$$|\boldsymbol{a}| = r.$$

Indeed, since $|\boldsymbol{a}| \leq r < r_0$, the operator $\varphi'(\boldsymbol{a})$ is invertible and therefore there exists a vector $\boldsymbol{h} \in \mathbb{R}^m$ such that

$$\varphi'(\boldsymbol{a})\boldsymbol{h} = \boldsymbol{c} - \varphi(\boldsymbol{a}).$$

Let $|\boldsymbol{a}| < r$ and let λ be a positive number so small that

$$|\boldsymbol{a} + \lambda\boldsymbol{h}| \leq r.$$

Since

$$\varphi(\boldsymbol{a} + \lambda\boldsymbol{h}) = \varphi(\boldsymbol{a}) + \lambda\varphi'(\boldsymbol{a})\boldsymbol{h} + \lambda\boldsymbol{\varepsilon}(\lambda),$$

where $\boldsymbol{\varepsilon}(\lambda) \to 0$ as $\lambda \to 0$, then

$$\varphi(\boldsymbol{a} + \lambda\boldsymbol{h}) - \boldsymbol{c} = (1 - \lambda)(\varphi(\boldsymbol{a}) - \boldsymbol{c}) + \lambda\boldsymbol{\varepsilon}(\lambda),$$

and therefore

$$|\varphi(\boldsymbol{a} + \lambda\boldsymbol{h}) - \boldsymbol{c}| \leq (1 - \lambda)|(\varphi(\boldsymbol{a}) - \boldsymbol{c})| + \lambda|\boldsymbol{\varepsilon}(\lambda)| \leq (1 - \lambda)\eta + \lambda|\boldsymbol{\varepsilon}(\lambda)|.$$

Therefore, if λ is small enough, then

$$|\varphi(\boldsymbol{a} + \lambda\boldsymbol{h}) - \boldsymbol{c}| \leq \eta.$$

This inequality contradicts the definition of the number η, and, consequently, the assumption $|\boldsymbol{a}| < r$ is false.

Let us now use inequality (4.14). According to this inequality,

$$|\varphi(\boldsymbol{t}_0) - \boldsymbol{c}| + |\varphi(\boldsymbol{t}_n) - \boldsymbol{c}| \geq |\varphi(\boldsymbol{t}_n) - \varphi(\boldsymbol{t}_0)| \geq \frac{m}{2}|\boldsymbol{t}_n - \boldsymbol{t}_0| \geq \frac{m}{2}(|\boldsymbol{t}_n| - |\boldsymbol{t}_0|).$$

Passing to the limit as $n \to \infty$ and taking into account that, as just proved, $|\boldsymbol{t}_n| \to |\boldsymbol{a}| = r$, we obtain from this that

$$|\varphi(\boldsymbol{t}_0) - \boldsymbol{c}| + \eta \geq \frac{m}{2}(r - |\boldsymbol{t}_0) = 2\delta.$$

Since $\eta \leq |\varphi(\boldsymbol{t}_0) - \boldsymbol{c}|$, this is possible only for

$$|\varphi(\boldsymbol{t}_0) - \boldsymbol{c}| \geq \delta.$$

To complete the proof of the theorem on inverse functions, it remains to prove that in the open set V_0 the map φ^{-1} is smooth. To this end, we first prove that it is differentiable, i.e., that for any point $\boldsymbol{t}_1 \in V_0$ there exists a linear operator $(\varphi^{-1})'(\boldsymbol{t}_1)$ such that for any two points $\boldsymbol{t}_1, \boldsymbol{t}_1 + \boldsymbol{h}_1 \in V_0$ the following equality holds:

$$\varphi^{-1}(\boldsymbol{t}_1 + \boldsymbol{h}_1) = (\varphi^{-1})'(\boldsymbol{t}_1)\boldsymbol{h}_1 + |\boldsymbol{h}_1|\boldsymbol{\eta}(\boldsymbol{h}_1, \boldsymbol{t}_1), \tag{4.15}$$

where $|\boldsymbol{\eta}(\boldsymbol{h}_1, \boldsymbol{t}_1)| \to 0$ as $|\boldsymbol{h}_1| \to 0$.

Let $\boldsymbol{t} = \varphi^{-1}(\boldsymbol{t}_1)$ and $\boldsymbol{t} + \boldsymbol{h} = \varphi^{-1}(\boldsymbol{t}_1 + \boldsymbol{h}_1)$. Then

$$\begin{aligned}\varphi^{-1}(\boldsymbol{t})^{-1}\boldsymbol{h}_1 &= \varphi^{-1}(\varphi(\boldsymbol{t} + \boldsymbol{h}) - \varphi(\boldsymbol{t})) = \boldsymbol{h} + |\boldsymbol{h}|\varphi'(\boldsymbol{t})^{-1}\boldsymbol{\varepsilon}(\boldsymbol{t} + \boldsymbol{h}) \\ &= \varphi^{-1}(\boldsymbol{t}_1 + \boldsymbol{h}_1) - \varphi^{-1}(\boldsymbol{t}_1) + |\boldsymbol{h}|\varphi'(\boldsymbol{t})^{-1}\boldsymbol{\varepsilon}(\boldsymbol{t} + \boldsymbol{h}).\end{aligned}$$

But this means that for the map φ^{-1} the equality (4.15) holds with

$$(\varphi^1(\boldsymbol{t}_1) = \varphi'(\boldsymbol{t})^{-1} \quad \text{and} \quad \boldsymbol{\eta}(\boldsymbol{h}_1, \boldsymbol{t}_1) = \frac{|\boldsymbol{h}|}{|\boldsymbol{h}_1|}\varphi'(\boldsymbol{t})^{-1}\boldsymbol{\varepsilon}(\boldsymbol{t} + \boldsymbol{h}).$$

Since $(\varphi^1(\boldsymbol{t}_1) = \varphi'(\boldsymbol{t})^{-1}$, it immediately follows from the elementary formulae for the matrix of the inverse operator that all elements of the matrix of the operator $(\varphi^1)'(\boldsymbol{t})$ are infinitely differentiable functions of the coordinates of the point $\boldsymbol{t}_1$. Consequently, the map φ^1 is smooth.

Thus, the inverse function theorem is completely proven.

Remark 4.16. In what follows, we will have to consider not a single smooth map φ, but a whole family φ_x of such maps. We will assume that

1) all these maps are defined on the same ball $|\boldsymbol{t}| \leq R$;
2) for any x, the Jacobian D_{φ_x} of the map φ_x at the point $\mathbf{0}$ is nonzero;
3) the parameter x ranges over some topological space X;
4) the maps φ_x depend continuously on the parameter x in the sense that the functions $u_x^i(t^1, \ldots, t^m)$ defining these maps depend continuously on x.

In this case, the Lipschitz constant M can obviously be chosen so that it depends continuously on the parameter x. Furthermore, it is clear that the number m also depends continuously on x. Therefore, the number r_0 will also depend continuously on x.

Thus,

Proposition 4.17. *in the case under consideration, the radii r_x of neighbourhoods of the point $\mathbf{0}$ in which the maps φ_x are diffeomorphisms can be chosen so that they depend continuously on the parameter x.*

4.3 Smooth manifolds

A smooth premanifold M is called an *m-dimensional smooth manifold* if each of its points p has a neighbourhood U diffeomorphic to some open subset of $\mathbb{R}^m$. For $m = 0$, each such premanifold is a discrete space, i.e., it consists of separate isolated points. Since this case is uninteresting, we will assume throughout that $m > 0$.

The neighbourhood U of the point p, provided by the definition of a smooth manifold, we will call the *coordinate neighbourhood* of this point, and its diffeomorphic map into the space $\mathbb{R}^m$ will be called the *coordinate diffeomorphism*. On each coordinate neighbourhood U, the coordinate diffeomorphism

$$\xi : U \to \mathbb{R}^m$$

defines, according to the formula

$$\xi(q) = (x^1(q), \ \ldots, \ x^m(q)), \quad q \in U, \tag{4.18}$$

m real functions $x^1 = x^1(q), \ \ldots, \ x^m = x^m(q)$, which we will call *local coordinates at the point p* corresponding to the given coordinate diffeomorphism ξ.

It is easy to see that

Proposition 4.19. *functions $x^1, \ \ldots, \ x^m$ defined in some neighbourhood U of a point p are local coordinates in the neighbourhood U if and only if*

1) the map

$$\xi : U \to \mathbb{R}^m$$

defined by formula (4.18) is a homeomorphism of the neighbourhood U onto some open set $U^0 \subset \mathbb{R}^m$;

2) the functions $x^1, \ \ldots, \ x^m$ are smooth on U;

3) each function f, smooth at a point $q \in U$, depends smoothly near this point on the functions $x^1, \ \ldots, \ x^m$.

Proof. Indeed, if the functions $x^1, \ \ldots, \ x^m$ are local coordinates in a neighbourhood U, then condition 1) is satisfied by definition, condition 2) is satisfied because $x^i = t^i \circ \xi$, where t^i are functions on $\mathbb{R}^m$ that associate with each point of $\mathbb{R}^m$ its i-th coordinate, and, finally, condition 3) is satisfied because, by the smoothness of the inverse map $\xi^{-1} : U^0 \to U$, the function $f \circ \xi^{-1}$ is smooth at the point $\xi(q) \in \mathbb{R}^m$, i.e., it coincides near this point with some smooth infinitely differentiable function on $\mathbb{R}^m$, and, consequently, in the corresponding neighbourhood of the point q the equality $f = u \circ \xi$ holds, i.e. the equality $f = u(x^1, \ \ldots, \ x^m)$.

Conversely, if the functions $x^1, \ \ldots, \ x^m$ satisfy conditions 1) - 3), then the map ξ will be a diffeomorphism, since condition 1) guarantees it is homeomorphic, condition 2) guarantees it is smooth (since $u \circ \xi = u(x^1, \ \ldots, \ x^m)$ for any function u on $\mathbb{R}^m$) and condition 3) guarantees the inverse map $\xi^{-1}S$ is also smooth (since the relation $f = u(x^1, \ \ldots, \ x^m)$ means that $u = f \circ \xi^1$). □

Since local coordinates $x^1, \ldots, x^m$ are obviously continuous, it follows directly from the proved assertion that

Proposition 4.20. *any smooth function on a smooth manifold is continuous on this manifold.*

It is clear that at any point p of a smooth manifold M, there exist infinitely many different systems of local coordinates. To obtain another system from a given system of local coordinates $x^1, \ldots, x^m$, defined, say, in a neighbourhood U, one can, for example, take an arbitrary neighbourhood V contained in U and consider the restrictions

$$y^1 = x^1|_V, \ \ldots \ y^m = x^m|_V$$

of the functions $x^1, \ldots, x^m$ to the neighbourhood V. It is clear that the functions $y^1, \ldots, y^m$ are local coordinates in a neighbourhood V. We can also take an arbitrary diffeomorphism φ of the open set $U^0 = \xi(U)$ (where ξ is the coordinate diffeomorphism defining the local coordinates $x^1, \ldots, x^m$) onto some other open set V_0 of the space $\mathbb{R}^m$ and consider the functions $y^1, \ldots, y^m$ defined by the diffeomorphism $\varphi \circ \xi$. These functions will be local coordinates in the same neighbourhood U as the coordinates $x^1, \ldots, x^m$. It is easy to see that

Proposition 4.21. *these two transformations are sufficient to relate two arbitrary systems of local coordinates at the point p.*

Proof. Indeed, if one system is defined in a neighbourhood U by a coordinate diffeomorphism ξ, and the other is defined in a neighbourhood V by a coordinate diffeomorphism η, then in a neighbourhood $U \cap V$ the equality $\eta = \varphi \circ \xi$ holds, where φ is a diffeomorphism $\eta \circ \xi^{-1} : \xi(U \cap V) \to \eta(U \cap V)$. □

These observations, together with the inverse function theorem, directly lead, in particular, to the following statement, known as the *local coordinate change theorem*:

Theorem 4.22. *Smooth functions at a point p*

$$y^i = u^i(x^1, \ \ldots, \ x^m), \qquad i = 1, \ \ldots, \ m,$$

are local coordinates at this point if and only if the Jacobian of the functions $u^1, \ldots, u^m$ at the point $\xi(p) \in \mathbb{R}^m$ is nonzero.

An example of local coordinates are functions defined on a linear space L that associate with each vector of this space its coordinates in a given basis. (The corresponding coordinate neighbourhood U coincides with the entire space L.) This shows that

Proposition 4.23. *any m-dimensional linear space is an m-dimensional smooth manifold.*

In particular, the arithmetic space $\mathbb{R}^m$ is an m-dimensional smooth manifold.

Pre-submanifolds of a smooth manifold, generally speaking, are not manifolds. A pre-submanifold of a smooth manifold that is itself a smooth manifold will be called a *submanifold* of that manifold. It is easy to see that

Proposition 4.24. *any open pre-submanifold W of a smooth manifold M is its submanifold.*

Proof. Indeed, for any point $p \in W$, the intersection $U \cap W$, where U is an arbitrary coordinate neighbourhood of the point p in the manifold M, is obviously a coordinate neighbourhood of the point p in the pre-manifold W. Moreover, the local coordinates in the neighbourhood $U \cap W$ are the restrictions $x^1|_{U\cap W}, \dots, x^m|_{U\cap W}$ of the local coordinates $x^1, \dots, x^m$ defined in the neighbourhood U. □

In particular,

Proposition 4.25. *any open subset of the space $\mathbb{R}^m$ is a smooth manifold.*

Let M and N be arbitrary smooth manifolds of dimensions m and n, respectively. Let us consider their topological product $M \times N$ and the natural projections

$$\pi : M \times N \to M, \quad \rho : M \times N \to N,$$

defined respectively by the formulae

$$\pi(p, q) = p, \quad \rho(p, q) = q, \qquad (p, q) \in M \times N.$$

For any function f on M, we assign a function $\widehat{f}$ to $M \times N$ by setting

$$\widehat{f} = f \circ \pi.$$

Similarly, for any function g on N, we assign a function $\widehat{g}$ to $M \times N$ by setting

$$\widehat{g} = g \circ \rho.$$

Thus,

$$\widehat{f}(p, q) = f(p), \quad \widehat{g}(p, q) = g(q).$$

for any point $(p, q) \in M \times N$.

Let $\mathcal{O}(M \times N)$ be the set of all functions on $M \times N$ that near any point $(p, q) \in M \times N$ depend smoothly on functions of type $\widehat{f}$, $f \in \mathcal{O}(M)$ and $\widehat{g}$, $g \in \mathcal{O}(N)$. (See Definition 4.3.)

It is clear that $\mathcal{O}(M \times N)$ is a smoothness on the space $M \times N$. Moreover,

Proposition 4.26. *The space $M \times N$ equipped with this smoothness is a smooth manifold of dimension $m + n$.*

Proof. Indeed, for any local coordinates $x^1, \dots, x^m$ and $y^1, \dots, y^n$ at the points p and q of the manifolds M and N respectively, the functions $\widehat{x}^1, \dots, \widehat{x}^m, \widehat{y}^1, \dots, \widehat{y}^n$ form, as is easy to see, a system of local coordinates at the point (p, q). □

We will call the constructed smooth manifold $M \times N$ the *product* of the manifolds M and N. By construction,

Proposition 4.27. *natural projections*

$$\pi : M \times N \to M, \quad \rho : M \times N \to N$$

are smooth maps.

The product $M_1 \times \cdots \times M_k$ of a finite number of smooth manifolds $M_1, \ldots, M_k$ is defined similarly. By definition, functions smooth on this product are functions that depend smoothly on functions of the form $f_i \circ \pi_i$ near any point of the product, where $\pi_i : M_1, \ldots, M_k \to M_i$ is the natural projection, and f_i is an arbitrary smooth function on M_i. The natural projections π_i in this case are also smooth maps.

Clearly, the presence of a smooth manifold structure on a topological space M imposes quite strong topological restrictions on this space. (The structure of a smooth premanifold does not impose such restrictions.) In particular, since any point of an arbitrary smooth manifold M has a neighbourhood homeomorphic to an open set of the space $\mathbb{R}^m$, this manifold is locally compact and has countable local weight. (Recall that the minimum cardinality of a base for its topology is called the "weight".) Furthermore, being Hausdorff by assumption, it is regular by virtue of its local compactness. Thus, any smooth manifold is a regular locally compact space of countable local weight.

In particular,

Proposition 4.28. *any smooth manifold is a compactly generated (aka ‘kaonic") space of the second category.*

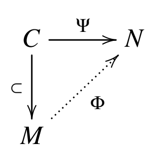
Moreover,

Proposition 4.29. *any separable smooth manifold is a paracompact space with a countable base.*

Remark 4.30. In particular, any separable manifold is a normal space. This is generally not true for non-separable manifolds.

4.4 E-manifolds

We call a smooth manifold N an *E-manifold* if, for any smooth manifold M, its any compact subset $C \subset M$, and any smooth map $\Psi : C \to N$, there exists a smooth map $\Phi : M \to N$ such that $\Psi = \Phi|_C$.



In this section, we prove that

Proposition 4.31. *every open convex subset of* $\mathbb{R}^m$ *is an E-manifold.*

To this end, we first prove that

Proposition 4.32. *for any compact subset* C *of an arbitrary smooth manifold* M *and any open set* $U \supset C$*, there exists a smooth function* f *on* M*, taking values from the interval* $I = [0, 1]$*, equal to one on* C *and equal to zero outside* U*.*

This statement can be viewed as a refinement of Urysohn's lemma for the case of smooth manifolds. We will therefore call the function f envisaged by it the *smooth Urysohn function* of the pair (U, C).

First, we consider the case $M = \mathbb{R}^m$.

Let $a \leq t \leq b$ and let

$$F(t) = k \int_a^b \exp\left[-\frac{1}{(t-a)(b-t)}\right] dt,$$

where the normalising factor k is chosen such that $F(b) = 1$. It is easy to see that all derivatives of the function $F(t)$ at points a and b are equal to zero. Therefore, by defining the function $F(t)$ for all values of t by the formula

$$F(t) = \begin{cases} 0, & \text{if} \quad t \leq a, \\ 1, & \text{if} \quad t \geq 1, \end{cases}$$

we obtain an infinitely differentiable function on the entire $\mathbb{R}$ axis, taking values from the I interval, equal to zero for $t \leq a$ and one for $t \geq b$.

Now let $\mathbb{E}$ and $\mathbb{E}' \supset \mathbb{E}$ be two (closed) concentric balls in $\mathbb{E}^m$ with radii $\rho > 0$ and $\rho' > \rho$, respectively. We define a numerical function g on $\mathbb{R}^m$ by setting for any point $\boldsymbol{t} \in \mathbb{R}^m$

$$g(\boldsymbol{t}) = F(r),$$

where F is the function we just constructed with $a = \rho$ and $b = \rho'$ and r is the distance of the point $\boldsymbol{t}$ from the common centre of the balls $\mathbb{E}$ and $\mathbb{E}'$.

Let us consider all possible closed balls of the space $\mathbb{R}^m$ that contain at least one point of the set C and are entirely contained in the open set U. Since the set C is, by assumption, compact, there exists a finite system $\mathbb{E}_1, \ldots, \mathbb{E}_k$ of such balls covering the set C. Since the set U is open, for each ball $\mathbb{E}_i$ there exists a concentric ball $\mathbb{E}_i'$ of larger radius such that $\mathbb{E}_i' \subset U$. Let g_i be the function g constructed for the balls $\mathbb{E}_i$ and $\mathbb{E}_i'$. Clearly, the function $g_1 \cdots g_k$ is infinitely differentiable, takes values from the interval I, is zero on the interior of each ball $\mathbb{E}_i$ and therefore on the set C, and is one outside all balls $\mathbb{E}_i'$ and therefore outside the set U. Therefore, the function $f = 1 - g_1 \cdots g_k$ is the desired smooth Urysohn function of the pair (U, C).

Let us now turn to the general case of an arbitrary manifold M. Let p be an arbitrary point of the set C, and let U_p be some of its coordinate neighbourhood. Since the manifold M is a regular locally compact space, the point p has neighbourhoods W_p and V_p such that $\overline{W}_p \subset V_p$, $\overline{V}_p \subset U_p \cap U$, and the closure v of the neighbourhood W_p is compact. Let W_p^0 and V_p^0 be the images of the neighbourhoods W_p and V_p under the coordinate diffeomorphism $\xi : U_p \to \mathbb{R}^m$. Since the set $\overline{W}_p^0 = \xi(\overline{W}_p)$ is compact, and the set V_p^0 is open and contains the set W_p^0, the already proven special case of the statement under consideration is applicable to these sets. Consequently, on the space $\mathbb{R}^m$ there exists an infinitely differentiable function h_p, taking values from the interval I, equal to one on the set $\overline{W}_p^0$ and zero outside the set V_p^0.

We define a numerical function g_p on the manifold M, setting for any point $q \in M$

$$g_p(q) = \begin{cases} h_p(x^1(q), \ldots, x^m(q)), & \text{if} \quad q \in U_p, \\ 0, & \text{if} \quad q \notin U_p, \end{cases}$$

where $x^1, \ldots, x^m$ are local coordinates defined in the neighbourhood U_p. It is clear that the function g_p is uniquely defined. Moreover, near any point $q \in M$, it either coincides with the smooth function $h_p(x^1, \ldots, x^m)$ at q, or with the smooth function 0 (recall that, by hypothesis, $\overline{V}_p \subset \overline{U}_p$). Consequently, this function is smooth on the manifold M. It takes values from the interval I, is equal to one on the set $\overline{W}_p$, and is zero outside the set V_p.

Since the set C is compact, it contains a finite system of points $p_1, \ldots, p_k$ such that the corresponding sets $\overline{W}_{p_1}, \ldots, \overline{W}_{p_k}$ form a covering of the set C. Let $V_{p_1}, \ldots, V_{p_k}$ be neighbourhoods of V_p, and $g_{p_1}, \ldots, g_{p_k}$ be functions of g_p corresponding to points $p_1, \ldots, p_k$. Clearly, the function

$$f = 1 - (1 - g_{p_1}) \cdots (1 - g_{p_k})$$

takes values from the interval I, is equal to one on any set $\overline{W}_{p_i}$, and therefore on the set C, and is equal to zero outside all sets V_{p_i}, and therefore outside the set U.

Thus, this function is the desired smooth Urysohn function of the pair (U, C).

Let us now consider the question of a similar refinement of the theorem on the existence of a partition of unity. Let $\{U_\alpha; \alpha \in A\}$ be an arbitrary locally finite open covering of M. A partition of unity $\{f_\alpha; \alpha \in A\}$ subordinate to the covering $\{U_\alpha; \alpha \in A\}$ will be called *smooth* if all its constituent functions f_α are smooth functions on M. From the just-proven refinement of Urysohn's lemma it follows immediately that

Proposition 4.33. *if the manifold M is a normal space, then for each of its open locally finite coverings $\{U_\alpha; \alpha \in A\}$, which has the property that for any $\alpha \in A$ the closure $\overline{U}_\alpha$ of the set U_α is compact, there exists a smooth partition of unity $\{f_\alpha; \alpha \in A\}$ subordinate to this covering.*

To prove this statement, it suffices to repeat (with appropriate trivial changes) the proof of the similar statement from §1.3.

We obtain an important modification of the previous proposition by considering an arbitrary compact subset C in the manifold M (no longer necessarily normal). Let $\{U_\alpha; \alpha \in A\}$ be an arbitrary family of open sets of M that is a cover of the set C (i.e., such that $C \subset \cap U_\alpha$). We will call a family $\{f_\alpha; \alpha \in A\}$

of smooth functions f_α on M subordinate to a covering $\{U_\alpha; \alpha \in A\}$, a *smooth partition of unity on the set* C, if

1) for any $\alpha \in A$ the function f_α is non-negative and its support $\overline{[f_\alpha \neq 0]}$ is contained in the set U_α;

2) for any point $p \in C$ the following equality holds

$$\sum_{\alpha \in A} f_\alpha(p) = 1.$$

Naturally, it is assumed that the last condition is meaningful. This assumption is certainly satisfied if the covering $\{U_\alpha; \alpha \in A\}$ is finite. On the other hand, given the compactness of the set C, the assumption that the covering $\{U_\alpha; \alpha \in A\}$ is finite essentially does not limit generality. From now on, we will always assume that this covering is finite.

Now suppose that for a (finite) covering $\{U_\alpha; \alpha \in A\}$ there exists an open (in M) covering $\{V_\alpha; \alpha \in A$ of C such that for any $\alpha \in A$ the set $\overline{V}_\alpha$ is compact and contained in U_α. (This condition now replaces both the condition of normality of the manifold M and the condition of compactness of the sets U_α.) It turns out that

Proposition 4.34. *under this assumption there exists a smooth partition of unity on the set C subordinate to the covering $\{U_\alpha; \alpha \in A\}$.*

Proof. Indeed, let h_α be a smooth Urysohn function of the pair $\{U_\alpha, \overline{V}_\alpha$. Consider the function

$$g_\alpha = \frac{h_\alpha}{\sum_{\alpha \in A} h_\alpha}.$$

Since $\sum_{\alpha \in A} h_\alpha \neq 0$ on $V = \cup_{\alpha \in A} V_\alpha$, the function g_α is defined on the open set V and is a smooth function on V. At the same time

$$\sum_{\alpha \in A} g_\alpha = 1.$$

Since the manifold M is a regular space, the compact set C has a neighbourhood W such that $\overline{W} \subset V$. Let h be the smooth Urysohn function of the pair (W, C). For any $\alpha \in A$, we define a function f_α on the manifold M by setting

$$f_\alpha() = \begin{cases} g_\alpha(p)h(p)), & \text{if} \quad p \in V, \\ 0, & \text{if} \quad p \notin V. \end{cases}$$

Clearly, this function is smooth on M and its support is contained in the set U_α. Furthermore,

$$\sum_{\alpha \in A} f_\alpha = 1 \quad \text{on} \quad C.$$

Thus, the above statement is completely proven. □

Now we can easily prove that

Proposition 4.35. *the number line $\mathbb{R}$ is an E-manifold.*

Proof. Indeed, let M be an arbitrary smooth manifold and C be its arbitrary compact subset. We must prove that for any smooth function g on C, there exists a smooth function f on M such that $f|_C = g$. Let p be an arbitrary point of C. By the definition of smooth functions on C, there exists a neighbourhood $U_p \subset M$ of the point p and a smooth function g_p on M such that

$$g|_{U_p \cap C} = g_p|_{U_p \cap C}.$$

Let V_p be a neighbourhood of the point p whose closure $\overline{V}_p$ is compact and contained in the neighbourhood U_p. Since the set C is compact, there exists a finite system of points $p_1, \ldots, p_k \in C$ such that the set

$$V = \cup_{i=1}^{n} V_{p_i}$$

contains the set C. It is clear that the corresponding neighbourhoods $U_{p_1}, \ldots, U_{p_n}$ form a covering of the set C that satisfies all the conditions of the previous proposition. Let $\{f_1, \ldots, f_n\}$ be the corresponding smooth partition of unity on the set C. Since for any point $p \in C$ and any $i = 1, \ldots, n$ either

$$g_{p_i} = g_p, \quad \text{if} \quad p \in U_i \cap C, \quad \textit{or}$$
$$f_i(p) = 0, \quad \text{if} \quad p \notin U_i \cap C,$$

then the formula

$$f = f_1 g_{p_1} + \cdots + f_n g_{p_n}$$

defines a smooth function on M possessing the required property (i.e., coinciding on C with the function g).

Thus, the above statement (which is an analogue of Tietze's theorem) is completely proven. □

Remark 4.36. Remark. It is clear that the support $\overline{[f \neq 0]}$ of the constructed function f is compact. Thus, we have in fact proved that

Proposition 4.37. *any smooth function on a compact subset C of M is the restriction of some smooth function on M with compact support.*

It follows from the proved statement that

Proposition 4.38. *for any $m \geq 1$, the m-dimensional arithmetic space $\mathbb{R}^m$ is an E-manifold.*

Proof. Indeed, any smooth map $C \to \mathbb{R}^m$ is defined by m smooth functions, and in order to extend this map to the entire manifold M, it is sufficient to extend these functions to M. □

Let us now consider the unit open ball $\mathbb{E}^m$ of the space $\mathbb{R}^m$. It is easy to see that

Proposition 4.39. *the ball* $\mathbb{E}^m$ *is diffeomorphic to the space* $\mathbb{R}^m$.

Proof. Indeed, by associating with each point $\boldsymbol{t} = (t^1, \ldots, t^m)$ of the space $\mathbb{R}^m$ a point $\boldsymbol{v} == (v^1, \ldots, v^m)$ of the ball $\mathbb{E}^n$ with coordinates

$$v^i = \frac{2}{\pi}\frac{\arctan|\boldsymbol{t}|}{|\boldsymbol{t}|}t^i \quad i = 1, \ldots, m,$$

we obviously obtain a diffeomorphism

$$\mathbb{R}^m \to \mathbb{E}^m.$$

□

Since a manifold diffeomorphic to an E -manifold is itself obviously an E-manifold, it follows that

Proposition 4.40. *the ball* $\mathbb{E}^m$ *is an E-manifold.*

Now let D be an arbitrary open bounded convex subset of the space $\mathbb{R}^m$. To prove that this subset is an E-manifold, it suffices, according to what was said above, to prove that it is diffeomorphic to the ball $\mathbb{E}^m$.

For this purpose, we choose an arbitrary point $\boldsymbol{x}_0 \in D$ and consider the function $\varphi(\boldsymbol{u})$ of vectors $\boldsymbol{u} \in \mathbb{S}^{m-1}$, which defines the boundary of the body $P = \overline{D}$ in "polar coordinates" with centre at the point $\boldsymbol{x}_0$. Simple elementary geometric considerations show that the function φ satisfies the Lipschitz condition, i.e., there exists a number L such that

$$|\varphi(\boldsymbol{u}_1) - \varphi(\boldsymbol{u}_2)| \leq L|\boldsymbol{u}_1 - \boldsymbol{u}_2|$$

for any points $\boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbb{S}^{m-1}$.

On the other hand, it is easy to see that

Proposition 4.41. *For any function* φ *satisfying the Lipschitz condition on the sphere* $\mathbb{S}^{m-1}$, *there exists a continuous function* g *in the closed unit ball* $\mathbb{E}^m$ *such that:*

1) the function g *in the open ball* $E^m = \mathring{\mathbb{E}}^m$ *is a smooth function;*

2) on the sphere $\mathbb{S}^{m-1}$, *the function* g *coincides with the function* φ;

3) at the point $\mathbf{0}$, *the function* g *is nonzero;*

4) along each radius of the ball $\mathbb{E}^m$, *the function* g *increases monotonically.*

Such a function g can, for example, be defined by the formula

$$g(v\boldsymbol{u}) = A\left(\exp\left(1 - \frac{1}{v^2}\right) - 1\right) +$$
$$\frac{2^n \exp(1 - \frac{1}{v^2})}{1 - v^n}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left[\prod_{i=1}^{m}\omega\left(\frac{2(t^i - u^i)}{1 - v}\right)\right]\varphi\left(\frac{\boldsymbol{t}}{|\boldsymbol{t}|}\right)dt^1\cdots dt^m,$$
$$0 \leq v \leq 1, \quad \boldsymbol{u} \in \mathbb{S}^{m-1},$$

where

a) A is some sufficiently large positive number;

b) $u^1, \ldots, u^m$ are the coordinates of the point $\boldsymbol{u} \in \mathbb{S}^{m-1}$,

c) $\boldsymbol{t}$ is a point in space $\mathbb{R}^m$ with coordinates $t^1, \ldots, t^m$;

d) $\omega(t)$ is a smooth function defined by the formula

$$\omega(t) = \begin{cases} \frac{1}{k}\exp\left(\frac{1}{t^2-1}\right), & \text{if } |t| \le 1, \\ 0, & \text{if } |t| \ge 1, \end{cases} \qquad k = \int_{-1}^{1} \exp\left(\frac{1}{t^2-1}\right) dt.$$

Verifying that the function g defined in this way satisfies all conditions 1) - 4) is quite elementary and we will omit it.

Having constructed a function g for the function φ corresponding to the convex set D, we define a map

$$\Pi : E^m = \mathring{\mathbb{E}}^m \to D,$$

by setting for any point $v\boldsymbol{u} \in \mathbb{E}^m$, $0 \le v < 1$, $\boldsymbol{u} \in \mathbb{S}^{m-1}$,

$$\Phi(v\boldsymbol{u}) = \boldsymbol{x}_0 + g(v\boldsymbol{u})\boldsymbol{u}.$$

It is easy to verify that the map Φ constructed in this way is a diffeomorphism. We have thus proved that

Proposition 4.42. *any open bounded convex subset D of $\mathbb{R}^m$ is diffeomorphic to the ball $E^m = \mathring{\mathbb{E}}^m$ and is therefore an E-manifold.*

Finally, let us consider the case where the open convex set D is unbounded. Let M be an arbitrary smooth manifold, C its arbitrary compact subset, and Ψ an arbitrary smooth map of C into the convex set D. Since the set C is compact and the map Ψ is continuous, the image $\Psi(C)$ of the set C under the map Ψ is compact and therefore bounded, i.e., it is contained in some open ball of the space $\mathbb{R}^m$ of sufficiently large radius. Let D' be the intersection of D with this ball. The set D' is open, bounded, and convex. Therefore, by what has already been proved, this set is a E-manifold. On the other hand, since $\Psi(C) \subset D'$, we can regard the map Ψ as a map $C \to D'$. Therefore, there exists a smooth map $\Phi : M \to D'$ such that $\Phi|_C = \Psi$. Finally, since $D' \subset D$, we can assume that Φ is a smooth map $M \to D$.

Thus, the proposition formulated at the beginning of this section is fully proven.

Remark 4.43. The case of an unbounded convex set D can be treated similarly to the case of a bounded set, by proving that any such set is also diffeomorphic to the ball $E^m = \mathring{\mathbb{E}}^m$. If the set D does not contain any lines (for example, the interior of a paraboloid of revolution belongs to this type), then by means of a suitable projective transformation it is immediately transformed into some bounded convex set. If the set D contains at least one line, then, as is easily proved, it is a cylinder over some $m-1$-dimensional convex set. In this case, a diffeomorphism of the set D onto the ball $E^m = \mathring{\mathbb{E}}^m$ is easily constructed by induction on the number m.

In conclusion, consider an arbitrary open submanifold W of some smooth manifold M. By definition, the restriction $|f_W$ of an arbitrary smooth function f on M is a smooth function on W. In other words, the correspondence $f \mapsto f|_W$ defines a certain map (the so-called *restriction map*)

$$\mathscr{O}(M) \to \mathscr{O}(W),$$

which is obviously a homomorphism of the algebra $\mathscr{O}(M)$ into the algebra $\mathscr{O}(W)$.

Generally speaking, the restriction map $\mathscr{O}(M) \to \mathscr{O}(W)$ is not an epimorphism, i.e. (in contrast to the case of compact submanifolds) not every smooth function on W is the restriction of some function that is smooth on M. However, from the results proved above, it immediately follows that this map is *locally epimorphic* in the sense that

Proposition 4.44. *whatever the open set $V \subset W$ whose closure $\overline{V}$ is compact and contained in W, for any smooth function g on W there exists a smooth function f on M such that $f|_V = g|_V$.*

4.5 Vector fields

Let M be an arbitrary smooth manifold. By a *vector field* on M we mean an arbitrary derivation of the algebra $\mathscr{O}(M)$, i.e., a linear map X of this algebra into itself such that

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

for any functions $f, g \in \mathscr{O}(M)$. Obviously, the set $\mathscr{O}^1(M)$ of all vector fields on the manifold M is a $\mathscr{O}(M)$-module with respect to the operations

$$(X+Y)g = Xg + Yg, \quad X, Y \in \mathscr{O}^1(M), \quad g \in \mathscr{O}(M),$$
$$(fX)g = f \cdot Xg, \quad X \in \mathscr{O}^1(M), \quad f, g \in \mathscr{O}(M).$$

Moreover, it is easy to see that for any vector fields $X, Y \in \mathscr{O}^1(M)$, the map

$$X, Y \in \mathscr{O}^1(M)$$

defined by the formula

$$[X, Y]g = X(Yg) - Y(Xg), \quad g \in \mathscr{O}(M),$$

is also a vector field, and the operation $[X, Y]$ defines the module $\mathscr{O}^1(M)$ as a Lie algebra, i.e., this operation is anti-commutative:

$$[X, Y] = -[Y, X], \quad X, Y \in \mathscr{O}^1(M)$$

and satisfies the Jacobi identity:

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0, \quad X, Y, Z \in \mathscr{O}^1(M).$$

Let W be an arbitrary open submanifold of M and let X be an arbitrary vector field on M. As noted at the end of the previous section, for any smooth function g on W and any point $p \in W$, there exists a smooth function f on M such that $g = f$ on some neighbourhood of p. Let f_1 and f_2 be two smooth functions on M with this property, and let V be a neighbourhood of the point p such that $f_1 = f_2 = g$ on V. Let, in addition, h be a smooth function on the manifold M equal to one at some point $q \in V$ and zero outside the neighbourhood V. Then $h(f_1 - f_2) = 0$ on M and therefore

$$h \cdot X(f_1 - f_2) + Xh \cdot (f_1 - f_2) = 0 \quad \text{on} \quad M.$$

Since

$$h(q) = 1, f_1(q) = f_2(q), \quad X(f_1 - f_2) = Xf_1 - Xf_2,$$

it follows that

$$(Xf_1)(q) - (Xf_2)(q) = (Xh)(q) \cdot (f_1(q) - f_2(q)) = 0,$$

i.e., that

$$(Xf_1)(q) = (Xf_2)(q).$$

Therefore, setting

$$\overline{g}(q) = (Xf)(q), \quad q \in V,$$

where f is an arbitrary smooth function on M for which $f|_V = g|_V$, we obtain on V some uniquely defined, smooth on V, function $\overline{g}$. It is clear that the functions $\overline{g}$ corresponding to different neighbourhoods V of the point p take the same value at the point p. Therefore, by setting

$$(Yg)(p) = \overline{g}(p), \quad p \in V,$$

we uniquely define on W some function Yg. This function near any point $p \in W$ coincides with a smooth function g at p and is therefore itself a smooth function on W. Therefore, the correspondence $g \mapsto Yg$ defines a map

$$Y : \mathcal{O}(W) \to \mathcal{O}(W),$$

and it is easy to verify that this map is a derivation of the algebra $\mathcal{O}(W)$, i.e., it is a vector field on the manifold W.

We will call the field Y the *restriction* of the field X on the open submaniold W and will denote it by the symbol $X|_W$ and even simply by the symbol X if there is no misunderstanding. It is easy to verify that the restriction map

$$\mathcal{O}^1(M) \to \mathcal{O}^1(W),$$

assigning to each field $X \in \mathcal{O}^1(M)$ a field $X|_W \in \mathcal{O}^1(W)$, reserves all the algebraic operations introduced above, i.e., it is both a module homomorphism (if the $\mathcal{O}(W)$-module $\mathcal{O}^1(W)$ is considered, via the restriction map $\mathcal{O}(M) \to \mathcal{O}(W)$, as a $\mathcal{O}(M)$-module), and a Lie algebra homomorphism.

Just as for functions, the restriction map $\mathcal{O}^1(M) \to \mathcal{O}^1(W)$ is, generally speaking, not epimorphic, i.e., not every field $Y \in \mathcal{O}^1(W)$ is a restriction of some field $X \in \mathcal{O}^1(M)$. It is only locally epimorphic in the sense that,

Proposition 4.45. *for any open set $V \subset W$ whose closure $\overline{V}$ is compact and contained in W, for any field $Y \in \mathcal{O}^1(W)$ there exists a field $X \in \mathcal{O}^1(M)$ such that*

$$X|_V = Y|_V.$$

This field X can be defined, for example, by the formula

$$(Xf)(p) = \begin{cases} h(p) \cdot (Y(f|_W))(p), & \text{if} \quad p \in W, \\ 0, & \text{if} \quad p \notin W, \end{cases}$$

where f is an arbitrary smooth function on M, p is an arbitrary point of the manifold M, and h is a smooth function on M equal to one on $\overline{V}$ and equal to zero outside some open set $U \supset \overline{V}$ whose closure $\overline{U}$ belongs to W.

In the special case where an open submanifold W is a coordinate neighbourhood U, one can easily describe all vector fields on W. Let $x^1, \ldots, x^m$ be local coordinates defined in a neighbourhood U. As we know, for any smooth function f on U and any point $p \in U$, there exists a neighbourhood $V \subset U$ of p such that on V the function f depends smoothly on the coordinates $x^1, \ldots, x^m$, i.e.,

$$f = u(x^1, \ldots, x^m) \quad \text{on} \quad V$$

where $u = u(t^1, \ldots, t^m)$ is some smooth function on $\mathbb{R}^m$ (uniquely determined by f). Let V_1 be another relative position of the point p with the same property, i.e. such that $f = u_1(x^1, \ldots, x^m)$ on $\mathbb{R}^m$ where $u_1 = u_1(t^1, \ldots, t^m)$ is some smooth function on $\mathbb{R}^m$. It is clear that the functions u and u_1 coincide in some neighbourhood of the point $\xi(p) = (x^1(p), \ldots, x^m(p))$ (namely, in the neighbourhood $\xi(V \cap V_1)$, which is the image under the coordinate diffeomorphism $\xi : U \to \mathbb{R}^m$ of the neighbourhood $V \cap V_1$ of the point p), and therefore their partial derivatives $\frac{\partial u}{\partial t^i}$ and $\frac{\partial u_1}{\partial t^i}$, $i = 1, \ldots, m$, take the same value at the point $\xi(p)$. It follows that formula

$$h_i(p) = \left(\frac{\partial u}{\partial t^i}\right)_{\xi(p)}, \quad i = 1, \ldots, m,$$

uniquely defines on U some functions $h_1, \ldots, h_m$. In a neighbourhood V, the functions h_i coincide with the smooth functions $\frac{\partial u}{\partial t^i}(x^1, \ldots, x^m)$ and are therefore smooth functions in p. Since the point p is arbitrary, this means that every function h_i is smooth on U. We will denote it by the symbol $\frac{\partial f}{\partial x^i}$ and call it the *partial derivative* of the function f with respect to the coordinate x^i. According to this definition,

Proposition 4.46. *if*

$$f = u(x^1, \ldots, x^m) \quad \textit{on} \quad V \subset U,$$

then

$$\frac{\partial f}{\partial x^i} = \frac{\partial u}{\partial t^i}(x^1, \ldots, x^m) \quad \textit{on} \quad V.$$

It is easy to check that the map

$$\frac{\partial}{\partial x^i} : f \mapsto \frac{\partial f}{\partial x^i}, \quad i = 1, \ldots, m,$$

is a vector field on U and that the vector fields constructed in this way

$$\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^m}$$

satisfy the relations

$$\frac{\partial x^j}{\partial x^i} = \delta_i^j, \quad i = 1, \ldots, m,$$

where δ_i^j is the Kronecker delta ($\delta_i^j = 0$, if $i \neq j$, and $\delta_i^j = 1$, if $i = j$.)

In particular, we see that

Proposition 4.47. *the vector fields $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^m}$ are linearly independent (in the $\mathcal{O}(U)$-module $\mathcal{O}^1(U)$).*

In the particular case $M = \mathbb{R}^m$, the neighbourhood U can be taken to be the entire space $\mathbb{R}^m$, and the local coordinates $x^1, \ldots, x^m$ can be taken to be the functions that associate with each point $\boldsymbol{t} \in \mathbb{R}^m$ its coordinates $t^1, \ldots, t^m$. Clearly, in this case,

Proposition 4.48. *the vector fields $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^m}$ represent the usual operations of partial differentiation with respect to the variables $t^1, \ldots, t^m$.*

In particular, for $m = 1$ there is only one coordinate $t = t^i$ and we obtain that

Proposition 4.49. *the vector field $\frac{\partial}{\partial t}$ represents the differentiation $\frac{d}{dt}$ with respect to the variable t.*

Returning to a smooth function f on U and a neighbourhood V in which $f = u(x^1, \ldots, x^m)$, we can, without loss of generality, obviously assume that the image $\xi(V)$ of the neighbourhood V under the coordinate diffeomorphism

$$\xi(q) = (x^1(q), \ldots, x^m(q)), \quad q \in V$$

is an open ball in the space $\mathbb{R}^m$ centred at $\mathbf{0}$. In this open ball the function $u(t^1, \ldots, t^m)$ admits the representation[1]

$$\begin{aligned} u(t^1, \ldots, t^m) &= u(0, \ldots, 0) + \int_0^1 \frac{\partial}{\partial s} u(st^1, \ldots, st^m) ds \\ &= u(0, \ldots, 0) + t^i \int_0^1 \frac{\partial u}{\partial t^i}(st^1, \ldots, st^m) ds. \end{aligned}$$

[1]Here, as everywhere below, we (following the convention adopted in tensor calculus) assume that summation is performed over two identical indices, one upper and the other lower (the limits of which are each time clear from the context).

Therefore, defining smooth functions on V $g_i = g_i(q)$ by the formula

$$g_i(q) = \int_0^1 \frac{\partial u}{\partial t^i}(sx^1, \ldots, sx^m)ds,$$

we obtain the relation

$$f = f(p) + x^i g_i \quad \text{on} \quad V, \tag{4.50}$$

where the number $f(p)$ is considered as a constant function on V, taking the same value $f(p)$ at all points $q \in V$.

Let V be an arbitrary vector field on a neighbourhood U. It is easy to see that the formula

$$X_p f = (Xf)(p), \quad p \in Y, \quad f \in \mathcal{O}(p),$$

where X on the right-hand side is the restriction of X to the corresponding neighbourhood of the point p, defines a certain linear map X_p of the algebra $\mathcal{O}(p)$ into the field $\mathbb{R}$, with the property that

$$X_p(fg) = X_p f \cdot g(p) + fp \cdot X_p g \tag{4.51}$$

for any functions $f, g \in \mathcal{O}(p)$.

From relation (4.51) it immediately follows that

Proposition 4.52. *$X_p(c) = 0$ for any constant c.*

Therefore, applying the map X_p to both parts of formula (4.50), we obtain that

$$X_p f = X_p x^i \cdot g_i(p).$$

But,

$$g_i(p) = \int_0^1 \frac{\partial u}{\partial t^i}(0, \ldots, 0)ds = \frac{\partial u}{\partial t^i}(0, \ldots, 0) = \frac{\partial f}{\partial x^i}(p).$$

Thus,

$$X_p f = X_p x^i \cdot \frac{\partial f}{\partial x^i}(p). \tag{4.53}$$

Since by definition $X_p = (Xf)(p)$ and since the function f is an arbitrary smooth function on U, and P is an arbitrary point in U, then

$$X = Xx^i \cdot \frac{\partial}{\partial x^i} \quad \text{on} \quad U. \tag{4.54}$$

Since the fields $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^m}$ are linearly independent, it is thus proved that

Proposition 4.55. *the module $\mathcal{O}^1(U)$ over the algebra $\mathcal{O}(U)$ has a basis $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^m}$ consisting of m elements; any field $X \in \mathcal{O}^1(U)$ can be expressed in terms of this basis by formula (4.54).*

In the special case when $M = \mathbb{R}^m$, we obtain from this that

Proposition 4.56. *the partial derivations $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^m}$ form a basis for the module $\mathcal{O}^1(\mathbb{R}^m)$.*

A basis $X_1, \dots, X_m$ of a module $\mathcal{O}^1(U)$ consisting of m fields will be called *holonomic* if in U there exists a system of local coordinates $x^1, \dots, x^m$ such that

$$X_1 = \frac{\partial}{\partial x^1}, \dots, X_m = \frac{\partial}{\partial x^m}.$$

Otherwise, we will call the basis $X_1, \dots, X_m$ *non-holonomic.*

In any basis $X_1, \dots, X_m$ of the module $\mathcal{O}^1(U)$ consisting of m fields, each field $X \in \mathcal{O}^1(U)$ has the form

$$X = X^i X_i,$$

where X^i are some smooth functions on U. These functions are uniquely determined by the field X and are called its *components* in the basis $X_1, \dots, X_m$. According to formula (4.54), for a holonomic basis

$$X^i = Xx^i,$$

Linear operations defined in the $\mathcal{O}^1(U)$ module are expressed in terms of components according to the formulas

$$(fX)^i = fX^i, \quad f \in \mathcal{O}(U), X \in \mathcal{O}^1(U),$$
$$(X+Y)^i = X^i + Y^i, \quad X, Y \in \mathcal{O}^1(U).$$

As for the Lie operation $[X, Y]$, the following statement holds:

Proposition 4.57. *If the basis $X_1, \dots, X_m$ is holonomial, then*

$$[X, Y]^i = X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j}. \tag{4.58}$$

Proof. Indeed,

$$[X, Y]^i = [X, Y]x^i = X(Yx^i) - Y(Xx^i) = X(Y^i) - Y(X^i).$$

□

From formula (4.58), in particular, it follows that

Proposition 4.59. *for any holonomic basis $X_1, \dots, X_m$, the equalities hold:*

$$[X_i, X_j] = 0. \tag{4.60}$$

In the general case,

$$[X_i, X_j] = c^k_{ij} X_k,$$

where c^k_{ij} are some smooth functions on U, and formula (4.58 is replaced by the formula

$$[X, Y]^i = X^j \cdot X_j(Y^i) - Y^j \cdot X_j(X^i) + c^i_{jk} X^j Y^k.$$

Remark 4.61. Formula (4.60eq:4-5-6) means that the result of applying two operations $\frac{\partial}{\partial x^i}$ and $\frac{\partial}{\partial x^j}$ to an arbitrary smooth function on U does not depend on which of these operations is applied first. In other words, as in elementary analysis,

$$\frac{\partial^2}{\partial x^i \partial x^j} = \frac{\partial^2}{\partial x^j \partial x^i}.$$

4.6 Vectors

The linear map introduced in the previous section

$$X_p : \mathscr{O}(p) \to \mathbb{R}$$

will be called the *vector of the field X at the point p*. A linear map $A : \mathscr{O}(p) \to \mathbb{R}$ will be called a *vector of the manifold M at the point p* if in some neighbourhood of the point p (or, equivalently, on the entire manifold M) there exists a vector field X such that $X_p = A$. It is easy to see that

Proposition 4.62. *a linear map $A : \mathscr{O}(p) \to \mathbb{R}$ is a vector of the manifold M at a point p if and only if*

$$A(fg) = Af \cdot g(p) + f(p) \cdot Ag$$

for any functions $f, g \in \mathscr{O}(p)$.

Proof. The necessity of this condition follows immediately from the relation (4.51). On the other hand, in proving formula (4.53), we used only relation (4.51), so for any linear map $A : \mathscr{O}(p) \to \mathbb{R}$ satisfying the conditions of the theorem, a similar formula will hold:

$$Af = a^i \cdot A\frac{\partial f}{\partial x^i}(p), \quad \text{where} \quad a^i = Ax^i, \tag{4.63}$$

and therefore the vector field X defined on the coordinate neighbourhood U by the formula

$$X = a^i \cdot A\frac{\partial}{\partial x^i},$$

will have the property that $X_p = A$. □

The set of all vectors of the manifold M at the point p is, obviously, a linear space over the field $\mathbb{R}$. This space is called the *tangent space* of the manifold M at the point p and is denoted by the symbol M_p or the symbol $\mathscr{O}^1(p)$. It contains, in particular, the vectors

$$\left(\frac{\partial}{\partial x^1}\right)_p, \ \ldots, \ \left(\frac{\partial}{\partial x^m}\right)_p, \tag{4.64}$$

where $x^1, \ \ldots, \ x^m$ is an arbitrary system of local coordinates at the point p. Vectors (4.64) are linearly independent and any vector $A \in M_p$ is expressed linearly through them:

$$A = Ax^i \cdot \left(\frac{\partial}{\partial x^i}\right)_p,$$

(see formula (4.63)). This proves that

Proposition 4.65. *for any point $p \in M$, the dimension of the space M_p is equal to the dimension m of the manifold M.*

Let, in particular, the manifold M be an m-dimensional linear space L. It is easy to see that for any vector $A \in L$ and any smooth function f on L, the function $\frac{\partial f}{\partial A}$ defined by the formula

$$\frac{\partial f}{\partial A}(B) = \frac{df(B+tA)}{dt}|_{t=0}, \quad B \in L,$$

is a smooth function on L, and the resulting map $\frac{\partial}{\partial A} : \mathcal{O}(L) \to \mathcal{O}L)$ is a vector field on L.

Let $A_1, \ldots, A_m$ be an arbitrary basis of the space L. Setting $f(B) = f(b^1, \ldots, b^m)$, where $b^1, \ldots, b^m$ are the coordinates of the vector B in the basis $A_1, \ldots, A_m$, we immediately obtain that

$$\left(\frac{\partial f}{\partial A_i}\right)_B = \frac{\partial f}{\partial b^i}.$$

It follows that

Proposition 4.66. *for any vector $B \in L$, the vectors $\left(\frac{\partial f}{\partial A_1}\right)_B, \ldots, \left(\frac{\partial f}{\partial A_m}\right)_B$ are linearly independent and form a basis of the tangent space L_B.*

This means, in particular, that

Proposition 4.67. *for any vector $B \in L$, the map $A \mapsto \left(\frac{\partial}{\partial A}\right)_B$ of the space L into the space L_B is an isomorphism.*

In the future, we will, as a rule, identify the vectors A and $\left(\frac{\partial}{\partial A}\right)_B$ and, thus, we will assume that $L = L_B$.

Returning to an arbitrary smooth manifold M, consider the set

$$M_* = \cup_{p \in M} M_p$$

of all its vectors at all points $p \in M$. Denoting for each vector $A \in M_*$ by the symbol $\mu(A)$, where $A \in M_p$ for the point p of the manifold M, we obtain a certain single-valued map

$$\mu : M_* \to M.$$

Let U be an arbitrary coordinate neighbourhood in the manifold M and let

$$U_* = \mu^1(U)$$

be its complete preimage under the map μ. By choosing local coordinates $x^1, \ldots, x^m$ in the neighbourhood U and constructing the corresponding coordinate homeomorphism ξ of the neighbourhood U onto an open set U^0 of $\mathbb{R}^m$, we define a map

$$\xi_* : U_* \to U^0 \times \mathbb{R}^m$$

of the set U_* into the topological product $U^0 \times \mathbb{R}^m \subset \mathbb{R}^{2m}$, setting for any point $A \in U_*$

$$\xi_*(A) = (\boldsymbol{t}, \boldsymbol{a}),$$

where $\boldsymbol{t}$ is the image of the point $p = \mu(A)$ under the homeomorphism ξ, and $\boldsymbol{a} = (a^1, \dots, a^m)$ is a point of $\mathbb{R}^m$ such that

$$A = a^i \left(\frac{\partial}{\partial x^i}\right)_p .$$

It is easy to see that the map ξ_* is bijective.

We introduce a topology into the set U_* with respect to which the map ξ_* is a homeomorphism. Clearly, this topology is uniquely defined and does not depend on the choice of the local coordinate system $x^1, \dots, x^m$. In particular, for any two coordinate neighbourhoods $U, V \subset M$, the topologies induced by the topologies of the sets U_* and V_* on the set $U_* \cap V_* = (U \cap V)_*$ coincide with each other, i.e., the intersection $U_* \cap V_*$ is a subspace of each of the spaces U_* and V_*.

Moreover, it is clear that this subspace is open in both U_* and V_*. Therefore (see §1.1) on the set M_* there exists a unique topology with respect to which this set is the free union of all subspaces of the form U_*. Obviously, in this topology, the set M_* is a Hausdorff space.

A real function F defined on the space M_* will be considered smooth if, for any coordinate neighbourhood $U \subset M$ and any local coordinate system $x^1, \dots, x^m$ defined on the neighbourhood U, the function $F \circ \xi_*^{-1}$ on the product $U^0 \times \mathbb{R}^m \subset \mathbb{R}^{2m}$, where ξ_* is the map $U_* \to U^0 \times \mathbb{R}^m$ constructed above, is a smooth function on $U^0 \times \mathbb{R}^m$. It is easy to verify that the set $\mathbb{O}(M_*)$ of all smooth functions on M_* is smooth on M_* and that the space M_* equipped with this smoothness is a smooth manifold. For coordinate neighbourhoods in M_* we can take, for example, subspaces U_*, and for coordinate homeomorphisms - maps

$$\xi_* : U_* \to \mathbb{R}^{2m}.$$

We will call the constructed $2m$-dimensional smooth manifold M_* the *manifold of tangent vectors* (or the *tangent bundle*) of the manifold M.

Continuous maps

$$X : M \to M_*$$

of a manifold M into the manifold M_* for which

$$\mu \circ X = 1_M,$$

i.e., such that $X(p) \in M_p$ for any point $p \in M$, we will call *continuous vector fields on the manifold* M. A continuous field X, which is a smooth map of the manifold M into the manifold M_*, we will call a *smooth vector field* on the manifold M.

The condition of smoothness of the field X means, by definition, that for any smooth function F on M_*, the function $f(p) = F(X(p))$ of the point $p \in M$ is a smooth function on M. Since for each coordinate homeomorphism $\xi_* : U_* \to \mathbb{R}^{2m}$ the relation holds, this condition is equivalent to the smoothness condition (for each coordinate neighbourhood $U \subset M$) of the map

$$\xi_* : X|_U \to \mathbb{R}^{2m}.$$

Let $x^1, \ldots, x^m$ be the local coordinates in the neighbourhood U with the help of which the homeomorphism ξ_* was constructed, and let

$$X^i(p) = X(p)x^i$$

be the coordinates of the vector $X(p)$, $p \in U$, in the basis $\left(\frac{\partial}{\partial x^1}\right)_p, \cdots, \left(\frac{\partial}{\partial x^m}\right)_p$. Then, according to the definition of the homeomorphism ξ_*,

$$(\xi_* \circ X)(p) = (x^1(p), \ldots, x^m(p), X^1(p), \ldots, X^m(p))$$

for any point $p \in U$. Since the local coordinates $x^1, \ldots, x^m$ are smooth functions on U, it follows that the map $\xi_* \circ X|_U$ is smooth if and only if all functions $X^i(p)$ are smooth. Since for any smooth function f on U the equality holds,

$$X(p)f = X^i(p)\frac{\partial f}{\partial x^i},$$

we thus finally obtain that

Proposition 4.68. *A continuous vector field*

$$X : M \to M_*$$

is smooth if and only if for any smooth function f on M, the function

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

at a point $p \in M$ is a smooth function on M.

Let us now consider on the manifold M an arbitrary vector field $X : \mathcal{O}(M) \to \mathcal{O}(M)$ in the sense of the previous section. It is clear that, setting

$$X(p) = X_p, \quad p \in M,$$

we obtain a certain map

$$X : M \to M_*,$$

which is a continuous vector field on the manifold M. Moreover, since for any smooth function f on M the equality holds

$$X(p)f = X_p f = (Xf)(p),$$

the continuous vector field constructed in this way $X : M \to M_*$ is smooth.

Conversely, let $X : M \to M_*$ be an arbitrary smooth vector field on the manifold M. We define a map $X : \mathcal{O}(M) \to \mathcal{O}(M)$ of the algebra $\mathcal{O}(M)$ into itself by setting

$$(Xf)(p) = X_p f$$

for any smooth function f on M and any point $p \in M$. It is clear that the map X of the algebra $\mathcal{O}(M)$ into itself defined in this way is linear. Moreover, for any functions $f, g \in \mathcal{O}(M)$, at each point $p \in M$ the equality holds

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

Consequently, the map X is a derivation of the algebra $\mathcal{O}(M)$, i.e., a vector field on the manifold M in the sense of the previous section. Thus,

Proposition 4.69. *the concept of a smooth vector field essentially coincides with the concept of a vector field as a derivation of the algebra $\mathcal{O}(M)$.*

The concept of a vector at a point allows us to specify a simple, but often useful, necessary and sufficient condition for a system of vector fields $X_1, \ldots, X_m$, defined on a coordinate neighbourhood U, to constitute a basis (not necessarily holonomic) of the module $\mathcal{O}^1(U)$. Namely, as is easy to see,

Proposition 4.70. *vector fields $X_1, \ldots, X_m \in \mathcal{O}^1(U)$ form a basis of the $\mathcal{O}(U)$-module $\mathcal{O}^1(U)$ if and only if for any point $p \in U$ the vectors*

$$(X_1)_p, \ \ldots, \ (X_m)_p$$

form a basis of the space M_p.

Proof. Indeed, let the fields $X_1, \ldots, X_m$ form a basis of the $\mathcal{O}(U)$-module $\mathcal{O}^1(U)$ and let A be an arbitrary vector in the space M_p. By definition, in the neighbourhood U there exists a vector field $X \in \mathcal{O}^1(U)$ such that $X_p = A$. Let

$$X = X^i X_i$$

be the decomposition of the field X in the basis $X_1, \ldots, X_m$. Then

$$A = X^i(p) \cdot (X_i)_p.$$

Therefore, the vectors $(X_1)_p, \ldots, (X_m)_p$ generate the space M_p. Since there are m of them, they are linearly independent, i.e., they form a basis.

Conversely, let fields $X_1, \ldots, X_m \in \mathcal{O}^1(U)$ have the property that for any point $p \in U$ the vectors $(X_1)_p, \ldots, (X_m)_p$ form a basis for the space M_p, and let X be an arbitrary vector field on the neighbourhood U. Then at any point $p \in U$ the expansion holds

$$X = \widehat{X}^i(p) \cdot (X_i)_p,$$

where $\widehat{X}^i(p)$ are some functions of the point p. Then for any $i = 1, \ldots, m$ the equality holds

$$X_p x^j = \widehat{X}^i(p) \cdot (X_i)_p x^j.$$

But the numbers $X_p x^j$ are, as we know, the values at the point p of the components $X^j = X x^j$ of the field X in the basis $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^m}$. Similarly, $(X_i)_p x^j = X_i^j(p)$, where X_i^j are the components of the field X_i, in the basis $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^m}$. Thus,

$$X^j(p) = \widehat{X}^i(p) \cdot X_i^j(p). \tag{4.71}$$

It is clear that the determinant $\det X_i^j(p)$ at any point $p \in U$ is nonzero. Therefore, using the well-known Cramer formula, we can express the numbers $\widehat{X}^i(p)$ in terms of the numbers $X^j(p)$ and $X_i^j(p)$. Since the components X^j and X_i^j are smooth functions of the point p, it follows that the functions $\widehat{X}^i = \widehat{X}^i(p)$ also depend smoothly on p. Thus,

$$X = \widehat{X}^i X_i$$

where $\widehat{X}^i$ are smooth functions on U. Consequently, the fields $X_1, \ldots, X_m \in \mathcal{O}^1(U)$ generate the $\mathcal{O}(U)$-module $\mathcal{O}^1(U)$. Since they are obviously linearly independent, they form a basis of this module. □

It immediately follows from the proved proposition that

Proposition 4.72. *for any coordinate neighbourhood U, each basis of the $\mathcal{O}(U)$-module $\mathcal{O}^1(U)$ consists of m fields.*

In connection with this proposition, it is also worth noting that, in view of the arbitrariness in the choice of local coordinates described in §4.3, the following statement holds:

Proposition 4.73. *whatever the point p of a smooth manifold M, for any basis $A_1, \ldots, A_m$ of the space M_p there exists at the point p a system of local coordinates $x^1, \ldots, x^m$ such that*

$$\frac{\partial}{\partial x^1}_p = A_1, \ \ldots, \ \frac{\partial}{\partial x^m}_p = A_m.$$

In conclusion, let us consider an arbitrary smooth map Φ at a point $p \in M$ of the manifold M into some manifold N. To any vector $A \in M_p$ we can assign the map

$$d\Phi_p(A) : \mathcal{O}(\Phi(p)) \to \mathcal{O}(\Phi(p)),$$

by setting

$$d\Phi_p(A)g = A(g \circ \Phi)$$

for any function g smooth at the point $\Phi(p) \in N$. It is easy to see that the map $d\Phi_p(A)$ constructed in this way is a vector of the manifold N at the point $\Phi(p)$. Thus, we obtain a certain (obviously linear) map

$$d\Phi_p : M_p \to N_{\Phi(p)}.$$

We will call this map the *differential of the map Φ at the point p.*

It is clear that

Proposition 4.74. *for any smooth manifolds M, N, and P and any smooth maps $\Phi : M \to N$ and $\Psi : N \to P$, at any point $p \in M$, we have the relation*

$$d(\Psi\Phi)_p = d\Psi_{\Phi(p)} d\Phi_p. \tag{4.75}$$

Let $x^1, \ldots, x^m$ be local coordinates at a point p of a manifold M, and $y^1, \ldots, y^n$ be local coordinates at a point $\Phi(p)$ of a manifold N. Since the map Φ is smooth, the functions $y^1 \circ \Phi, \ldots, y^m \circ \Phi$ (defined in some neighbourhood of the point p) are smooth functions at p and therefore depend smoothly near the point p on the coordinates $x^1, \ldots, x^m$, i.e., there exist smooth functions $u^1(t^1, \ldots, t^m), \ldots, u^n(t^1, \ldots, t^m)$ of the real variables $t^1, \ldots, t^m$ such that

$$y^i \circ \Phi = u^1(x^1, \ldots, x^m), \quad i = 1, \ldots, n, \tag{4.76}$$

near the point p. Following the accepted custom, we will write these equalities in the form

$$y^i = u^i(x^1, \ldots, x^m), \quad i = 1, \ldots, n, \tag{4.77}$$

and we will say that they *represent a notation of the map* Φ *in local coordinates* $x^1, \ldots, x^m$ *and* $y^1, \ldots, y^n$.

From relations (4.76) and the definition of the map $d\Phi_p$, it easily follows that

Proposition 4.78.

$$d\Phi_p \left(\frac{\partial}{\partial x^i}\right)_p = \left(\frac{\partial(y^j \circ \Phi)}{\partial x^i}\right)_p \cdot \left(\frac{\partial}{\partial y^j}\right)_{\Phi(p)}, \qquad i = 1, \ldots, m, \quad j = 1, \ldots, n. \tag{4.79}$$

Proof. Indeed, let

$$g = u(y^1, \ldots, y^n)$$

be an arbitrary smooth function at the point $\Phi(p)$ on the manifold N. By definition,

$$\frac{\partial g}{\partial y^j} = \frac{\partial u}{\partial t^j}(y^1, \ldots, y^n).$$

On the other hand,

$$g \circ \Phi = u(y^1 \circ \Phi, \ldots, y^n \circ \Phi) = u(u^1(x^1, \ldots, x^m), \ldots, u^n(x^1, \ldots, x^m))$$

and therefore

$$\frac{\partial g \circ \Phi}{\partial y^j} = \frac{\partial u}{\partial t^j}(y^1 \circ \Phi, \ldots, y^n \circ \Phi) \cdot \frac{\partial u^j}{\partial t^i}(x^1, \ldots, x^m).$$

Therefore,

$$d\Phi_p \left(\frac{\partial}{\partial x^i}\right)_p g = \left(\frac{\partial(g \circ \Phi)}{\partial x^i}\right)_p = \left(\frac{\partial g}{\partial y^j}\right)_{\Phi(p)} \cdot \left(\frac{\partial(y^j \circ \Phi)}{\partial x^i}\right)_p,$$

since relations (4.76) yield

$$\frac{\partial u^j}{\partial t^i}(x^1, \ldots, x^m) = \frac{\partial(y^j \circ \Phi)}{\partial x^i}.$$

Thus, relations (4.79) are completely proven. □

In accordance with the abbreviated notation (4.77), we will, as a rule, write formulae (4.79) in the form

$$d\Phi_p \left(\frac{\partial}{\partial x^i}\right)_p = \left(\frac{\partial(y^j \circ \Phi)}{\partial x^i}\right)_p \cdot \left(\frac{\partial}{\partial y^j}\right)_{\Phi(p)}. \tag{4.80}$$

The matrix of order $n \times m$ with elements

$$\left(\frac{\partial y^j}{\partial x^i}\right)_p = \left(\frac{\partial (y^j \circ \Phi)}{\partial x^i}\right)_p, \qquad i = 1, \ldots, m, \quad j = 1, \ldots, n.$$

i.e., the matrix of linear relations (4.79), we will call the *matrix of the map* Φ *at the point* p (with respect to the local coordinates $x^1, \ldots, x^m$ and $y^1, \ldots, y^n$). For $m = n$, we will call the determinant of this matrix the *Jacobian* of the map Φ at the point p (with respect to the local coordinates $x^1, \ldots, x^m$ and $y^1, \ldots, y^n$). Obviously, it coincides with the Jacobian (in the sense of §4.2) of the map $\eta \circ \xi^{-1}$, where η and ξ are coordinate diffeomorphisms defining the local coordinates $y^1, \ldots, y^n$ and $x^1, \ldots, x^m$, respectively.

A smooth map

$$\Phi : M \to N$$

of an m-dimensional manifold M into an n-dimensional manifold N is called *regular at a point* $p \in M$ if the map

$$d\Phi_p : M_p \to N_{\Phi(p)}.$$

has the maximum possible rank, i.e., if it is monomorphic for $m \leq n$ and epimorphic for $m \geq n$ (and hence isomorphic for $m = n$). From the elementary properties of linear maps it immediately follows that

Proposition 4.81. *a map $\Phi : M \to N$ is regular at a point $p \in M$ if and only if the rank of its matrix at this point (calculated in arbitrary systems of local coordinates $x^1, \ldots, x^m$ and $y^1, \ldots, y^n$ at the points p and $\Phi(p)$ of the manifolds M and N) is equal to m for $m \leq n$ and n for $m \geq n$).*

By way of the theorem on the replacement of local coordinates, it follows that

Proposition 4.82. *for $m \leq n$, the map $\Phi : M \to N$ is regular at a point $p \in M$ if and only if, for any system of local coordinates $y^1, \ldots, y^n$ at the point $\Phi(p)$ of the manifold N, from the functions $y^1 \circ \Phi, \ldots, y^n \circ \Phi$, one can select m functions that constitute a system of local coordinates at the point p of the manifold M.*

Besides,

Proposition 4.83. *for $m = n$ the map $\Phi : M \to N$ is regular at the point $p \in M$ if and only if its Jacobian at this point is non-zero.*

The inverse function theorem from §4.2 states that for maps of domains of the same Euclidean space, being regular at a point is equivalent to being diffeomorphic at that point. Due to the local nature of this statement, it is obviously also valid for maps of any manifolds (of the same dimension). Thus,

Proposition 4.84. *A smooth map Φ of a manifold M into a manifold N of the same dimension is diffeomorphic at a point $p \in M$ if and only if it is regular at that point.*

Points p of the manifold M at which the map $\Phi : M \to N$ is not regular will be called *critical points* of this map. Maps that have no critical points, i.e., are regular at all points $p \in M$, will be called *regular* (*on* M).

According to the previous proposition,

Proposition 4.85. *A smooth map Φ of a manifold M into a manifold N of the same dimension is regular if and only if it is locally diffeomorphic.*

Let us now consider the product $M \times N$ of smooth manifolds M and N (see §4.3). Let

$$\pi : M \times N \to M, \quad \rho : M \times N \to N$$

be the natural projections of the product $M \times N$ onto manifolds M and N, respectively. Since these projections are smooth maps, for any point $(p, q) \in M \times N$ the followwng maps are defined

$$d\pi_{(p,q)} : (M \times N)_{(p,q)} \to M_p, \quad d\rho_{(p,q)} : (M \times N)_{(p,q)} \to N_q.$$

Let A be an arbitrary vector of the manifold $M \times N$ at the point (p, q) and let

$$B = d\pi_{(p,q)}(A), \quad C = d\rho_{(p,q)}(A).$$

It is clear that for any system of local coordinates $x^1, \ldots, x^m$ at a point p of the manifold M the following equalities hold:

$$A\widehat{x}^i = Bx^i, \quad i = 1, \ldots, m.$$

Similarly, for any system of local coordinates $y^1, \ldots, y^n$ at a point q of the manifold N the following equalities hold:

$$A\widehat{y}^j = By^j, \quad j = 1, \ldots, n.$$

This shows, firstly, that $A = 0$ when and only when $B = 0$ and $C = 0$, and, secondly, that for any vectors $B \in M_p$ and $C \in N_q$ there exists a vector $A \in (M \times N)_{(p,q)}$ such that $B = d\pi_{(p,q)}(A)$ and $C = d\rho_{(p,q)}(A)$. In other words,

Proposition 4.86. *the maps*

$$M_p \xleftarrow{d\pi_{(p,q)}} (M \times N)_{(p,q)} \xrightarrow{d\rho_{(p,q)}} N_q.$$

define a decomposition of the space $(M \times N)_{(p,q)}$ into a direct sum of the spaces M_p and N_q.

It is clear that this decomposition is consistent with the actions of vectors A on functions smooth at the point (p, q), i.e., for any such function f, the equality

$$Af = Bf + Cf,$$

holds where, say, Bf is the result of applying the vector $B = d\pi_{(p,q)}(A)$ to the function f, considered as a function of one point $p \in M$ (for a fixed point $q \in N$).

In the future, when it is convenient for us, we will denote the vector $A \in (M \times N)_{(p,q)}$, for which $B = d\pi_{(p,q)}(A)$ and $C = d\rho_{(p,q)}(A)$, by the symbol (B, C).

It is clear that similar direct sum decompositions also hold for products of any finite number of smooth manifolds.

4.7 Linear differential forms

Let, as above, M be an arbitrary smooth manifold of dimension m. By a *linear differential form* on a manifold M we mean a derivative $\mathcal{O}(M)$-linear map of the $\mathcal{O}(M)$-module $\mathcal{O}^1(M)$ into the algebra $\mathcal{O}(M)$ (considered as a $\mathcal{O}(M)$-module). Thus, each form ω associates with any vector field $X \in \mathcal{O}^1(M)$ some smooth function $\omega(X)$ on M, and for any smooth function f on M and any fields $X, X_1, X_2 \in \mathcal{O}^1(M)$ the following equalities hold:

$$\omega(fX) = f\omega(X),$$
$$\omega(X_1 + X_2) = \omega(X_1)) + \omega(X_2).$$

Obviously, the set $\mathcal{O}_1(M)$ of all linear differential forms on the manifold M is an $\mathcal{O}(M)$-module with respect to the operations

$$(f\omega)(X) = f\omega(X), \quad f \in \mathcal{O}(M), \omega \in \mathcal{O}_1(M), X \in \mathcal{O}^1(M),$$
$$(\omega_1 + \omega_2)(X) = \omega_1(X) + \omega_2(X), \quad \omega_1, \omega_2 \in \mathcal{O}_1(M), X \in \mathcal{O}^1(M).$$

Let W be an arbitrary open submanifold of M and let ω be an arbitrary linear differential form on M. Further, let Y be an arbitrary vector field on W. For any point $p \in W$, consider its neighbourhood V whose closure $\overline{V}$ is compact and contained in W. As we know, on the manifold M there exist vector fields X such that $X|_V = Y|_V$. Let X_1 and X_2 be two fields on M that have this property, and let h be a smooth function on M that is equal to one at the point p and zero outside the neighbourhood V. Then $h \cdot (X_1 - X_2) = 0$ on M and therefore

$$h(\omega(X_1) - \omega(X_2)) = 0.$$

It follows that at the point p the functions $\omega(X_1)$ and $\omega(X_2)$ coincide. Since the point p was an arbitrary point of the submanifold W, it was thus proved that the formula

$$\omega|_W(Y)(p) = \omega(X)(p),$$

where X is an arbitrary vector field on M that coincides with the field Y on V, identically defines on the submanifold W some (as is easily seen, smooth) function $\omega|_W(Y)$. Thus, we obtain some (obviously linear) map

$$\omega|_W : \mathcal{O}^1(W) \to \omega(W),$$

i.e., a linear differential form on the submanifold W. We will call this form the *restriction* of the form ω on the open submanifold W and, when there is no possibility of misunderstanding, we will denote it by the same symbol ω.

Just as for vector fields, the restriction map $\omega \to \omega|_W$ is, generally speaking, not epimorphic, but only locally epimorphic in the sense that,

Proposition 4.87. *Whatever the open set $V \subset W$ whose closure $\overline{V}$ is compact and contained in W, for any linear differential form $\theta \in \mathcal{O}_1(W)$ there exists a linear differential form $\omega \in \mathcal{O}_1(M)$ such that*

$$\omega|_V = \theta|_V.$$

This form ω is defined by the formula

$$\omega(X) = \begin{cases} g \cdot \theta(X|_V) & \text{on} \quad W, \\ 0 & \text{outside} \quad W, \end{cases}$$

where g is a smooth function on M that is equal to one on $\overline{V}$ and zero outside some open set whose closure is contained in W.

In the special case where W is a coordinate neighbourhood of U, one can easily describe all linear differential forms on U.

Let $X_1, \ldots, X_m$ be an arbitrary basis of the $\mathcal{O}(U)$-module $\mathcal{O}^1(U)$. Clearly, for any $i = 1, \ldots, m$, the relations

$$\omega^i(X_j) = \delta^i_j, \quad j = 1, \ldots, m$$

uniquely determine on U some linear differential form ω^i. The forms $\omega^1, \ldots, \omega^m$ are linearly independent and have the property that for any vector field $X \in \mathcal{O}^1(U)$, the following equality holds

$$X^i = \omega^i(X),$$

where X^i are the components of X in the basis $X_1, \ldots, X_m$. It follows directly from this that

Proposition 4.88. *the forms $\omega_1, \ldots, \omega_m$ form a basis for the module $\mathcal{O}_1(U)$, and*

$$\omega = \omega(X_i)\omega i$$

for any form $\omega \in \mathcal{O}_1(U)$.

We will call the functions $\omega(X_i)$ the *components* of the form ω in the basis $X_1, \ldots, X_m$ (or in the basis $\omega_1, \ldots, \omega_m$).

If the basis $X_1, \ldots, X_m$ is holonomic, i.e., if

$$X_1 = \frac{\partial}{\partial x^1}, \ \ldots, \ X_m = \frac{\partial}{\partial x^m},$$

where $x^1, \ldots, x^m$ is some system of local coordinates in the neighbourhood U, then we will also call the basis a *holonomic* basis. In this case, the forms $\omega^1, \ldots, \omega^m$ we will, as a rule, denote by the symbols

$$dx^1, \ \ldots, \ dx^m.$$

The components of the form $\omega\left(\frac{\partial}{\partial x^i}\right)$ in the basis $dx^1, \ldots, dx^m$ we will sometimes call its *coefficients* (in the system of local coordinates $x^1, \ldots, x^m$).

We obtain an important example of a linear differential form on a manifold M by setting for some smooth function f on M and any field $X \in \mathcal{O}^1(M)$

$$df(X) = X(f).$$

It is clear that this formula defines a certain linear differential form df on M. We will call the form df the *differential* of the function f.

By definition, for any smooth function f on the manifold M (or at least on the coordinate neighbourhood U), the following equality holds

$$df\left(\frac{\partial}{\partial x^i}\right) = \frac{\partial f}{\partial x^i}, \quad i = 1, \ \ldots, \ m,$$

so that, as in elementary analysis,

$$df = \frac{\partial f}{\partial x^i} dx^i.$$

In particular, we see that the forms $dx^1, \ \ldots, \ dx^m$ of the holonomic basis are nothing more than differentials of the local coordinates $x^1, \ \ldots, \ x^m$.

Linear functionals defined on the space $M(p) = \mathcal{O}^1(p)$, i.e., linear maps of this space into the field $\mathbb{R}$, we will call *covectors* of the manifold M at the point p. They form a m-dimensional linear space M_p^*, dual to the space M_p. We will also denote the space M_p^* by the symbol $\mathcal{O}_1(p)$.

It is easy to see that for any form $\omega \in \mathcal{O}_1(M)$ and any field $X \in \mathcal{O}^1(M)$, the equality $X_p = 0$ implies the equality $\omega(X)(p) = 0$. Indeed, in the coordinate neighbourhood U of the point p, the following relation holds

$$\omega(X) = \omega\left(\frac{\partial}{\partial x^i}\right) \cdot Xx^i,$$

and therefore

$$\omega(X)(p) = \omega\left(\frac{\partial}{\partial x^i}\right)(p) \cdot X_p x^i = 0.$$

It follows that, by setting for any vector $A \in M_p$

$$\omega_p(A) = \omega(X)(p),$$

where X is an arbitrary vector field on M for which $X_p = A$, we uniquely determine on M_p some covector ω_p.

In particular, the differentials $dx^1, \ \ldots, \ dx^m$ of the local coordinates $x^1, \ \ldots, \ x^m$ define covectors

$$(dx^1)_p, \ \ldots, \ (dx^m)_p.$$

These covectors form a basis for the space M_p^* (there are m of them, and they are linearly independent). Any covector $a \in M_p^*$ can be decomposed in this basis according to the formula

$$\alpha = \alpha\left(\left(\frac{\partial}{\partial x^i}\right)_p\right) \cdot (dx^i)_p.$$

From this, in particular, it follows directly that

Proposition 4.89. *for any covector* $\alpha \in \mathcal{O}_1(p)$ *there exists on the manifold* M *a linear differential form* ω *such that* $\omega_p = \alpha$.

Proof. Indeed, according to the local epimorphism property of the restriction map, it suffices to construct the form ω on some coordinate neighbourhood U of the point p. On the other hand, the formula

$$\omega = a_i dx^i,$$

where $a_i = \alpha\left(\left(\frac{\partial}{\partial x^i}\right)_p\right)$, obviously defines a form ω on a neighbourhood U for which $\omega_p = \alpha$. □

Analogously to the set of all vectors, the set M_*^* of all covectors of the manifold M at all its points $p \in M$ can be defined as a $2m$-dimensional smooth manifold, after which the concept of a continuous (or smooth) covector field ω on the manifold M is naturally defined as a continuous (respectively, smooth) map $\omega : M \to M_*^*$ such that $\omega(p) \in M_p^*$ for any point $p \in M$. Moreover, it can be easily shown (see the corresponding arguments for vector fields) that a covector field ω is smooth if and only if for any vector field $X \in \mathcal{O}^1(M)$ the numerical function $\omega(p)(X_p)$ of a point p of a manifold M is a smooth function on M.

However, to save space, we prefer not to introduce the manifolds M_*^* explicitly here, but instead adopt the latter property of smooth covector fields as their definition. Thus, by a *smooth covector field* on a manifold M we will define an arbitrary function $\omega X \in \mathcal{O}^1(M)$, the numerical function $\omega(p)(X_p)$ of the point $p \in M$ is a smooth function on M.

It is clear that for any differential form $\omega \in \mathcal{O}_1(M)$, the formula

$$\omega(p) = \omega_p, \quad p \in M,$$

defines a smooth covector field on the manifold M. Conversely, every smooth covector field $\omega \in \mathcal{O}_1(M)$ uniquely determines, according to the formula

$$\omega(X)(p) = \omega(p)(X_p), \quad X \in \mathcal{O}^1(M), \in M,$$

a linear differential form $\omega \in \mathcal{O}_1(M)$. Thus,

Proposition 4.90. *the concept of a linear differential form essentially coincides with the concept of a smooth covector field.*

Remark 4.91. Any smooth function f on M defines a covector $(df)_p$ at each point $p \in M$. On the other hand, considering the function f as a smooth map $M \to \mathbb{R}$, we can speak of its differential $df_p : M_p \to \mathbb{R}_{f_p}$, where $\mathbb{R}_{f_p}$ is the tangent space of $\mathbb{R}$ at the point $f(p)$. The space $\mathbb{R}_{f_p}$ is one-dimensional and is generated by the vector $\left(\frac{\partial}{\partial t}\right)_{f(p)}$. It is clear that if we identify each vector $a\left(\frac{\partial}{\partial t}\right)_{f(p)} \in \mathbb{R}_{f_p}$ with a number a, then the differential df_p coincides with the covector $(df)_p$. In this sense, the definition of the differential of a function is consistent with the definition of the differential of a smooth map.

In conclusion, let us consider an arbitrary smooth map Φ of some manifold N into a manifold M. Let ω be an arbitrary linear differential form on the manifold M. At each point $p \in N$, we can define a covector $(\Phi^*\omega)_p$ by setting

$$(\Phi^*\omega)_p(A) = \omega_p(d\Phi_p(A))$$

for any vector $A \in N_p$. Thus, on the manifold N, we obtain a (obviously smooth) covector field $\Phi^*(\omega)$, i.e., a linear differential form. It is easy to verify that the map

$$\Phi^* : \mathcal{O}^1(M) \to \mathcal{O}^1(N)$$

defined in this way is a homomorphism of $\mathcal{O}(M)$-modules (if the $\mathcal{O}(N)$-module $\mathcal{O}^1(N)$ is considered as a $\mathcal{O}(N)$-module via the map $\Phi^* : \mathcal{O}(M) \to \mathcal{O}(N)$, which associates with each function $f \in \mathcal{O}(M)$ the function $\Phi^* f = \circ\Phi$).

If Φ is an inclusion of an open submanifold W into a manifold M, then the map Φ^* is nothing more than a restriction map $\mathcal{O}_1(M) \to \mathcal{O}_1(W)$.

4.8 Tensors and tensor fields

Let r and s be arbitrary non-negative integers. By a *tensor field of type* (r, s) on a manifold M we mean an arbitrary function

$$T(\omega^1, dotsc, \ \omega^r; X_1, \ \dots, \ X_s)$$

of $r+s$ arguments, the first r of which are linear differential forms $\omega^1, \ \dots, \ \omega^r \in \mathcal{O}_1(M)$, and the last s are vector fields $X_1, \ \dots, \ X_s \in \mathcal{O}^1(M)$.

We will denote the set of all tensor fields of type (r, s) on a manifold M by the symbol $\mathcal{O}^r_s(M)$. Clearly, this set is naturally defined as an $\mathcal{O}(M)$-module.

By definition, tensor fields of type $(0, 0)$ are nothing more than smooth functions on M, so that

$$\mathcal{O}^0_0(M) = \mathcal{O}(M),$$

and tensor fields of type $(0, 1)$ are nothing more than linear differential forms, so that

$$\mathcal{O}^0_1(M) = \mathcal{O}_1(M).$$

Moreover, any vector field $X \in \mathcal{O}^1(M)$ can be considered as a tensor field of type $(1, 0)$, i.e., as a $\mathcal{O}(M)$-linear map of the $\mathcal{O}(M)$-module $\mathcal{O}_1(M)$ into the algebra $\mathcal{O}(M)$, defining its value $X(\omega)$ on an arbitrary form $\omega \in \mathcal{O}_1(M)$ by the formula

$$X(\omega) = \omega(X).$$

Since $\omega(X) = 0$ for all forms $\omega \in \mathcal{O}_1(M)$ only when $X = 0$, the module $\mathcal{O}^1(M)$ is embedded in the module $\mathcal{O}^1_0(M)$. It turns out that these modules actually coincide:

$$\mathcal{O}^1_0(M) = \mathcal{O}^1(M), \tag{4.92}$$

i.e.,

Proposition 4.93. *vector fields exhaust all tensor fields of type* $(0, 1)$.

We will prove this important statement a little later, after we have studied in more detail the structure of arbitrary tensor fields.

Let W be an arbitrary open submanifold of M, and let T be an arbitrary tensor field of type (r, s) on M. It is easy to see that

Proposition 4.94. *if at least one of the arguments* $\omega, \ldots, \omega^r, X_1, \ldots, X_s$ *of* T *is zero on* W *(i.e., its restriction to* W *is zero), then*

$$T(\omega^1, \ldots, \omega^r; X_1, \ldots, X_s) = 0 \quad on \quad W.$$

Proof. Indeed, let p be an arbitrary point of W and let V be some neighbourhood of it whose closure $\overline{V}$ is compact and contained in W. Further, let f be a smooth function on M equal to one at p and zero outside V. Assuming, for definiteness, that $\omega^1 = 0$ on W, we obtain that $f\omega^1 = 0$ on M, and therefore

$$T(f\omega^1, \omega^2, \ldots, omega^r; X_1, \ldots, X_s) = 0.$$

But

$$T(f\omega^1, \omega^2, \ldots, \omega^r; X_1, \ldots, X_s) = fT(\omega^1, \ldots, \omega^r; X_1, \ldots, X_s).$$

so that

$$T(\omega^1, \ldots, \omega^r; X_1, \ldots, X_s)(p) = f(p)T(\omega^1, \ldots, \omega^r; X_1, \ldots, X_s)(p) = 0.$$

since $f(p) = 1$. Thus,

$$T(\omega^1, \ldots, \omega^r; X_1, \ldots, X_s) = 0 \quad \text{on} \quad W.$$

□

Now let $\theta^1, \ldots, \theta^r$ be arbitrary linear differential forms, and $Y_1, \ldots, Y_s$ be arbitrary vector fields on an open submanifold W. For any point $p \in W$, consider on M the forms $\omega^1, \ldots, \omega^r$ and the fields $X_1, \ldots, X_s$ that coincide in some neighbourhood of the point p with the forms $\theta^1, \ldots, \theta^r$ and the fields $Y_1, \ldots, Y_s$, respectively. From the statement just proved, it immediately follows (see the similar arguments above for vector fields and differential forms) that the formula

$$T|_W(\theta^1, \ldots, \theta^r; Y_1, \ldots, Y_s)(p) = T(\omega^1, \ldots, \omega^r; X_1, \ldots, X_s)(p)$$

uniquely defines on W some smooth function

$$T|_W(\theta^1, \ldots, \theta^r; Y_1, \ldots, Y_s).$$

It is clear that this function depends linearly on the forms $\theta^1, \ldots, \theta^r$ and the fields $Y_1, \ldots, Y_s$, i.e., it is a tensor field of type (r, s) on the manifold W. We will call the field $T|_W$ the *restriction* of the field T on the open submanifold W

and, when there is no possibility misunderstandings, we will denote it by the former symbol T.

Obviously, the restriction map $\mathcal{O}_s^r(M) \to \mathcal{O}_s^r(W)$, which associates with each field $T \in \mathcal{O}_s^r(M)$ the field $T|_W \mathcal{O}_s^r(W)$, is a $\mathcal{O}(M)$-homomorphism of the $\mathcal{O}(M)$-module $\mathcal{O}_s^r(M)$ into the $\mathcal{O}(M)$-module $\mathcal{O}_s^r(W)$ (considered, via the restriction map $\mathcal{O}(M) \to \mathcal{O}(W)$, as an $\mathcal{O}(M)$-module). Moreover, similar to the case of vector zeros and linear differential forms, this homomorphism obviously has the property of local epimorphism, i.e.,

Proposition 4.95. *for any open set $V \subset W$ whose closure $\overline{V}$ is compact and contained in W, for any field $S \in \mathcal{O}_s^r(W)$ there exists a field $T \in \mathcal{O}_s^r(M)$, such that*

$$S|_V = T|_V.$$

Let us consider, in particular, the restriction $T|_U$ of the field $T \in \mathcal{O}_s^r(M)$ to some coordinate neighbourhood $U \subset M$. Let $Y_1, \ldots, Y_m$ be an arbitrary basis of the module $\mathcal{O}^1(U)$, and $\theta^1, \ldots, \theta^m$ be a basis of the module $\mathcal{O}_1(U)$ defined by the conditions $\theta^i(Y_j) = \delta^i_j$; $i, j = 1, \ldots, m$. Then, for any forms $\omega^1, \ldots, \omega^r \in \mathcal{O}_1(U)$ and any fields $X_1, \ldots, X_s \in \mathcal{O}^1(U)$, the following relation holds:

$$T(\omega^1, \ldots, \omega^r; X_1, \ldots, X_s) = T^{k_1 \cdots k_r}_{\ell_1 \cdots \ell_s} \omega^1_{k_1} \cdots \omega^r_{k_r} X_1^{\ell_1} \cdots X_s^{\ell_s} \quad \text{on} \quad U, \tag{4.96}$$

where

$$T^{k_1 \cdots k_r}_{\ell_1 \cdots \ell_s} = T(\theta^{k_1}, \ldots, \theta^{k_r}; Y_{\ell_1}, \ldots, Y_{\ell_s}), \tag{4.97}$$

and $\omega^i_k = \omega^i(Y_k)$, and $X^\ell_j = \theta^\ell(X_j)$ are the components of the forms ω^i and the fields X_j in the basis $Y_1, \ldots, Y_m$. We will call functions (4.97) the *components* of the tensor field T on a neighbourhood U in the basis $Y_1, \ldots, Y_m$. They completely determine the field T on U.

In particular, if the field T is of type $(1, 0)$, its components are the functions $T^k = T(\theta^k)$. We define a vector field X_k on the neighbourhood U by setting

$$X_U = T^k Y_k,$$

where, as above, $Y_1, \ldots, Y_m$ is an arbitrary basis of the $\mathcal{O}(U)$-module $\mathcal{O}^1(U)$ (relative to which the components T^k are calculated). Since for any form $\omega \in \mathcal{O}_1(U)$ the following equality

$$T(\omega) = \omega(Y_k) T(\theta^k) = \omega(T^k Y_k) = \omega(X_U)$$

holds on U, then on the neighbourhood U the tensor field T coincides with the vector field X_U. In particular, this implies that the field X_U is uniquely defined (independent of the choice of the basis $Y_1, \ldots, Y_m$). Therefore, setting for any point $p \in M$

$$X_p = (X_U)_p,$$

where U is an arbitrary coordinate neighbourhood of the point p, we obtain a uniquely defined (obviously smooth) vector field X on the manifold M. Since

$T(\omega) = \omega(X_U)$ on U for any form $\omega \in \mathcal{O}_1(M)$, then $T(\omega)(p) = \omega(X_U)(p) = \omega_p((X_U)_p) = \omega_p(X_p) = \omega(X)(p)$, i.e.

$$T(\omega) = \omega(X) \quad \text{on} \quad M.$$

In other words, $T = X$. Thus, relation (4.92) is completely proven.

Let p be an arbitrary point of a smooth manifold M. By a *tensor of type* (r, s) at the point p we mean an arbitrary functional linear in each argument

$$L(\alpha^1, \ldots, \alpha^r; A_1, \ldots, A_s),$$

whose first r arguments are covectors $\alpha^1, \ldots, \alpha^r \in M_p^*$, and the last s are vectors $A_1, \ldots, A_s \in M_p$. All tensors of type (r, s) at the point p obviously form a linear space $\mathcal{O}_s^r(p)$. It is clear that

$$\mathcal{O}_0^1(p) = \mathcal{O}^1(p), \quad \mathcal{O}_1^0(p) = \mathcal{O}_1(p).$$

On the other hand, $\mathcal{O}_0^0(p) = \mathbb{R}$ (and not $\mathcal{O}(p)$).

Let $B_1, \ldots, B_m$ be an arbitrary basis of the space M_p and let $\beta^1, \ldots, \beta m$ be a basis of the space M_p^* defined by the conditions $\beta^i(B_j) = \delta_j^i$, $j = 1, \ldots, m$. Then, for any tensor L of type (r, s) at the point p, the following equality holds

$$L(\alpha^1, \ldots, \alpha^r; a_1, \ldots, A_s) = L_{\ell_1 \cdots \ell_s}^{k_1 \cdots k_r} \alpha_{k_1}^1 \cdots \alpha_{k_r}^r A_1^{\ell_1} \cdots A_s^{\ell_s}, \tag{4.98}$$

where $\alpha_k^i = \alpha^i(B_k)$ and $A_j^{\ell}) = \beta^{\ell}(A_j)$ are the components of the covectors α^i and the vectors A_j in the basis $B_1, \ldots, B_m$, and

$$L_{\ell_1 \cdots \ell_s}^{k_1 \cdots k_r} = L(\beta^{k_1}, \ldots, \beta^{k_r}; B_{\ell_1}, \ldots, B_{\ell_s}). \tag{4.99}$$

We will call the numbers (4.99) the *components* of the tensor L in the basis $B_1, \ldots, B_m$. These numbers do not, generally speaking, satisfy any relations, i.e., for any set m^{r+s} of numbers $L_{\ell_1 \cdots \ell_s}^{k_1 \cdots k_r}$, formula (4.98) defines a tensor $L \in \mathcal{O}_s^r(p)$ with components $L_{\ell_1 \cdots \ell_s}^{k_1 \cdots k_r}$.

Just as for vectors, the set of all tensors of a given type (r, s) at all points of a manifold M can be naturally defined as a smooth manifold, after which one can define the concept of a continuous (or smooth) tensor field (r, s) on a manifold M as a continuous (respectively, smooth) map T of a manifold M into the manifold of all tensors such that $T(p) \in \mathcal{O}_s^r(p)$ for any point $p \in M$.
Without dwelling on this, as in the case of forms, we prefer to independently define smooth tensor fields T of type (r, s) on a manifold M as functions T that associate with each point $p \in M$ some tensor $T(p)$ of type (r, s) at this point such that for any forms $\omega^1, \ldots, \omega^r \in \mathcal{O}_1(M)$ and any vector fields $X_1, \ldots, X_s \in \mathcal{O}^1(M)$ the numerical function

$$f(p) = T(p)((\omega^1)_p, \ldots, (\omega^r)_p; (X_1)_p, \ldots, (X_s)_p)$$

at the point $p \in M$ is a smooth function on M.

For any tensor field $T \in \mathcal{O}^r_s(M)$, it immediately follows from formula (4.96) that if at least one of its arguments $\omega^1, \dots, \omega^r$ and $X_1, \dots, X_s$ is equal to zero at the point p, then

$$T(\omega^1, \dots, \omega^r; X_1, \dots, X_s) = 0.$$

Therefore, the formula

$$T_p(\alpha^1, \dots, \alpha^r; A_1, \dots, A_s) = T(\omega^1, \dots, \omega^r; X_1, \dots, X_s)(p),$$

where the forms $\omega^1, \dots, \omega^r$ and the fields $X_1, \dots, X_s$ are chosen such that $\omega^i_p = \alpha^i$, $i = 1, \dots, r$, and $(X_j)_p = A_j$, $j = 1, \dots, s$, uniquely determines a tensor T_p of type (r, s) at the point p. Clearly, the resulting function $T(p) = T_p$ is a smooth tensor field in the sense of the previous definition. Conversely, any smooth tensor field $T(p) \in \mathcal{O}^r_s(p)$ is defined by the formula

$$T(\omega^1, \dots, \omega^r; X_1, \dots, X_s) = T(p)((\omega^1)_p, \dots, (\omega^r)_p; (X_1)_p, \dots, (X_s)_p),$$

where $\omega^1, \dots, \omega^r \in \mathcal{O}_1(M)$ and $X_1, \dots, X_s \in \mathcal{O}^1(M)$, is some field $T \in \mathcal{O}^r_s(M)$. Thus,

Proposition 4.100. *the concept of a tensor field, as defined at the beginning of this section, essentially coincides with the concept of a smooth tensor field defined above.*

This assertion has many important consequences. For example, it immediately implies the equality (4.92) already proven above. Indeed, any tensor of type $(1, 0)$ at a point p is, by definition, a linear functional on the space M^*_p, and therefore, by the inherent property of reflexivity of finite-dimensional linear spaces, each such tensor can be considered as a vector at a point p.

Furthermore, any tensor $L = L(\alpha; A)$ of type $(1, 1)$ with a fixed argument A is a tensor of type $(1, 0)$, i.e., a vector. Therefore, each such tensor can be considered as a linear transformation of the space M_p. With this in mind, we will call an arbitrary function T that associates with each point $p \in M$ some linear transformation $T_p : M_p \to M_p$ the *field of linear transformations* on the manifold M. We will call such a field *smooth* if, for any vector field $X \in \mathcal{O}^1(M)$, the field

$$T(X)_p = T_p(X_p), \quad p \in M,$$

is a smooth vector field on M. From what has been said above it follows directly that the concepts of a tensor field of type $(1, 1)$ and a smooth field of linear transformations essentially coincide. On the other hand, every smooth field T of linear transformations defines by the formula

$$T(X)_p = T_p(X_p), \quad p \in M, X \in \mathcal{O}^1(M),$$

some $\mathcal{O}(M)$-linear map of the $\mathcal{O}(M)$-module $\mathcal{O}^1(M)$ into itself. Conversely, any such map $T : \mathcal{O}^1(M) \to \mathcal{O}^1(M)$ is defined by the formula

$$T_p(A) = T(X)_p, \quad A \in M_p,$$

where X is an arbitrary field on M for which $X_p = A$, is some smooth field T of linear transformations.

Thus,

Proposition 4.101. *The following concepts are essentially the same:*

1) a tensor field of type $(1,1)$ *on a manifold* M*;*

2) a smooth field of linear transformations on M*;*

3) an $\mathcal{O}(M)$*-linear map of the* $\mathcal{O}(M)$*-module* $\mathcal{O}^1(M)$ *into itself.*

In this case, the tensor field $T \in T\mathcal{O}_1^1(M)$ is defined by the linear map $T : \mathcal{O}^1(M) \to \mathcal{O}^1(M)$ by the formula

$$T(\omega; X) = \omega(T(X)), \quad \omega \in \mathcal{O}_1(M), X \in \mathcal{O}^1(M),$$

and the field of linear transformations $T_p : M_p \to M_p$ is defined by the formula

$$T_p(A) = T(X)_p, \quad A \in M_p,$$

where X is an arbitrary vector field on the manifold M for which $X_p = A$.

Remark 4.102. By fixing the argument α in the tensor $L(\alpha; A)$ of type $(1,1)$ at the point p, we can consider it as a linear transformation of the space M_p^*. This linear transformation is conjugate, in the sense of the theory of linear spaces, to the linear transformation of the space M_p considered above.

After tensors of type $(1,1)$, the next most complex ones are tensors of type $(1,2)$. Reasoning as above, we easily see that

Proposition 4.103. *the following concepts are essentially the same:*

1) a tensor field $T(\omega; X; Y)$ *of type* $(1,2)$ *on a manifold* M*;*

2) a smooth field of bilinear functions $T_p(A,B)$*,* $A, B \in M_p$*, taking values in the space* M_p*;*

3) an $\mathcal{O}(M)$*-linear function* $T(X,Y)$*,* $X, Y \in \mathcal{O}^1(M)$*, in both arguments, taking values in the* $\mathcal{O}(M)$*-module* $\mathcal{O}^1(M)$*.*

In this case, the following relations hold:

$$T(\omega; X, Y) = \omega(T(X,Y)), \quad \omega \in \mathcal{O}_1(M), X, Y \in \mathcal{O}^1(M),$$
$$T_p(X_p, Y_p) = T(X,Y)_p, \quad p \in M, X, Y \in \mathcal{O}^1(M).$$

Similarly,

Proposition 4.104. *the following concepts are essentially the same:*

1) a tensor field $T(\omega; X; Y; Z)$ *of type* $(1,3)$ *on a manifold* M*;*

2) a smooth field of bilinear functions $T_p(A,B)$, $A,B \in M_p$, taking values in linear transformations of the space M_p;

3) an $\mathcal{O}(M)$-linear function $T(X,Y)$, $X,Y \in \mathcal{O}^1(M)$, in both arguments, taking values in the $\mathcal{O}(M)$-linear maps of the $\mathcal{O}(M)$-module $\mathcal{O}^1(M)$ into itself.

In this case, the following relations hold:

$$T(\omega; X,Y,Z) = \omega(T(X,Y)Z), \quad \omega \in \mathcal{O}_1(M),\ X,Y,Z \in \mathcal{O}^1(M),$$
$$T_p(X_p,Y_p)Z_p = (T(X,Y)Z_p), \quad p \in M,\ X,Y,Z \in \mathcal{O}^1(M).$$

4.9 Operations on tensors and tensor fields

Let L and K be arbitrary tensors at a point p of a smooth manifold M, having types (r,s) and (u,v), respectively. It is clear that the formula

$$(L \otimes K)(\alpha^1,\ \ldots,\ \alpha^{r+u}; A_1,\ \ldots,\ A_{s+v}) =$$
$$L(\alpha^1,\ \ldots,\ \alpha^r;\ A_1,\ \ldots,\ A_s)\cdot K(\alpha^{r+1},\ \ldots,\ \alpha^{r+u};\ A_{s+1},\ \ldots,\ A_{s+v}),$$

where $\alpha^1,\ \ldots,\ \alpha^{r+u} \in \mathcal{O}_1(p)$ and $A_1,\ \ldots,\ A_{s+v} \in \mathcal{O}^1(p)$ defines a tensor $L \otimes K$ of the type $(r+u, s+v)$. We will call it the *product* of the tensors L and K. On tensors of type $(0,0)$, i.e., constant numbers, it coincides with the ordinary product. In an arbitrary basis $B_1,\ \ldots,\ B_m$ of the space M_p, its components are expressed, as is easy to see, by the formula

$$(L \otimes K)^{k_1,\ \ldots\ k_{r+u}}_{\ell_1,\ \ldots,\ \ell_{s+v}} = L^{k_1,\ \ldots\ k_r}_{\ell_1,\ \ldots,\ \ell_s} K^{k_{r+1},\ \ldots\ k_{r+u}}_{\ell_{s+1},\ \ldots,\ \ell_{s+v}}.$$

Moreover, it is easy to verify that

Proposition 4.105. *with respect to the multiplication $\otimes$, the direct sum*

$$\mathcal{O}^*(p) = \sum_{r,s=0}^{\infty} \mathcal{O}^r_s(p)$$

of linear spaces $\mathcal{O}^r_s(p)$ represents an associative algebra with unity over the field $\mathbb{R}$.

Let $B_1,\ \ldots,\ B_m$ be an arbitrary basis of the space M_p and let $\beta^1,\ \ldots,\ \beta^m$ be a basis of the space M_p^* such that $\beta^i(B_j) = \delta^i_j$. Let, further, L be an arbitrary tensor of type (r,s) at the point p and let $L^{k_1,\ \ldots\ k_r}_{\ell_1,\ \ldots,\ \ell_s}$ be its components in the basis $B_1,\ \ldots,\ B_m$. Consider the tensor

$$L^{k_1,\ \ldots\ k_r}_{\ell_1,\ \ldots, ell_s} \beta^{\ell_1} \otimes \cdots \otimes \beta^{\ell_s} B_{k_1} \otimes \cdots \otimes B_{k_r}.$$

It is easy to see that the components of this tensor are precisely the numbers $L^{k_1,\ \ldots\ k_r}_{\ell_1,\ \ldots,\ \ell_s}$. Therefore,

$$L = L^{k_1,\ \ldots\ k_r}_{\ell_1,\ \ldots,\ \ell_s} \beta^{\ell_1} \otimes \cdots \otimes \beta^{\ell_s} B_{k_1} \otimes \cdots \otimes B_{k_r}.$$

Thus, it is proved that

Proposition 4.106. *the algebra $\mathcal{O}^*(p)$ is generated over the field $\mathbb{R}$ by the vectors $B_1, \dots, B_m$ and the covectors $\beta^1, \dots, \beta^m$.*

Let i and j be arbitrary integers satisfying the inequalities $1 \le i \le s$ and $1 \le j \le r$, respectively (here we assume that $s > 0$ and $r > 0$). By fixing, in some tensor $L(\alpha^1, \dots, \alpha^r; A_1, \dots, A_s)$ of type (r, a), all arguments $\alpha^1, \dots, \alpha^r; A_1, \dots, A_s$ except for the i-th vector argument A_i and the j-th covector argument α^j, we obtain from the tensor L some tensor of type $(1,1)$, i.e., a linear transformation of the space M_p. This transformation depends, of course, on the covectors $\alpha^1, \dots, \widehat{\alpha}^j, \dots, \alpha^r$ and the vectors $A_1, \dots, \widehat{A}^i, \dots, A_s$ (the sign "$\wedge$" indicates that the corresponding covector - or vector - must be omitted). Let

$$(C^i_j L)(\alpha^1, \dots, \widehat{\alpha}^j, \dots, \alpha^r; A_1, \dots, \widehat{A}^i, \dots, A_s)$$

be the trace of this linear transformation. It is a linear function of the $r-1$ covectors $\alpha^1, \dots, \widehat{\alpha}^j, \dots, \alpha^r$ and the $s-1$ vectors $A_1, \dots, \widehat{A}^i, \dots, A_s$, i.e., it is a tensor of type $(r-1, s-1)$. We will say that this tensor is obtained from the tensor L *by contraction over the i-th upper and j-th lower indices.* In each basis of the space M_p, its components $(C^i_j L)^{k_1, \dots, k_r}_{\ell_1, \dots, \ell_s}$ are expressed in terms of the components $L^{k_1, \dots, k_r}_{\ell_1, \dots, \ell_s}$ of the tensor L according to the formula

$$(C^i_j L)^{k_1, \dots, k_{r-1}}_{\ell_1, \dots, \ell_{s-1}} = L^{k_1, \dots, k_{i-1}, \dots, hk_i, \dots, k_{r-1}}_{\ell_1, \dots, \ell^{j-1}, \dots, h\ell_j, \dots, \ell_{s-1}}.$$

The constructed map

$$C^i_j : \mathcal{O}^r_s(p) \to \mathcal{O}^{r-1}_{s-1}(p),$$

is obviously linear. We will call it the *contraction map.*

All the above constructions are immediately transferred from tensors to arbitrary tensor fields. For example, the product $T \otimes S$ of tensor fields $T \in \mathcal{O}^r_s(M)$ and $S \in \mathcal{O}^u_v(M)$ can be defined either by the formula

$$\begin{aligned}
&(T \otimes S)(\omega^1, \dots, \omega^{r+u}; X_1, \dots, X_{s+v}) = \\
&T(\omega^1, \dots, \omega^r; X_1, \dots, X_s) S(\omega^{r+1}, \dots, \omega^{r+u}; X_{s+1}, \dots, X_{s+v}), \\
&\omega^1, \dots, \omega^{r+u} \in \mathcal{O}_1(M);\ X_1, \dots, X_{s+v} \in \mathcal{O}^1(M),
\end{aligned}$$

or by the formula

$$(T \otimes S)_p = T_p \otimes S_p,$$

where p is an arbitrary point of the manifold M. From the corresponding statement for tensors, it follows immediately that

Proposition 4.107. *with respect to the multiplication $\otimes$, the direct sum*

$$\mathcal{O}^*(M) = \sum_{r,s=0}^{\infty} \mathcal{O}^r_s(M)$$

of $\mathcal{O}(M)$-modules, $\mathcal{O}^r_s(M)$, is an associative algebra with unity over the ring $\mathcal{O}(M)$.

Consider, in particular, the algebra $\mathcal{O}^*(U)$, where U is an arbitrary coordinate neighbourhood in the manifold M. Let $x^1, \ldots, x^m$ be an arbitrary system of local coordinates on the neighbourhood U. By comparing the components, we can easily obtain (see the similar arguments above for the algebra $\mathcal{O}^*(p)$) that for any tensor field $T \in \mathcal{O}^r_s(U))$ the following formula holds

$$T = T^{k_1, \ldots, k_{r-1}}_{\ell_1, \ldots, \ell_{s-1}} dx^{\ell_1} \otimes \cdots \otimes dx^{\ell_s} \otimes \frac{\partial}{\partial x^{k_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{k_r}},$$

where $T^{k_1, \ldots, k_{r-1}}_{\ell_1, \ldots, \ell_{s-1}}$ are the components of the field T in the basis $\frac{\partial}{\partial x^1} \otimes \cdots \otimes \frac{\partial}{\partial x^m}$. Therefore,

Proposition 4.108. *the algebra $\mathcal{O}^*(U)$ is generated over the ring $\mathcal{O}(U)$ by the forms $dx^1, \ldots, dx^m$ and the fields $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^m}$.*

The contraction operation C^i_j for tensor fields is most easily defined by the formula

$$(C^i_j T)_p = C^i_j T_p,$$

where p is an arbitrary point of the manifold M. Clearly, the resulting contraction map

$$C^i_j : \mathcal{O}^r_s(M) \to \mathcal{O}^{r-1}_{s-1}(M), \quad 1 \le i \le s,\ 1 \le j \le r,$$

is a homomorphism of $\mathcal{O}(M)$-modules.

Of particular interest is the case $r = s$. In this case, by applying r folds of the folding map to the field T, we obtain a field of type $(0, 0)$, i.e., a smooth function on M. We will say that this function is obtained from the field T as a result of a *complete fold*. Clearly, in general, the result of a complete fold depends on which pairs of indices were used to perform the r partial folds.

Consider, in particular, for an arbitrary tensor field T of type (r, s), the field

$$S = T \otimes \omega^1 \otimes \cdots \otimes \omega^r \otimes X_1 \otimes \cdots \otimes X_s$$

of type $(r + s, r + s)$, where $\omega^1, \ldots, \omega^r \in \mathcal{O}_1(M)$ and $X_1, \ldots, X_s \in \mathcal{O}^1(M)$.

Let $p \in M$. Denoting by ω^i_j and X^k_ℓ the components of the covectors $(\omega^i)_p$ and the vectors $(X_\ell)_p$ in some basis of the space M_p, we obtain that the components of the tensor S_p in this basis are expressed by the formula

$$(S_p)^{k_1, \ldots, k_{r+s}}_{\ell_1, \ldots, \ell_{r+s}} = (T_p)^{k_1, \ldots, k_r}_{\ell_1, \ldots, \ell_s} \omega^1_{\ell_{s+1}}, \ldots, \omega^r_{\ell_{r+s}} X^{k_{r+1}}_1, \ldots, X^{k_{r+s}}_s,$$

where $(T_p)^{k_1, \ldots, k_r}_{\ell_1, \ldots, \ell_s}$ are the components of the tensor T_p. Therefore, contracting the tensor S_p over the indices $(1, s + 1), (2, s + 2), \ldots, (r, s + r), (r + 1, 1), (r + 2, 2), \ldots, (r + s, s)$, we obtain the number

$$(T_p)^{k_1, \ldots, k_r}_{\ell_1, \ldots, \ell_s} \omega^1_{\ell_{s+1}}, \ldots, \omega^r_{\ell_{r+s}} X^{k_{r+1}}_1, \ldots, X^{k_{r+s}}_s =$$
$$(T_p)((\omega^1_p, \ldots, (\omega^r)_p;\ (X_1)_p, \ldots, (X_s)_p) = T(\omega^1, \ldots, \omega^r;\ X_1, \ldots, X_s)(p).$$

This proves that

Proposition 4.109. *As a result of the complete contraction of the tensor field*

$$T \otimes \omega^1 \otimes \cdots \otimes \omega^r \otimes X_1 \otimes \cdots \otimes X_s$$

by the indices $(1, s+1), (2, s+2), \ldots, (r, s+r), (r+1, 1), (r+2, 2), \ldots, (r+s, s)$, *we obtain the value*

$$T(\omega^1, \ldots, \omega^r; X_1, \ldots, X_s)$$

of the field T *on the forms* $\omega^1, \ldots, \omega^r$ *and the fields* $X_1, \ldots, X_s$.

Let us consider in conclusion an arbitrary smooth map Φ of some manifold N into a manifold M. Let T be a tensor field of type $(0, s)$ on a manifold M. We define a tensor $(\Phi^*T)_p$ of type $(0, s)$ at each point $p \in N$ by setting

$$(\Phi^*T)_p(A_1, \ldots, A_s) = T_p(d\Phi_p(A_1), \ldots, d\Phi_p(A_s))$$

for any vectors $A_1, \ldots, A_s \in N_p$. It is easy to verify that we thereby obtain a smooth tensor field Φ^*T on the manifold N and that the map

$$\Phi^* : \mathcal{O}_s^0(M) \to \mathcal{O}_s^0(N)$$

is a homomorphism of $\mathcal{O}(M)$)-modules (if the $\mathcal{O}(N)$-module $\mathcal{O}_s^0(N)$ is considered as an $\mathcal{O}(M)$-module via the map $\Phi^* : \mathcal{O}(M) \to \mathcal{O}(N)$).

Obviously, in the case where the manifold N is an open submanifold W of the manifold M, and the map Φ is an inclusion map, the map Φ^* coincides with the restriction map $\mathcal{O}_s^0(M) \to \mathcal{O}_s^0(N)$. On the other hand, for tensor fields of type $(0, 1)$, i.e., for linear differential forms, the map Φ^* coincides with the map Φ^* constructed in §4.7. For fields of type (r, s) with $\neq$, the map Φ^* is not defined.

4.10 Riemannian spaces

A smooth m-dimensional manifold M is called a *Riemannian space* if it is given a smooth field of symmetric bilinear positive definite forms, i.e., a tensor field g of type $(0, 2)$ such that

1) for any vector fields $X, Y \in \mathcal{O}^1(M)$ the following equality holds:

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

2) for any point $p \in M$ the bilinear form g_p on the space M_p is positive definite.

We will call the tensor field $g(X, Y)$ the *metric tensor field* of the Riemannian space M. To simplify formulae, we will generally denote the function $g(X,)$ by the symbol (X, Y). It depends symmetrically on X and Y and is non-negative for $X = Y$ (and is equal to zero only at those points $p \in M$ for which $X_p = 0$).

For any point p in a Riemannian space M, the bilinear form g_p defines a scalar product in the tangent space M_p

$$(A, B) = g_p(A, B), \quad A, , B \in M_p,$$

which transforms this space into a Euclidean space. This allows us, in particular, to talk about the length of vectors in M_p and the angle between them. We will denote the length $\sqrt{(A,A)}$ of a vector $A \in M_p$ by the symbol $|A|$.

Every vector field X on a Riemannian space M defines, by the formula

$$\omega_X(Y) = (X,Y), \quad Y \in \mathcal{O}^1(M),$$

a certain linear differential form ω_x. Conversely, by the well-known properties of Euclidean spaces, every covector $\omega_p \in M_p$ uniquely corresponds to a vector $X_p \in M_p$ such that

$$\omega_p(B) = (X_p, B)$$

for any vector $B \in M_p$. Moreover, it is easily verified that if the covectors ω_p form a field ω that is smooth on M, i.e., a linear differential form, then the corresponding vectors X_p form a smooth vector field X for which $\omega_X = \omega$. Thus,

Proposition 4.110. *for Riemannian spaces, the concepts of a linear differential form and a vector field essentially coincide.*

In particular, for any smooth function f on M, there is a vector field grad f associated with the differential df of this function by the formula

$$df(X) = (\operatorname{grad} f, X), \quad X \in \mathcal{O}^1(M).$$

We will call this the *gradient field* of f.

Remark 4.111. Similar arguments show that on the Riemannian space M, we can identify tensor fields of any type (r,s) with fields of every type (r',s') for which $r'+s' = r+s$. This identification is known as the operation of “raising and lowering indices.” We will need it only in the case considered above $r+s=1$.

Postponing a more or less in-depth study of Riemannian spaces until Chapter 7, we will limit ourselves here to proving that

Proposition 4.112. *any smooth separable manifold M can be defined as a Riemannian space, i.e., there exists at least one field g on it that has the above properties 1) and 2).*

To this end, we first note that, since any separable manifold M has a countable base, it can be covered by a countable system of coordinate neighbourhoods U_n, $n \geq 1$, and these neighbourhoods can be chosen so that the closure $\overline{U}_n$ of each of them is compact and so that the local coordinates $x_n^1, \ldots, x_n^m$ defined in the neighbourhood U_n are also local coordinates in some larger neighbourhood containing the closure $\overline{U}_n$ of U_n.

Having chosen such a system of coordinate neighbourhoods U_n, we consider the sets

$$A_n = \cup_{k=1}^n \overline{U}_k.$$

Each of these sets is compact, and their union is the entire manifold M. First, let $n = 1$. We construct a tensor field g on the set $A_i = \overline{U}_i$, assuming that its components

$$g_{ij} = g\left(\frac{\partial}{\partial x_1^i}, \frac{\partial}{\partial x_1^j}\right), \quad i, j = 1, \ \ldots, \ m,$$

in the basis $\frac{\partial}{\partial x_1^1}, \ \ldots, \ \frac{\partial}{\partial x_1^m}$ are the numbers δ^i_j. Clearly, the field g is thus completely defined on A_1 and satisfies all the necessary conditions (symmetric and positive definite).

Now let the field g already be constructed on the set A_{n-1}, where $n > 1$. Then, in particular, it is also constructed on the set $A_{n-1} \cap \overline{U}_n$, and therefore for any point $p \in A_{n-1} \cap \overline{U}_n$ the numbers are defined

$$g_{ij}(p) = g\left(\frac{\partial}{\partial x_n^i}, \frac{\partial}{\partial x_n^j}\right), \quad i, j = 1, \ \ldots, \ m.$$

The set of $\frac{m(m+1)}{2}$ smooth functions $g_{ij}(p) = g_{ji}(p)$ can be considered as a smooth map of the compact set $A_{n-1} \cap \overline{U}_n$ into the set D of all symmetric positive definite matrices of order m. The space of all symmetric matrices is isomorphic to the $\frac{m(m+1)}{2}$-dimensional arithmetic space $\mathbb{R}^{\frac{m(m+1)}{2}}$ and, as is easy to see, the set D is a convex open subset of this space. Therefore (see §4.4) the set D is an E-manifold. Consequently, the smooth map $A_{n-1} \cap \overline{U}_n \to D$ constructed above is the restriction of some smooth map $M \to D$, i.e., on the manifold M there exist $\frac{m(m+1)}{2}$ smooth functions $g_{ij}(p) = g_{ji}(p)$ that coincide on $A_{n-1} \cap \overline{U}_n$ with the functions $g_{ij}(p)$ constructed above, $\in A_{n-1} \cap \overline{U}_n$, and have the property that for any point $p \in M$ the matrix $g_{ij}(p)$ is symmetric and positive definite. Taking the numbers $g_{ij}(p)$ for any point $p \in \overline{U}_n$ as components in the basis $\left(\frac{\partial}{\partial x_n^1}\right)_p, \ \ldots, \ \left(\frac{\partial}{\partial x_n^m}\right)_p$ of some tensor, we obtain a tensor field g on $\overline{U}_n$ that coincides on $A_{n-1} \cap \overline{U}_n$ with the field gg previously constructed on A_{n-1}. We have thus extended the field g from A_{n-1} to $A_n = A_{n-1} \cup \overline{U}_n$. (In this argument, we implicitly assumed that $A_{n-1} \cap \overline{U}_n \neq \varnothing$. In the case when $A_{n-1} \cap \overline{U}_n = \varnothing$, the field g on A_n is constructed in an obvious way.)

Thus, the field g is constructed by induction on all sets A_n, and therefore on the entire manifold M.

Remark 4.113. Generally speaking, separability (or, equivalently, the existence of a countable base) is not a necessary condition for constructing a metric tensor field on a manifold M. Indeed, for example, it is clear that the topological sum of any number of Riemannian spaces is automatically defined as a Riemannian space. However, as we will show in Chapter 7, if M is connected, then this condition is necessary, i.e., any connected Riemannian space is necessarily a space with a countable base.

Chapter 5

Smooth Manifolds. II

This chapter focuses primarily on the concept of a submanifold.

§5.1 presents the basic definitions related to the concepts of a curve and a surface (of any dimension) in an arbitrary smooth manifold.

§5.2 proves a theorem on the possibility of extending an arbitrary tensor field defined on a sufficiently "small" curve to the entire manifold.

§5.3 examines the interpretation of vector fields on a smooth manifold as "fields of infinitesimal transformations" of one-parameter diffeomorphism groups. In this regard, in particular, the existence of integral curves for an arbitrary vector field is proved.

§5.4 presents the simplest properties of submanifolds. In particular, it is proved that any submanifold can be locally defined by a system of independent equations.

In §5.5, the existence of tubular neighbourhoods is proved for closed submanifolds of Euclidean spaces.

In §5.6, Whitney's theorem on the embedding of any compact manifold into Euclidean space is proved. In this section, Sard's well-known theorem on critical points of smooth maps of manifolds is proved.

In §5.7, Whitney's embedding theorem is proved for any separable manifold M. To this end, the manifold M is first embedded in a countable-dimensional Euclidean space $\mathbb{R}^\infty$. The final transition from the space $\mathbb{R}^\infty$ to the space $\mathbb{R}^n$ is accomplished through a series of suitably chosen projections.

5.1 Curves and surfaces

By a *continuous curve* on a smooth manifold M we will mean an arbitrary continuous map $\gamma : J \to M$ of some open interval J of the real axis $\mathbb{R}$ into the manifold M. The points $p \in M$ for which there exists a number $t \in J$ such that $p = \gamma(t)$ will be called *points of the curve* γ and we will also say that the curve γ *passes* through these points.

The restriction $\gamma|_C$ of the curve $\gamma : J \to M$ on some closed interval $C \subset J$ will be called a *segment* of the curve γ (in this case, we will generally assume that the interval C is not reducible to a single point). We will call the segment $\gamma|_C$ *finite* if the interval C is finite.

Allowing a certain freedom of speech, in what follows we will also call continuous maps into the manifold M of closed intervals $C \subset \mathbb{R}$ *curves.* By this convention, the paths introduced in §1.5 are a special case of curves (obtained when $C = [0, 1]$). For each curve $\gamma : [a, b] \to M$, we will call the points $p = \gamma(a)$ and $q = \gamma(b$ its *initial* and *final* points, respectively. We will also say that the curve $\gamma : [a, b] \to M$ *connects* the point $p = \gamma(a)$ with the point $q = \gamma(b)$.

Since curves $\gamma : J \to M$ are, by definition, maps of smooth manifolds, it makes sense to speak of curves that are *smooth* (at a given point $t \in J$ or on the entire interval J). A curve that is smooth at all points of the interval J, with the exception of a finite number of them, is called a *piecewise smooth* curve. A piecewise smooth curve can be thought of as a finite sequence of smooth curves with the property that the initial point of each subsequent curve in this sequence coincides with the endpoint of the previous curve.

An example of a smooth curve is the curve $\delta_p : J \to M$, which is a constant map of the interval J to some point $p \in M$. We will call curves of this kind *degenerate.*

A smooth manifold M is called *connected* if any two of its points can be connected by a piecewise smooth curve. Every such manifold is automatically connected as a topological space, i.e., it has no nonempty proper subsets that are simultaneously closed and open. (See §1.2) It is easy to see that the converse is also true, i.e.,

Proposition 5.1. *any smooth manifold M that is connected as a topological space is a connected manifold.*

Proof. Indeed, let N be the subset of M consisting of all points $p \in M$ that can be connected by a piecewise smooth curve to some fixed point $p_0 \in M$. Clearly, $N \neq \varnothing$ (since $p_0 \in N$). Furthermore, it is easy to see that the subset N is open in M.
(Indeed, each point $p \in N$ has in M a neighbourhood U diffeomorphic to an open ball of Euclidean space and therefore consisting of points q that can be connected to the point p by a smooth curve; combining this curve with a piecewise smooth curve connecting the point p_0 with the point p, we obtain a piecewise smooth curve connecting the point p_0 with the point q.)
By similar considerations, the complement $M \setminus N$ of N is also open, i.e., N is closed. Being a nonempty, simultaneously closed, and open subset of a connected topological space, N must coincide with the entire space. Thus, any point of the manifold M can be connected by a piecewise smooth curve to some fixed point p_0. But it is clear that this is only possible when the manifold M is connected. □

Remark 5.2. It can be proven that a smooth manifold is connected if and only if any two of its points can be connected by a smooth (and even regular, see below) curve. We won't need this fact, so we'll leave it unproven.

For any smooth function f on M, each curve $\gamma : J \to M$ defines a numerical function on the interval J

$$f(t) = f(\gamma(t)), \quad t \in J.$$

This function is smooth (infinitely differentiable) if the curve γ is smooth.

In particular, if the curve γ is contained in the coordinate neighbourhood U, i.e., if $\gamma(t) \subset U$, then on the interval J the functions

$$x^1(t) = x^1(\gamma(t)), \ \ldots, \ x^m(t) = x^m(\gamma(t)),$$

where $x^1, \ \ldots, \ x^m$ are local coordinates in the neighbourhood U. These functions uniquely define the curve γ. We will say that they perform a *parametric definition* of this curve in the neighbourhood U. It is clear that

Proposition 5.3. *a curve γ is smooth if and only if the functions $x^1(t), \ \ldots, \ x^m(t)$ are smooth functions of the parameter t.*

We can parametrically define an arbitrary curve on a manifold γ on each coordinate neighbourhood through which this curve passes.

Let M_* be the manifold of tangent vectors of M (see §4.6) and let

$$\mu : M_* \to M$$

be the natural projection ($\mu(A) = p$ if $A \in M_p$). By a *vector field* on a curve $\gamma : J \to M$ we mean an arbitrary curve

$$X : J \to M_*$$

of M_* that projects onto the curve γ, i.e., has the property that

$$\mu \circ X = \gamma.$$

In other words, the vector field $X(t)$ on the curve $\gamma(t)$ is a function that associates with each point $t \in J$ a certain vector $X(t) \in M_{\gamma(t)}$. On each coordinate neighbourhood U, an arbitrary vector field $X(t)$ has the form

$$X(t) = X^i(t) \cdot \left(\frac{\partial}{\partial x^i}\right)_{\gamma(t)},$$

where $X^i(t) = X(t)x^i$ are certain smooth numerical functions (*components* of the field $X(t)$ in the local coordinate system under consideration). It is easy to see that a vector field $X(t)$ on a smooth curve $\gamma(t)$ is smooth (i.e., is a smooth curve of the manifold M_*) if and only if all its components $X^1(t), \ \ldots, \ X^m(t)$ are smooth functions of the parameter t.

An example of a smooth vector field on a smooth curve $\gamma : J \to M$ is the field

$$\dot{\gamma}(t) = d\gamma_t \left(\frac{\partial}{\partial t}\right).$$

We will call this field the *field of tangent vectors to the curve* γ. It is identically equal to zero if and only if the curve γ is degenerate. A curve $\gamma : J \to M$ we will call *regular at a point* $t \in J$ if $\dot\gamma(t) \neq$ (this definition agrees with the general definition of a regular map; see §3.6). We will call a curve $\gamma : J \to M$, regular at all points of the interval J, a *regular curve*, and a curve $\gamma : J \to M$, regular at all points of the interval J, with the exception of a finite number of them, a *piecewise regular curve.*

Since for numerical functions of the parameter t the operation $\frac{\partial}{\partial t}$ coincides with ordinary differentiation with respect to t and since for any function f that is smooth at a point $\gamma(t)$ of the curve $\gamma : J \to M$, the following equality holds

$$\frac{\partial}{\partial t}(f(\gamma(t))) = d\gamma_t\left(\frac{\partial}{\partial t}\right)f = \dot\gamma(t)f,$$

thus

$$\dot\gamma(t)f = \lim_{h\to 0}\frac{f(\gamma(t+h)) - f(\gamma(t))}{h}. \tag{5.4}$$

In particular, the derivatives $\dot x^i(t) = \frac{dx^i}{dt}$ of the functions $x^i(t) = x^i(\gamma(t))$, which parametrically define the curve $\gamma(t)$ in a certain system of local coordinates $x^1, \dots, x^m$ are expressed by the formulae

$$\dot x^i(t) = \dot\gamma(t)x^i, \quad i = 1, \dots, m.$$

This means that

$$\dot\gamma(t) = \dot x^i(t)\cdot\left(\frac{\partial}{\partial x^i}\right)_{\gamma(t)},$$

i.e., that

Proposition 5.5. *the functions $\dot x^i(t)$ are components of the field $\dot\gamma(t)$ in the system of local coordinates $x^1(t), \dots, x^m(t)$.*

Therefore, for any smooth function f on M, the derivative $\dot f(t) = \gamma(t)f$ of the function $f(t) = f(\gamma(t))$ can be calculated using the usual "rule for differentiating a composite function";

$$\dot f(t) = \left(\frac{\partial f}{\partial x^i}\right)_{\gamma(t)}\cdot\frac{dx^i(t)}{dt}., \tag{5.6}$$

For any smooth curve $\gamma : J \to M$ and any smooth map $\Phi : M \to N$ on the manifold N, the curve is defined:

$$\gamma_\Phi = \Phi\circ\gamma.$$

From formula (4.79) of §4.6 it follows directly that for any $t \in J$ the following equality holds

$$\dot\gamma_\Phi = d\Phi_{\gamma t}(\dot\gamma(t)). \tag{5.7}$$

Along with vector fields, we will also consider (smooth) tensor fields of any given type (r, s) on the (smooth) curve $\gamma : J \to M$. Since we have not introduced an analogue of the manifold M_* for arbitrary tensors, we are forced to give an independent definition to this concept here. Namely, by a *tensor field $T(t)$ of type (r,s) on a curve $\gamma(t)$* we will call an arbitrary function that associates with each point $t \in J$ a certain tensor $T(t) \in \mathcal{O}_s^r(\gamma(t))$. A tensor field $T(t)$ of type (r,s) on a smooth curve $\gamma(t)$ will be called *smooth* if for any linear differential forms $\omega^1, \ \ldots, \ \omega^r \in \mathcal{O}_1(M)$ and any vector fields $X_1, \ \ldots, \ X_s \in \mathcal{O}^1(M)$ the numerical function

$$T(t)((\omega^1)_{\gamma(t)} \cdots (\omega^r)_{\gamma(t)}; \ (X_1)_{\gamma(t)} \cdots (X_s)_{\gamma(t)})$$

at a point $t \in J$ is a smooth function on J. If the curve $\gamma(t)$ is contained in a coordinate neighbourhood U, then for any $t \in J$ the following decomposition holds:

$$T(t) = T^{k_1, \ \ldots, \ k_{r-1}}_{\ell_1, \ \ldots, \ \ell_{s-1}}(t)(dx^{\ell_1})_{\gamma(t)} \otimes \cdots \otimes (dx^{\ell_s})_{\gamma(t)} \otimes \left(\frac{\partial}{\partial x^{k_1}}\right)_{\gamma(t)} \otimes \cdots \otimes \left(\frac{\partial}{\partial x^{k_r}}\right)_{\gamma(t)},$$

where $x^1, \ \ldots, \ x^m$ are local coordinates in the neighbourhood U, and

$$T^{k_1, \ \ldots, \ k_{r-1}}_{\ell_1, \ \ldots, \ \ell_{s-1}}(t) = T(t)\left(dx^{\ell_1})_{\gamma(t)} \cdots (dx^{\ell_s})_{\gamma(t)}; \ \left(\frac{\partial}{\partial x^{k_1}}\right)_{\gamma(t)}, \ \ldots, \ \left(\frac{\partial}{\partial x^{k_r}}\right)_{\gamma(t)}\right).$$

We will call the functions $T^{k_1, \ \ldots, \ k_{r-1}}_{\ell_1, \ldots, \ell_{s-1}}(t)$ the components of the field $T(t)$ in the system of local coordinates $x^1, \ \ldots, \ x^m$. It is easy to see that the field $T(t)$ is smooth if and only if all its components $T^{k_1, \ldots, k_{r-1}}_{\ell_1, \ \ldots, \ \ell_{s-1}}(t)$ are smooth functions on J.

When considering tensor fields on the curve $\gamma : J \to M$, we will generally assume that $(r,s) \neq (0,0)$, since tensor fields of type $(0,0)$ are nothing more than numerical functions defined on the interval J and therefore, essentially, not connected in any way with the manifold M and the curve γ.

A generalisation of the concept of a smooth curve is the concept of an *n-dimensional (smooth) surface*, i.e., a smooth map of some connected open set $G \subset \mathbb{R}^n$ into a manifold M. As a rule, we will consider the set G to be a rectangle $J_1 \times \cdots \times J_n$, i.e., a product of n open intervals $J_1, \ \ldots, \ J_n \subset \mathbb{R}^n$. Furthermore, to simplify the notation, we will assume that $n = 2$, although everything said below will be true (with the appropriate obvious modifications) for any larger n.

Let

$$\varphi : J_1 \times J_2 \to M$$

be an arbitrary surface. By a *vector field* on a surface φ we mean an arbitrary (smooth) surface

$$X : J_1 \times J_2 \to M_*$$

of a manifold M_* with the property that

$$\mu \circ X = \varphi,$$

where $\mu : M_* \to M$ is the natural projection. In other words, the vector field $X(t^1, t^2)$ on the surface $\varphi(t^1, t^2)$ is a function that associates with each point

$(t^1, t^2) \in J_1 \times J_2$ a certain vector $X(t^1, t^2) \in M_{\varphi(t^1,t^2)}$ and has the property that for any function f that is smooth at the point $p = \varphi(t^1, t^2)$, the numerical function $g(t^1, t^2) = X(t^1, t^2)f$ is smooth at he point (t^1, t^2).

Similarly, we can define the concept of a tensor field on a surface φ of any given type (r, s). We will not dwell on this, since we will not need this concept further.

At each point (t^1, t^2) of the rectangle $J_1 \times J_2$, two vectors $\frac{\partial}{\partial t^1}$ and $\frac{\partial}{\partial t^2}$ are defined, forming a basis for the space $(J_1 \times J_2)_{(t^1,t^2)}$. The images of these vectors under the map

$$d\varphi_{(t^1,t^2)} : (J_1 \times J_2)_{(t^1,t^2)} \to M_{\varphi(t^1,t^2)}$$

We will denote them by the symbols $\frac{\partial\varphi}{\partial t^1}(t^1, t^2)$ and $\frac{\partial\varphi}{\partial t^2}(t^1, t^2)$ respectively and will call them the *basis tangent vectors* of the surface φ at the point (t^1, t^2). Clearly, this gives us two vector fields $\frac{\partial\varphi}{\partial t^1}$ and $\frac{\partial\varphi}{\partial t^2}$ on the surface φ.

We will say that a curve $\gamma : J \to M$ belongs to a surface $\varphi : J_1 \times J_2 \to M$ if there exist smooth functions $t^1(t) \in J_1$, $t \in J$, and $t^2(t) \in J_2$, $t \in J$ such that

$$\gamma(t) = \varphi(t^1(t), t^2(t))$$

for any $t \in J$. It is easy to verify that, as in elementary analysis, the tangent vectors of such a curve are defined by the formula

$$\dot\gamma(t) = \frac{\partial\varphi}{\partial t^1}(t^1, t^2) \cdot \dot t_1(t) + \frac{\partial\varphi}{\partial t^2}(t^1, t^2) \cdot \dot t_2(t).$$

For any n the corresponding formula is:

$$\dot\gamma(t) = \frac{\partial\varphi}{\partial t^i}(t^1, t^2) \cdot \dot t_i(t).$$

where

$$\gamma(t) = \varphi(t^1(t),\ \ldots,\ t^n(t))$$

Among the curves belonging to the surface $\varphi : J_1 \times J_2 \to M$, the so-called "coordinate curves" are of particular importance.

$$\varphi_{t^1} : J_2 \to M, \quad \varphi_{t^2} : J_1 \to M.$$

The first of these curves is defined for each fixed $t^1 \in J_1$ by the formula

$$\varphi_{t^1}(t^2) = \varphi(t^1, t^2), \quad t^2 \in J_2,$$

and the second is defined for each fixed $t^2 \in J_2$ by the formula

$$\varphi_{t^2}(t^1) = \varphi(t^1, t^2), \quad t^1 \in J_1.$$

The tangent vectors of these curves are the base tangent vectors of the surface φ:

$$\dot\varphi_{t^1}(t^2) = \frac{\partial\varphi}{\partial t^2}(t^1, t^2), \quad \dot\varphi_{t^2}(t^1) = \frac{\partial\varphi}{\partial t^1}(t^1, t^2).$$

If the surface $\varphi : J_1 \times J_2 \to M$ is contained in the coordinate neighbourhood U (i.e., if $\varphi(J_1 \times J_2) \subset U$), then then on the rectangle $J_1 \times J_2$ the functions

$$x^1(t^1, t^2) = x^1(\varphi(t^1, t^2)), \ \ldots, \ x^m(t^1, t^2) = x^m(\varphi(t^1, t^2)).$$

are said to be implementing the *parametric definition* of the surface φ in the neighbourhood U. Moreover, in this case, any vector field $X(t^1, t^2)$ on the surface φ has the form

$$X(t^1, t^2) = X^i(t^1, t^2) \left(\frac{\partial}{\partial x^i} \right)_{\varphi(t^1, t^2)},$$

where $X^i(t^1, t^2) = X(t^1, t^2)x^i$ are some smooth functions on $J_1 \times J_2$ (*components* of the field $X(t^1, t^2)$ in the local coordinate system under consideration). In particular,

$$\frac{\partial \varphi}{\partial t^1}(t^1, t^2) = \frac{\partial x^i(t^1, t^2)}{\partial t^1} \cdot \left(\frac{\partial}{\partial x^i} \right)_{\varphi(t^1, t^2)},$$
$$\frac{\partial \varphi}{\partial t^2}(t^1, t^2) = \frac{\partial x^i(t^1, t^2)}{\partial t^2} \cdot \left(\frac{\partial}{\partial x^i} \right)_{\varphi(t^1, t^2)}.$$

5.2 Extension of tensor fields

We will say that a tensor field $T(t)$ defined on a curve $\gamma(t)$ of a manifold M is *extendable* on a segment $\gamma|_C$ of this curve if on M there exists a (smooth) tensor field T - an *extension* of the field $T(t)$ - such that

$$T(t) = T_{\gamma(t)}$$

for any $t \in C$.

Moreover, we will say that a curve $\gamma : J \to M$ (or its segment $\gamma|_C$) has no self-intersections if $\gamma(t_1) \neq \gamma(t_2)$ for $t_1 \neq t_2$ (where $t_1, t \in J$ or $t_1, t_2 \in C$, respectively). We will call a finite segment $\gamma|_C$ of a curve γ *elementary* if it has no self-intersections and is contained in some coordinate neighbourhood U of the manifold M.

In this section, we will show that

Proposition 5.8. *any tensor field $T(t)$ on a regular curve $\gamma(t)$ is extendable on each of its elementary segments $\gamma|_C$.*

It is clear that to prove this statement it is sufficient to construct an extension T of the field $T(t)$ only on the coordinate neighbourhood $U \supset \gamma(C)$. Indeed, since the segment C is finite and the curve γ is continuous, the set $\gamma(C)$ is compact and therefore has a neighbourhood $V \subset U$ with compact closure $\overline{V} \subset U$. Therefore, according to the local epimorphism property of the restriction map (see §4.8), the field T defined on U coincides on V with some field defined on the entire manifold M. This latter field will be the desired extension of the field $T(t)$ on the interval $\gamma|_C$.

On the other hand, in order to construct an extension of the field $T(t)$ to a neighbourhood U, it is sufficient for each of its components $T^{k_1,\dots,k_r}_{\ell_1,\ \dots,\ \ell_s}(t)$ (in some local coordinate system $x^1, \dots, x^m$ defined in U) to find on M (or at least on U) a smooth function $T^{k_1,\ \dots,\ k_r}_{\ell_1,\ \dots,\ \ell_s}$ such that

$$T^{k_1,\ \dots,\ k_r}_{\ell_1,\ \dots,\ \ell_s}(t) = T^{k_1,\ \dots,\ k_r}_{\ell_1,\ \dots,\ \ell_s}(\gamma(t))$$

for any $t \in C$. Indeed, then the formula

$$T = T^{k_1,\ \dots,\ k_r}_{\ell_1,\ \dots,\ \ell_s} dx^{\ell_1} \otimes \cdots \otimes dx^{\ell_s} \otimes \frac{\partial}{\partial x^{k_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{k_r}},$$

will define the extension T of the field $T(t)$ to the neighbourhood U.

Thus, to prove the statement under consideration, it suffices to prove that

Proposition 5.9. *for any smooth function $g(t)$ on C, there exists a smooth function f on the manifold M such that*

$$g(t) = f(\gamma(t))$$

for any $t \in C$.

Consider the functions

$$x^1(t) = x^1(\gamma(t)), \ \dots, \ x^m(t) = x^m(\gamma(t)),$$

that parametrically define the curve $\gamma(t)$ in the neighbourhood U. Since the curve $\gamma(t)$ is, by assumption, regular, then for any point $t_0 \in C$ there exists an index i_0 such that $\dot{x}^{i_0}(t_0) \neq 0$. Therefore, according to the inverse function theorem, the function $x^{i_0}(t)$ is a diffeomorphism on some neighbourhood $|t-t_0| < \delta$ of the point t_0 onto some neighbourhood of the point $x^{i_0}(t_0)$. In other words, in the neighbourhood of the point $x^i(t_0) \in \mathbb{R}$ there exists a smooth function y such that $(x^{i_0}(t)) = t$ for $|t - t_0| < \delta$. Consider the function

$$\widetilde{f}_{t_0}(p) = g(y(x^{i_0}(p))).$$

This function is defined in some neighbourhood of the point $p_0 = \gamma(t_0)$ and is a smooth function at p_0. Therefore, the point p_0 has a neighbourhood $\widetilde{U}_{t_0}$ in which the function $\widetilde{f}_{t_0}$ coincides with some smooth function f_{t_0} on M.

We can assume that for $|t-t_0| < \delta$ the inclusion $\gamma(t) \in \widetilde{U}_{t_0}$ holds and therefore

$$f_{t_0}(\gamma(t)) = \widetilde{f}_{t_0}(\gamma(t)) = g(y(x^{i_0}(\gamma(t))) = g(y(x^{i_0}(t))) = g(t).$$

Since the segment C is finite, the set $C^{'}$ obtained by removing from C the points t for which $|t-t_0| < \delta$ is compact, and therefore its image $\gamma(C^{'})$ under the map γ is closed. Moreover, it does not contain the point p_0, since the segment $\gamma|_C$ has no self-intersections, by assumption. Therefore, the point p_0 has a neighbourhood U_{t_0} such that

$$U_{t_0} \cap \gamma(C^{'}) = \varnothing.$$

Therefore, if a number $t \in C$ has the property that $\gamma(t) \in U_{t_0}$, then $|t - t_0| < \delta$ and therefore, according to what was said above,

$$f_{t_0}(\gamma(t)) = g(t).$$

Now let W_{t0} be an arbitrary neighbourhood of the point $p_0 = \gamma(t_0)$ whose closure $\overline{W}_{t_0}$ is compact and contained in the neighbourhood U_{t_0}. Since the set $\gamma(C)$ is compact, there exists a finite system of points $t_1, \ldots, t_r$ on the segment C such that the corresponding neighbourhoods

$$W_1 = W_{t_1}, \ \ldots, \ W_r = W_{t_r}$$

cover the set $\gamma(C)$. Then the neighbourhoods

$$U_1 = U_{t_1}, \ \ldots, \ U_r = U_{t_r}$$

form a covering of the set $\gamma(C)$ that satisfies all the conditions of the proposition 4.34 proved in §4.4. Therefore, there exists a partition of unity $\{h1, \ \ldots, \ h_r\}$ on the set $\gamma(C)$ subordinate to this covering. Let us consider the function

$$f = f_1 h_1 + \cdots + f_r h_r.$$

This function is smooth on M. Moreover, since, by construction, $f_i(\gamma(t) = g(t)$ if $\gamma(t) \in U_i$, and $h_i(\gamma(t)) = 0$ otherwise, then

$$f(\gamma(t)) = g(t).$$

for any $t \in C$.

Thus, the above statement is completely proven.

Remark 5.10. A similar statement holds for any segment $\gamma|_C$ of the curve γ that has no self-intersections. We will not need this fact, and therefore we will leave it without proof.

To understand the generality of the proved statement, it is useful to note that

Proposition 5.11. *any point $p_0 = \gamma(t_0)$ of the regular curve $\gamma : J \to M$ belongs to some elementary segment $\gamma|_C$.*

Proof. Indeed, let $x^1, \ \ldots, \ x^m$ be an arbitrary system of local coordinates at the point p_0. As we have already seen, the regularity condition for the curve γ implies that among the functions $x^i(t)$ that parametrically define this curve, there exists a function $x^{i_0}(t)$ whose derivative $\dot{x}^{i_0}(t)$ is nonzero at the point t_0, and therefore also on some closed segment $C \ni t_0$. Since $\dot{x}^{i_0}(t) \neq$ on C, we have $\dot{x}^{i_0}(t_1) \neq \dot{x}^{i_0}(t_2)$, and therefore $\gamma(t_1) \neq \gamma(t_2)$ for any distinct points $t_1, t_2 \in C$. Consequently, the segment $\gamma|_C$ of the curve γ is elementary. □

5.3 One-parameter diffeomorphism groups and integral curves

A smooth map

$$\varphi : \mathbb{R} \times M \to M$$

of the product $\mathbb{R} \times M$ into a manifold M will be called a *one-parameter diffeomorphism group* of the manifold M if

1) $\varphi(0, p) = p$ for any point $p \in M$;

2) $\varphi(t, \varphi(s, p)) = \varphi(t + s, p)$ for any point $p \in M$ and any numbers $t, s \in \mathbb{R}$.

For each fixed number $t \in \mathbb{R}$, the group φ defines by the formula

$$\varphi_t(p) = \varphi(t, p), \qquad p \in M$$

some (obviously smooth) map

$$\varphi_t : M \to M.$$

Conditions 1) and 2), respectively, mean that

1) the map φ_0 is the identity map of the manifold M;

2) for any numbers $t, s \in \mathbb{R}$, the following equality holds

$$\varphi_t \circ \varphi_s = \varphi_{t+s}.$$

From these conditions it immediately follows that for any number $t \in \mathbb{R}$, the map φ_t has an inverse map φ_{-t} and is therefore a diffeomorphism.

For each fixed point $p \in M$, the group φ defines by the formula

$$\varphi_p(t) = \varphi(t, p), \qquad t \in \mathbb{R},$$

a smooth curve

$$\varphi_p : \mathbb{R} \to M$$

on the manifold M. We will call this curve the *trajectory* of the group φ defined by the point p. For $t = 0$, it passes through the point p. We define a vector field X on the manifold M by setting for any point $p \in M$

$$X_p = \dot{\varphi}_p(0).$$

According to formula (5.4) in §5.1, for any function f smooth at the point p, the equality holds

$$X_p f = \lim_{t \to 0} \frac{f(\varphi_t(p)) - f(p)}{t}. \tag{5.12}$$

In particular, it follows directly that for every smooth function f on M, the function $g(p) = X_p f$ is also smooth on M, i.e., that the field X is a smooth vector field on the manifold M. We will say that this field is *generated* by a given one-parameter group φ of diffeomorphisms of the manifold M.

Generally speaking, not every vector field $X \in \mathcal{O}^1(M)$ is generated by some one-parameter group of diffeomorphisms of the manifold M. However, it can be shown that locally this is true. To give this statement a precise meaning, we consider an arbitrary point p_0 of the manifold M, some of its neighbourhood U_0 and some interval J of the axis $\mathbb{R}$ containing the point 0.

$$\varphi : J \times \mathbb{R} \to M$$

of the product $J \times \mathbb{R}$ into the manifold M will be called a *local one-parameter group of local diffeomorphisms* if

1) $\varphi(0, p) = p$ for any point $p \in U$;

2) $\varphi(t, \varphi(s, p)) = \varphi(t + s, p)$ for any point $p \in U$ of any numbers $t, s \in J$ with the property that $t + s \in J$ and $\varphi(s, p) \in U$.

For every fixed number $t \in J$, the local group φ defines by the formula

$$\varphi_p(t) = \varphi(t, p), \qquad p \in U,$$

a smooth map

$$\varphi_t : U \to M$$

of the neighbourhood U into the manifold M. Conditions (1) and 2 respectively mean that

1) the map φ_0 is the identity map of the neighbourhood U;

2) for any numbers $t, s \in J$, and any point $p \in U$ with the property that $t+s \in J$ and $\varphi_s(p) \in U$ the following equality holds

$$\varphi_t(\varphi_s(p)) = \varphi_{t+s}(p).$$

Elementary considerations using the continuity of the map φ and the local compactness of the real axis $\mathbb{R}$ show that there exists a neighbourhood $U_0 \subset U$ of the point p_0 and an interval $J_0 \subset J$ of the axis $\mathbb{R}$ centred at the point 0 such that $\varphi_t(U_0) \subset U$ for any $t \in J_0$. It follows that for $t \in J_0$ and for any point $p \in U_0$ the equality following holds

$$\varphi_{-t}(\varphi_t(p)) = \varphi_t(\varphi_{-t}(p)) = p.$$

In other words, the restriction $\varphi_t,|_{U_0}$ of the map φ_t, $t \in J_0$, to the neighbourhood U_0 has an inverse map φ_t, and therefore is a diffeomorphism of this neighbourhood onto the neighbourhood $\varphi_t(U_0)$ of the point $\varphi_t(p_0)$.

For each fixed point $p \in U$, the local group φ defines by the formula

$$\varphi_p(t) = \varphi(t, p), \qquad t \in J,$$

a certain smooth curve

$$\varphi_p : J \to M.$$

As in the non-local case, we will call this curve the *trajectory* of the group φ defined by the point p. For $t = 0$, it passes through the point p. We define a vector field X on the neighbourhood U, setting for any point $p \in U$

$$X_p = \dot{\varphi}_p(0).$$

As above, it is immediately verified that the vector field X constructed in this way is smooth. We will say that it is *generated* by the local group φ of local diffeomorphisms of the neighbourhood U.

We call a smooth curve $\gamma : J \to M$ an *integral curve* of a vector field $X \in \mathcal{O}^1(M)$ if

$$\dot{\gamma}(t) = X_{\gamma(t)}$$

for any $t \in J$. It is easy to see that

Proposition 5.13. *every trajectory φ_p of a local group φ of local diffeomorphisms of a neighbourhood U is an integral curve of the vector field $\mathcal{O}^1(U)$ generated by this group.*

Proof. Indeed, in view of condition 2), for any number $t \in J$ and any sufficiently small (in absolute value) number $s \in \mathbb{R}$, the following equality holds

$$\varphi_{\varphi_p(t)}(s) = \varphi_p(t+s).$$

Assuming here that t is constant and the number s is a parameter on the curve and moving to tangent vectors at the point $s = 0$, we immediately obtain that

$$X_{\varphi_p(t)} = \dot{\varphi}_{\varphi_p}(0) = \dot{\varphi}(t).$$

□

Let us now assume that the neighbourhood U specified by the definition of the local one-parameter group

$$\varphi : J \times U \to M$$

is a coordinate neighbourhood of the point p_0. Without loss of generality, we can assume that the local coordinates $x^1, \ldots, x^m$ defined in the neighbourhood U are equal to zero at the point p_0, i.e., that the corresponding coordinate homeomorphism $\xi : U \to \mathbb{R}^m$ maps the point p_0 to the point $0 \in \mathbb{R}^m$. Fixing such a system of local coordinates and setting

$$\varphi^1(t, \boldsymbol{t}) = x^1(\varphi(t, \xi^{-1}(\boldsymbol{t}))), \ \ldots, \ \varphi^m(t, \boldsymbol{t}) = x^m(\varphi(t, \xi^{-1}(\boldsymbol{t}))),$$

we immediately see that defining a local one-parameter group $\varphi : J \times U \to M$ is equivalent to defining, in the neighbourhood $U^0 = \xi(U)$ of the point 0, a system of m smooth functions

$$\varphi^1(t, \boldsymbol{t}), \ldots, \varphi^m(t, \boldsymbol{t}), \quad t \in J, \boldsymbol{t} \in U^0,$$

having the following properties:

1) for any point $\boldsymbol{t} = (t^1, \ldots, t^m) \in U^0$, the equalities

$$\varphi^1(0, \boldsymbol{t}) = t^1, \ \ldots, \ \varphi^m(0, \boldsymbol{t}) = t^m$$

hold;

2) for any numbers $t, s \in J$, and any point $\boldsymbol{t} \in U^0$ with the property that $t + s \in J$ and the point $\boldsymbol{\varphi}(s, \boldsymbol{t}) = \varphi^1(s, \boldsymbol{t}), \ldots, \varphi^m(s, \boldsymbol{t}) \in \mathbb{R}^m$ belong to the neighbourhood U^0, then

$$\varphi^i(t, \boldsymbol{\varphi}(s, \boldsymbol{t})) = \varphi^i(t + s, \boldsymbol{t})$$

for any $i = 1, \ \ldots, \ m$.

On the other hand, defining a vector field $X \in \mathcal{O}^1(U)$ in a neighbourhood U is equivalent to defining a system of m smooth functions in a neighbourhood U^0

$$X^1(\boldsymbol{t}) = X^1(t^1, \ \ldots, \ t^m), \ldots, X^m(\boldsymbol{t}) = X^m(t^1, \ \ldots, \ t^m),$$

related to the components $X^1 = Xx^1, \ \ldots, \ X^m = Xx^m$ of X by the relations

$$X^1 = X^1(x^1, \ \ldots, \ x^m), \ldots, X^m = X^m(x^1, \ \ldots, \ x^m).$$

Moreover, the field X is generated by a local one-parameter group φ if and only if, for each point $\boldsymbol{t} \in U^0$, the functions

$$\varphi^1(t) = \varphi^1(t, \boldsymbol{t}), \ \ldots, \ \varphi^m(t) = \varphi^m(t, \boldsymbol{t})$$

satisfy the differential equations

$$\dot{\varphi}^i(t) = X^i(\varphi^i(t), \ \ldots, \ \varphi^m(t)), \qquad t \in J$$

expressing the fact that the trajectories $\varphi_p(t)$ of the group φ are integral curves of the field X.

From these considerations it easily follows that

Proposition 5.14. *any vector field $X \in \mathcal{O}^1(M)$ in some neighbourhood U of each point $p_0 \in M$ is generated by one and only one local one-parameter group $\varphi : J \times U \to M$.*

Proof. Indeed, we choose a neighbourhood U_1 of the point p_0 and a system of local coordinates $x^1, \ \ldots, \ x^m$ in this neighbourhood so that the corresponding neighbourhood $U_1^0 = \xi(U_1)$ of the point $\xi(p_0) = \mathbf{0}$ is a cube $|t^i| < 1$, $i = 1, \ \ldots, \ m$, of the space $\mathbb{R}^m$. In this cube, we consider a system of differential equations

$$\dot{\varphi}^i(t) = X^i(\varphi^i(t), \ \ldots, \ \varphi^m(t)), \quad i = 1, \ \ldots, \ m, \tag{5.15}$$

with respect to the unknown functions $\varphi^i(t), \ \ldots, \ \varphi^m(t)$, where $X^i(t^1, \ \ldots, \ t^m$, $i = 1, \ldots, m$, are functions on U_1^0 corresponding, as described above, to the given vector field X. According to the fundamental theorem of existence and uniqueness of solutions of systems of differential equations, there exist positive numbers ε and $\delta < 1$ such that for any point $\boldsymbol{t} = (t^1, \ldots, t^m)$ of the cube $U^0 \subset U_1^0$ defined by the inequalities $|t^i < 1|$, $i = 1, \ldots, m$, on the interval $J = (-\varepsilon, \varepsilon)$ of the $\mathbb{R}$ axis there exists one and only one solution $\varphi 1(t) = \varphi^1(t, \boldsymbol{t}), \ldots \varphi m(t) = \varphi^m(t, \boldsymbol{t})$ of system (5.15) satisfying the initial conditions

$$\varphi^i(0, \boldsymbol{t}) = t^1, \ldots, \varphi^m(0, \boldsymbol{t}) = t^m.$$

Moreover, according to the theorem on the dependence of solutions of differential equations on the initial condition, the functions $\varphi^1(t, \boldsymbol{t}), \ldots \varphi^m(t, \boldsymbol{t})$ smoothly depend on $t, t^1, \ldots, t^m$.

Thus, on the interval J, we have found m smooth functions

$$\varphi^1(t, \boldsymbol{t}), \ldots \varphi^m(t, \boldsymbol{t}), \tag{5.16}$$

that smoothly depend on the point $\boldsymbol{t} \in U^0$ and satisfy system (5.15) and the above condition 1). Therefore, in accordance with what was said above, to prove the existence of a local one-parameter group φ that generates a field X on a neighbourhood $U = \xi^{-1}(U^0)$, it is sufficient for us to show that functions (5.16) also satisfy condition 2).

Let t and s be integers in the interval J such that $t + s \in J$, and let $\boldsymbol{t}$ be a point in the cube U^0 such that

$$\boldsymbol{\varphi}(s, \boldsymbol{t}) = \varphi^1(s, \boldsymbol{t}), \ldots, \varphi^m(s, \boldsymbol{t}) \in U^0.$$

Consider the functions

$$\psi^1(t) = \varphi^1(t + s, \boldsymbol{t}), \ldots, \psi^m(t) = \varphi^m(t + s, \boldsymbol{t}).$$

Clearly, these functions satisfy system (5.15) with the initial conditions

$$\psi^1(0) = \varphi^1(s, \boldsymbol{t}), \ldots, \psi^m(0) = \varphi^m(s, \boldsymbol{t}).$$

Therefore, by the uniqueness theorem for solutions of system (5.15),

$$\psi^i(t) = \varphi^i(t, \boldsymbol{\varphi}(s, \boldsymbol{t})), \quad i = 1, \ldots, m.$$

But this also means that functions (5.16) satisfy condition 2).

The uniqueness of the constructed local one-parameter group immediately follows from the uniqueness of the solutions of system (5.15). □

From the proved statement, in particular, it follows directly that

Proposition 5.17. *for any vector field $X \in \mathcal{O}^1(M)$ and any point $p \in M$, there exists one and only one integral curve $\gamma : J \to M$ of the field X, defined on some interval J of the $\mathbb{R}$ axis containing the point 0, and possessing the property that $\gamma(0) = p$.*

Moreover, according to formula (5.12),

Proposition 5.18. *for any function $f \in \mathcal{O}^1(p)$, the following equality holds*

$$X_p f = \lim_{t\to 0} \frac{f(\gamma(t)) - f(p)}{t}. \tag{5.19}$$

In what follows, we will also need the following somewhat more general proposition, which follows directly from formula (5.4) in §5.1:

For any vector $C \in M_p$, any function $f \in \mathcal{O}^1(p)$, and any curve $\gamma : J \to M$ defined on an interval $J \subset \mathbb{R}$ containing the point 0 and possessing the property that $\gamma(0) = p$, $\dot{\gamma}(0) = C$, the following equality holds

$$Cf = \lim_{t\to 0} \frac{f(\gamma(t)) - f(p)}{t}.$$

Consider now two points $p_0^{'}$ and $p_0^{''}$ of the manifold M and their neighbourhoods $U^{'}$ and $U^{''}$, in which the vector field X is generated by local one-parameter groups $\varphi^{'} : J^{'} \times U^{'} \to M$ and $\varphi^{''} : J^{''} \times U^{''} \to M$, respectively. From the uniqueness of these groups it immediately follows that on the intersection $U^{'} \cap U^{''}$ the groups $\varphi^{'}$ and $\varphi^{''}$ coincide, i.e., either this intersection is empty, or $\varphi_t^{'}|_{U' \cap U''} = \varphi_t^{''} p|_{U' \cap U''}$ for any number $t \in J = J^{'} \cap J^{''}$. Therefore, putting

$$\varphi_t(p) = \begin{cases} \varphi_t^{'}(p), & \text{if} \quad U^{'}, \\ I\varphi_t^{''}(p), & \text{if} \quad U^{''}, \end{cases} \qquad t \in J,$$

we define on $U = U^{'} \cup U^{''}$ a local one-parameter group $\varphi : J \times U \to M$, which generates on U the field X.

This remark immediately implies that for any open set $U \subset M$ whose closure $\overline{U}$ is compact (and therefore covered by a finite number of neighbourhoods in which the field X is generated by some group), there exists one and only one local one-parameter group

$$\varphi : J \times U \to M,$$

where J is some interval of the $\mathbb{R}$-axis containing the point 0, generating a given vector field $X \in \mathcal{O}^1(M)$ on U.

For every vector field $X \in \mathcal{O}^1(M)$, the closure of the set of all points $p \in M$ for which $X_p \neq 0$ will be called the *support* of this field. From the above, it easily follows that

Proposition 5.20. *any vector field $X \in \mathcal{O}^1(M)$ whose support is compact is generated by some one-parameter group*

$$\varphi : \mathbb{R} \times M \to M$$

of diffeomorphisms of the manifold M.

Proof. Indeed, by compactness, the support of the field X has a neighbourhood U with compact closure. On the other hand, according to what was said above, the field X on the neighbourhood U is generated by some local one-parameter group

$$\varphi : J \times U \to M.$$

We extend this group to the entire manifold M, setting

$$\varphi_t(p) = p$$

for all numbers $t \in J$ and all points $p \in M \setminus V$. Clearly, as a result we obtain a local one-parameter group

$$\varphi : J \times M \to M.$$

generating the field X on the entire manifold M.

Let t_0 be a non-zero number in the interval J such that $t_0 \in J$. Any number $t \in \mathbb{R}$ can be written as the sum of some integer multiple of kt_0 of the number t_0 and a remainder r satisfying the inequality $|r| < t_0$ and therefore belonging to the interval J. We define a map $\varphi_t : M \to M$ by setting

$$\varphi_t = \begin{cases} \varphi_{t_0} \circ \varphi_{t_0} \circ \cdots \circ \varphi_{t_0} \circ \varphi_r, & \text{if} \quad k \geq 0, \\ \varphi_{-t_0} \circ \varphi_{-t_0} \circ \cdots \circ \varphi_{t_0} \circ \varphi_r, & \text{if} \quad k < 0, \end{cases}$$

where the map φ_{t_0} (or, respectively, the map φ_{-t_0}) is repeated k times (respectively, $-k$ times). A simple check shows that this gives us a certain one-parameter group

$$\varphi_t : \mathbb{R} \times M \to M$$

of diffeomorphisms of the manifold M. Obviously, this group is an extension (from J to $\mathbb{R}$) of the local group $\varphi : J \times M \to M$ constructed above and therefore also generates a vector field X.

Thus, the above statement is fully proven. □

In particular, we see that

Proposition 5.21. *on a compact manifold M, any vector field $X \in \mathcal{O}^1(M)$ is generated by some one-parameter group*

$$\varphi : \mathbb{R} \times M \to M$$

of diffeomorphisms of the manifold M.

5.4 Submanifolds

We will call a smooth manifold N a *(smooth) submanifold* of a smooth manifold M if it, as a topological space, is contained in the manifold M, its dimension n does not exceed the dimension m of the manifold M, and the inclusion map $\iota : N \subset M$ is smooth and regular.

Remark 5.22. It can be shown that condition $n \leq m$ follows from condition $N \subset M$.

For each point p of a submanifold N in the space M_p, a subspace $d_{\iota_p}(N_p)$ is defined that is isomorphic to the space N_p. We will say that the vectors of this subspace are *tangent* to the submanifold N at the point p. Furthermore, as a rule, we will identify this subspace with the space N_p, i.e., we will consider the monomorphic map d_{ι_p} as an inclusion.

In the special case where $N = M$, we obtain, according to this definition, that any vector of the space M_p is tangent to the manifold M at the point p. This explains the name "tangent space" for the space M_p.

Remark 5.23. The definition of submanifold adopted here differs in form from the definition adopted in §4.3. However, it turns out that both of these definitions are in fact equivalent, i.e., firstly,

Proposition 5.24. *for any sub-premanifold N of M that is a manifold, the inclusion map $\iota : N \subset M$ is regular*

and, secondly,

Proposition 5.25. *any submaniold N of M is its sub-premanifold.*

To prove the first assertion, we note that, according to the definition of smoothness on a submanifold, for any point $p \in N$, every function g smooth at p on N is, near the point p, the restriction of some function f smooth at p on M, i.e., near this point it has the form $f \circ \iota$. On the other hand, by the definition of the map d_{ι_p} for any vector $A \in N_p$, the following equality holds

$$d_{\iota_p}(A)(f) = A(f \circ \iota).$$

Therefore, if $d_{\iota_p}(A) = 0$, then $A = 0$.

To prove the second assertion, we note that, due to the regularity of the inclusion map $\iota : N \subset M$, for any system $x^1, \ldots, x^m$ of local coordinates on the manifold M at an arbitrary point $p \in N$, from the functions $x^1 \circ \iota, \ldots, x^m \circ \iota$ we can choose n functions that form a system of local coordinates at the point p on the submanifold N. Let these be the functions $x^1 \circ \iota, \ldots, x^n \circ \iota$. Then, for any function g smooth at the point p on the submanifold N near this point, the equality

$$g = v(x^1 \circ \iota, \ldots, x^n \circ \iota),$$

where v is some smooth function of n real variables, holds. But then it is clear that the function

$$f = v(x^1, \ldots, x^m)$$

on the manifold M is smooth at the point p and has the property that $f|_N = g$. We have thus proved that any function on N that is smooth at some point of the submanifold N is the restriction of some function on the entire manifold M that is smooth at this point. Since the inclusion map $\iota : M \subset M$ is smooth, the converse is also true, i.e., a function g on N is smooth at a point $p \in N$ if and only if it is the restriction of some function f that is smooth on M near this point. But this also means that the submanifold N is a sub-premanifold of M.

Note that along the way we also proved that

Proposition 5.26. *for any system of local coordinates $x^1, \ldots, x^m$ on the manifold M at a point $p \in N$, from the functions*

$$x^1 \circ \iota, \ \ldots, \ x^m \circ \iota$$

one can choose n functions that form a system of local coordinates at a point p on the submanifold N.

In particular,

Proposition 5.27. *for any point $p \in N$ on the manifold M, there exists a system of local coordinates $x^1, \ldots, x^m$ at the point p such that the functions*

$$x^1|_N, \ \ldots, \ x^n|_N$$

form a system of local coordinates at the point p on the submanifold N.

This property is possessed, for example, by every system of local coordinates $x^1, \ldots, x^m$ for which the vectors $\left(\frac{\partial}{\partial x^1}\right)_p, \ldots, \left(\frac{\partial}{\partial x^n}\right)_p$ belong to the subspace N_p (and therefore constitute its basis).

From the proposition proved above, it also easily follows that

Proposition 5.28. *a smooth map*

$$\Phi : N \to M$$

of a manifold N into a manifold M is a diffeomorphism of the manifold N onto the submanifold $\Phi(N)$ if and only if this map is regular and is a homeomorphism of the manifold N onto the subspace $\Phi(N)$ of the manifold M.

Proof. Indeed, let the map Φ be a diffeomorphism of the manifold N onto the submanifold $\Phi(N)$, i.e., let

$$\Phi = \iota \circ \Psi$$

where $\Psi : N \to \Phi(N)$ is some diffeomorphism, and ι is the inclusion map $\Phi(N) \subset M$. Since the sub-premanifold $\Phi(N)$ is diffeomorphic to the manifold M, it is a manifold and therefore a submanifold of M. Consequently, the map ι is regular. The map Ψ, being a diffeomorphism, is also regular. Therefore, the map Φ is also regular.

Conversely, let Φ be a regular map and let it be a homeomorphism of N onto the subspace $\Phi(N)$ of M. We introduce smoothness on the space $\Phi(N)$ by defining a function f on $\Phi(N)$ to be smooth if and only if the function $f \circ \Phi$ on N is smooth. Clearly, this defines the subspace $\Phi(N)$ as a smooth manifold, and the map

$$\Psi : N \to \Phi(N),$$

induced by Φ is a diffeomorphism and therefore a regular map. Let $\iota : \Phi(N) \subset M$ be the inclusion map. Since $\iota = \Phi \circ \Psi^{-1}$ and the map Φ is, by assumption, regular, the map ι is also regular. Therefore, the manifold $\Phi(N)$ is a submanifold, and therefore a pre-submanifold of the manifold M. Thus, our proposition is fully proven. □

In particular, this proposition implies that for any coordinate neighbourhood U on a submanifold N of a manifold M, the map

$$\xi^{-1} : U^0 \to M,$$

inverse to a coordinate diffeomorphism

$$\xi : U \to \mathbb{R}^n,$$

is a diffeomorphism of the open set $U^0 = \xi(U) \subset \mathbb{R}^n$ into the manifold M and is therefore an n-dimensional surface of this manifold. Conversely, for any n-dimensional surface

$$\varphi : G \to M, \qquad G \subset \mathbb{R}^n,$$

that is a diffeomorphism of the set G onto the set $\varphi(G) \subset M$, the set $\varphi(G)$ is an n-dimensional submanifold of the manifold M. This submanifold is a coordinate neighbourhood of each of its points, and the corresponding coordinate diffeomorphism is the diffeomorphism φ^{-1}.

In particular, for every curve

$$\gamma : J \to M,$$

that is a diffeomorphism of the interval J onto the set $\gamma(J)$, this latter set is a one-dimensional submanifold of the manifold M. On the other hand, every regular curve $\gamma : J \to M$ that has no self-intersections on any closed interval $C \subset J$ is a homeomorphism of this interval into the manifold M (since the interval C is compact, and the manifold M is Hausdorff) and therefore on every interval $J_0 \subset C$ is a diffeomorphism of this interval into the manifold M. By the results of point §5.2, it follows that

Proposition 5.29. *for each point $p = \gamma(t_0)$ of an arbitrary regular curve $\gamma : J \to M$ there exists an interval J_0 containing the point t_0 such that the set $\gamma(J_0)$ is a one-dimensional submanifold of the manifold M.*

It is clear that

Proposition 5.30. *for each point $p = \gamma(t)$, $t \in J_0$, the tangent space $\gamma(J_0)_p$ of the submanifold $\gamma(J_0)$ to the point p is generated by the tangent vector $\dot{\gamma}(t)$.*

In those cases where this will not cause misunderstanding, we will identify each curve $\gamma : J \to M$ that is a diffeomorphism with the submanifold $\gamma(J)$.

Clearly, every open submanifold of M has the same dimension m as M itself. Conversely,

Proposition 5.31. *every submanifold N of M whose dimension n is equal to the dimension m of M is an open submanifold of M.*

Proof. Indeed, since $m = n$ and the inclusion map $\iota : N \subset M$ is regular, at each point $p \in N$ this map is a local diffeomorphism (see §5.3). Therefore, in particular, every point $p \in N$ has a neighbourhood in N, which is its neighbourhood in M. But this means that N is open in M. □

We obtain an important example of an $m-1$-dimensional submanifold by considering an arbitrary smooth function f on the manifold M. For each real number a, we will, following the accepted convention (and thereby admitting some vagueness in terminology), call the sub-premanifold $[f=a]$ of M consisting of all points $p \in M$ for which $f(p)=a$ the *level surface* of the function f corresponding to the given number a. We will call a level surface $[f=a]$ *regular* if the function f, considered as a smooth map $M \to \mathbb{R}$, is regular on $[f=a]$, i.e., if $df_p \neq 0$ for each point $p \in ([f=a])$. We will show that

Proposition 5.32. *any regular level surface $[f=a]$ is an $m-1$-dimensional submanifold of M.*

Proof. Indeed, let p be an arbitrary point on the level surface $[f=a]$ and let $x^1, \ldots, x^m$ be an arbitrary system of local coordinates on the manifold M at the point p. Since the level surface $[f=a]$ is, by assumption, regular, at least one partial derivative $\frac{\partial f}{\partial x^i}$ of the function f is nonzero at the point p. Let, for definiteness,

$$\left(\frac{\partial f}{\partial x^1}\right)_p \neq 0.$$

Then, by the theorem on the change of local coordinates, the functions $f, x^2, \ldots, x^m$ also form a system of local coordinates at the point p. Let U be the corresponding coordinate neighbourhood and $\xi : U \to \mathbb{R}^m$ the corresponding coordinate diffeomorphism. Let, further,

$$y^1 = x^2|_{[f=a]}, \ \ldots, \ y^{m-1} = x^m|_{[f=a]}$$

and

$$V = U \cap [f=a].$$

By definition, the functions $y^1, \ldots, y^{m-1}$ are smooth at any point $q \in V$. Moreover, every smooth function φ on $[f=a]$ at $q \in V$, being the restriction of some smooth function $\psi = u(x^1, \ldots, x^m)$ to U at q, satisfies the relation

$$\varphi = u(a, y^1, \ldots, y^{m-1}),$$

in some neighbourhood of point q, i.e., it depends smoothly on the functions $y^1, \ldots, y^{m-1}$ near q. Finally, the map

$$\eta : V \to \mathbb{R}^{m-1},$$

defined by the formula

$$\eta(q) = (y^1(q), \ \ldots, \ y^{m-1}(q)), \quad q \in V,$$

is related to the coordinate diffeomorphism $\xi : U \to \mathbb{R}^m$ by the relation

$$\xi(q) = (a, \eta(q)), \quad q \in V.$$

Therefore, this map is a diffeomorphism of the neighbourhood V onto the section of the set $\xi(U)$ by the hyperplane $t^1 = a$. Thus, it is proved that the functions $y^1, \ldots, y^{m-1}$ are local coordinates in the neighbourhood V of the point p on the surface $[f = a]$. Since the point p is arbitrary, this proves that the subpremanifold $[f = a]$ is a smooth manifold and therefore, by the statement proved above, a submanifold of the manifold M. □

At each point $p \in [f = a]$, the space $[f = a]_p$, tangent to the submanifold $[f = a]$, is an $m - 1$-dimensional subspace of M_p. Let $\iota : [f = a] \to M$ be the inclusion map. Since $f \circ \iota = a$, we have $d(f \circ \iota)_p = df_p \circ d\iota_p = 0$, i.e., $df_p(A) = 0$ for any vector $A \in [f = a]$. Since the covector df_p, by assumption, is nonzero, the vectors $A \in M_p$ for which $df_p(A) = 0$ form an $m-1$-dimensional subspace of M_p. Since the space $[f = a]_p$ is also $m - 1$-dimensional, it immediately follows that

Proposition 5.33. *a vector $A \in M_p$ belongs to the subspace $[f = a]_p$, i.e., it is tangent to the submanifold $f = a]$, if and only if when*

$$df_p(A) = 0.$$

We obtain an example of a regular level surface by considering the function on the space $\mathbb{R}^m$.

$$f = (t^1)^2 + \cdots + (t^m)^2.$$

For any $a > 0$, the corresponding level surface $[f = a]$ is a sphere of radius $\sqrt{a}$ centred at 0. Since

$$df = 2t^1 dt^1 + \cdots + 2t^m dt^m,$$

this level surface is regular for all $a > 0$. Thus,

Proposition 5.34. *spheres of the space $\mathbb{R}^m$ are its submanifolds.*

Now let $f^1, \ldots, f^r$ be an arbitrary finite system of smooth functions on a manifold M and let $a^1, \ldots, a^r$ be arbitrary real numbers. Consider the submanifold N of M consisting of all points $p \in M$ for which

$$f^1(p) = a^1, \ \ldots, \ f^r(p) = a^r,$$

holds, i.e., the intersection

$$[f^1 = a^1] \cap \cdots \cap [f^r = a^r]$$

of the level surfaces $[f^1 = a^1], \ \ldots, \ [f^r = a^r]$.

We call functions $f^1, \ldots, f^r$ *independent* at a point $\in M$ if the covectors $df_p^1, \ldots, df_p^r$ are linearly independent elements of the space $\mathcal{O}_1(p)$. It turns out that

Proposition 5.35. *if functions $f^1, \ldots, f^r$ are independent at each point p of the submanifold*

$$N = [f^1 = a^1] \cap \cdots \cap [f^r = a^r],$$

then this submanifold is an $m-r$-dimensional submanifold of M, and at any point $p \in N$ the subspace N_p of M_p is defined by the equations

$$df^1_p(A) = 0, \ \dots, \ df^r_p(A) = 0, \quad A \in M_p. \tag{5.36}$$

For $r = 1$, this statement reduces to the proposition about regular level surfaces already proven above. Arguing by induction, we assume that for $r-1$, the statement in question has already been proven and that, consequently, the submanifold

$$N' = [f^1 = a^1] \cap \dots \cap [f^{r-1} = a^{r-1}],$$

is an $m-r-+1$-dimensional submanifold of M. Consider the restriction f of the function f^r to the submanifolds N'. Since at each point $p \in N$ the subspace N'_p of M_p is defined, according to the induction hypothesis, by the equations

$$df^1_p(A) = 0, \ \dots, \ df^{r-1}_p(A) = 0, \quad A \in M_p, \tag{5.37}$$

then (by the independence condition) the covector

$$df_p = d(f^r \circ \iota)_p = df^r|_{N'_p}$$

is not identically equal to zero, and therefore the level surface $[f = a^r]$ of the function f on the manifold N' is regular. Consequently, this level surface (which obviously coincides with the set N) is an $m-r$-dimensional submanifold of the submanifold N, and therefore of the manifold M. To complete the proof, it remains to note that the vector $A \in M_p$ satisfies equations (5.36) if and only if it satisfies equations (5.37) and the equation $df_p(A) = 0$.

It is clear that if the functions $f^1, \dots, f^r$ are independent at a point $p \in M$, then they are also independent at any point in some neighbourhood U of the point p. Applying the just proven statement to the manifold $M = U$, we immediately obtain from here, that

Proposition 5.38. *For any smooth and independent functions $f^1, \dots, f^r$ at a point $p \in M$, there exists a neighbourhood U of this point such that the set N of all points $q \in U$ for which*

$$f^1(q) = a^1, \ \dots, \ f^r(q) = a^r,$$

where $a^1 = f^1(q), \dots, a^r = f^r(q)$, is an $m-r$-dimensional submanifold (of the neighbourhood U, and therefore of the manifold M) containing the point p.

Generally speaking, not every n-dimensional submanifold N of M can be represented as an intersection $[f^1 = a^1] \cap \dots \cap [f^r = a^r]$ of level surfaces $r = m-n$ independent functions on N. This can only be done *locally*, i.e.,

Proposition 5.39. *for any point $p \in N$, there exists on the manifold M a system of $r = m-n$ functions $f^1, \dots, f^r$ independent at the point p and a neighbourhood U of the point p such that the point $q \in U$ belongs to the submanifold N if and only if*

$$f^1(q) = a^1, \ \dots, \ f^r(q) = a^r,$$

where $a^1 = f^1(q), \dots, a^r = f^r(q)$.

In this case, we will say that the *submanifold N in a neighbourhood U is defined by the equations*

$$f^1 = a^1, \ \dots, \ f^r = a^r.$$

We will prove even more, namely, that

Proposition 5.40. *at any point $p \in N$, there exists on the manifold M a system*

$$x^1, \ \dots, \ x^m \tag{5.41}$$

of local coordinates such that in the corresponding coordinate neighbourhood U, the submanifold N is defined by the equations

$$x^{n+1} = 0, \ \dots, \ x^m = 0.$$

Proof. To this end, we consider an arbitrary system of local coordinates $x^1, \ \dots, \ x^m$ on the manifold M at a point p such that the functions

$$y^1 = x^1|_N, \ \dots, \ y^n = x^n|_N$$

are local coordinates at the point p on the submanifold N. Let

$$y^{n+1} = x^{n+1}|_N, \ \dots, \ y^m = x^m|_N$$

Since the functions $y^{n+1}, \ \dots, \ y^m$ are smooth at the point $p \in N$, they depend smoothly near this point on the functions $y^1, \ \dots, \ y^n$, i.e., there exist smooth functions $u^{n+1}, \ \dots, \ u^m$ of real variables $t^1, \dots, t^n$ such that near the point p on the submanifold N the following equalities hold:

$$y^{n+1} = u^{n+1}(y^1, \ \dots, \ y^n), \ \dots, \ y^m = u^m(y^1, \ \dots, \ y^n).$$

Let

$$\begin{aligned} x'^1 &= x^1, \\ \cdots &= \cdots \\ x'^n &= x^n, \\ x'^{n+1} &= x^{n+1} - u^{n+1}(y^1, \ \dots, \ y^n), \\ \cdots &= \cdots \\ x'^m &= x^m - u^m(y^1, \ \dots, \ y^n). \end{aligned}$$

The theorem on the change of local coordinates immediately implies that in some neighbourhood U' of the point p, the functions $x'^1, \ \dots, \ x'^m$ are local coordinates on the manifold M. Consider the functions $x'^{n+1}, \ \dots, \ x'^m$. Since these functions are obviously independent at the point p, the equations

$$x'^{n+1} = 0, \ \dots, \ x'^m = 0. \tag{5.42}$$

define in the neighbourhood U' an n-dimensional submanifold N' of the manifold M. From the above, it immediately follows that $N \subset N'$. Since $\dim N = \dim N'$, this is possible only when the submanifold N is open in the submanifold $N^{'}$, i.e., in other words, when the submanifold N and $N^{'}$ coincide in some neighbourhood U of the point p. Thus, in the neighbourhood U, the submanifold N is defined by equations (5.42). To complete the proof, we can denote the coordinates $x^{'1}, \dots, x^{'m}$ by the same symbols $x^1, \dots, x^m$. □

The construction of local coordinates (5.41) contains considerable arbitrariness and therefore, in certain specific situations, they can be further specialised.

For example, let us assume that, along with the submanifold N, we are given another submanifold $N^{'}$ (of dimension $n^{'}$), also containing the point p and possessing, in addition, the property that its tangent space $N^{'}_p$ at this point intersects the space N_p only at zero. Then, as we will now show,

Proposition 5.43. *the local coordinates (5.41) can be chosen such that they additionally possess the property that the manifold $N^{'}$ in a neighbourhood U is defined by the equations*

$$x^1 = 0, \ \dots, \ x^n = 0, \quad x^{n+n^{'}+1} = 0, \ \dots, \ x^m = 0.$$

Proof. Indeed, by changing, if necessary, the numbering of the coordinates $x^{n+1}, \dots, x^m$, we can assume, without loss of generality, that the coordinates (5.41) have the property that the functions $x^{n+1}|_{N'}, \dots, x^{n+n^{'}}|_{N'}$ are local coordinates on the submanifold $N^{'}$ at the point p. But then, repeating the arguments already used above in relation to the submanifold N, we immediately see that there exist smooth functions $v^1, \dots, v^n, v^{n+n^{'}+1}, \dots, v^m$ of real variables $t^1, \dots, t^{n^{'}}$ such that, firstly, the functions defined by the formula

$$x^{'i} = \begin{cases} x^i & \text{if} \quad n+1 \le i \le n+n^{'}, \\ x^i - v^i(x^{n+1}, \dots, x^{n+n^{'}}) & \text{if either} \quad 1 \le i \le n \quad \text{or} \quad n+n^{'}+ \le i \le m, \end{cases}$$

are local coordinates on the manifold M at the point p and, secondly, near this point the submanifold $N^{'}$ is defined by the equations

$$x^{'1} = 0, \ \dots, \ x^{'n} = 0, \quad x^{'n+n^{'}+1} = 0, \ \dots, \ x^{'m} = 0.$$

It is clear that, with respect to the submanifold N, the coordinates $x^{'1}, \dots, x^{'m}$ still have the properties of coordinates (5.41). Therefore, our assertion is completely proven. □

In the particular case when the submanifold $N^{'}$ is one-dimensional and has the form $\gamma(J)$, where $\gamma : J \to M$ is some regular curve on the manifold M, the condition

$$N_p \cap N^{'}_p = 0.$$

means that at the point p the curve γ is not tangent to the submanifold N, i.e., that its tangent vector at this point does not belong to the subspace N. Thus

Proposition 5.44. *for any submanifold N of manifold M, any point $p \in N$ and any regular curve $\gamma : J \to M$ passing through the point p and not tangent to submanifold N at this point, there exists on the manifold M a system of local coordinates $x^1, \dots, x^m$ at the point p such that in some neighbourhood of the point p the submanifold N is defined by equations $x^{n+1} = 0, \dots, x^m = 0$, and the curve γ is defined by equations $x^1 = 0, \dots, x^n = 0, x^{n+} = 0, \dots, x^m = 0$ (or, if you prefer, by equations $x^1 = 0, \dots, x^n = 0, x^{n+1} = 0, \dots, x^{m-1} = 0$).*

In conclusion of this section, to avoid possible misunderstandings, we note that in the theory of smooth manifolds, the term "submanifold" is often understood in a somewhat more general sense than ours. Specifically, in many questions, it is convenient not to assume that submanifolds are subspaces of the ambient manifold M. It can easily be shown that all the fundamental properties of submanifolds are fully preserved if by submanifolds of M we mean smooth manifolds N, located as sets in M and possessing the property that the corresponding inclusion maps $\iota : N \subset M$ are smooth (and therefore continuous) and regular. We will not need this general understanding of the term "submanifold" (since, in essence, we will henceforth only deal with compact submanifolds), and therefore we will not dwell on it further.

5.5 Submanifolds of Euclidean spaces

The study of submanifolds is especially simple in the case where the enclosing manifold is Euclidean space $\mathbb{R}^n$. For each submanifold M of the space $\mathbb{R}^n$, the tangent space M_p at an arbitrary point $p \in M$ is a subspace of the tangent space $(\mathbb{R}^n)_p$. But the latter space is naturally identified with the space $\mathbb{R}^n$ itself (see §4.6). Thus, for each manifold $M \subset \mathbb{R}^n$, all spaces M_p corresponding to all possible points $p \in M$ can be considered as subspaces of the same space $\mathbb{R}^n$.

This fact, while significantly simplifying the study of manifolds M embedded in the space $\mathbb{R}^n$, simultaneously significantly enriches their theory.

Let p_0 be an arbitrary point of the manifold $M \subset \mathbb{R}^n$ and let $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ be a basis of the space $\mathbb{R}^n$ such that the first m of its vectors $\boldsymbol{e}_1, \dots, \boldsymbol{e}_m$ form a basis of the subspace $M_{p_0} \subset \mathbb{R}^n$. According to the results of §5.4, the coordinates $t^1, \dots, t^n$ of points in the space $\mathbb{R}^n$ in the basis $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ have the property that the constraints

$$x^1 = t^1|_M, \ \dots, \ x^m = t^m|_M \tag{5.45}$$

of the first m of these coordinates constitute a system of local coordinates at the point p_0 on the manifold M. In this section we will call each neighbourhood U of the point p_0 in the manifold M, on which functions (5.45) constitute a system of local coordinates, a *normal* neighbourhood of the point p_0 (with respect to the given basis $\boldsymbol{e}_1, \dots, \boldsymbol{e}_m$ of the space $\mathbb{R}^n$).

For each such normal neighbourhood U we consider the set U of all pairs $(p, \boldsymbol{u})$, where p is an arbitrary point of the neighbourhood U and $\boldsymbol{u}$ is an arbitrary vector of the space $\mathbb{R}^n$ orthogonal to the subspace $M_p \subset \mathbb{R}^n$ (we will call such vectors $\boldsymbol{u} \in \mathbb{R}^n$ vectors *orthogonal to the manifold M* at the point p). By definition, a vector u is orthogonal to the manifold M at the point p if and only if

$$\left(\left(\frac{\partial}{\partial x^i}\right)_p, \boldsymbol{u}\right) = 0.$$

for any $i = 1, \ldots, m$. But, according to the formula in §3.6, the vectors $\left(\frac{\partial}{\partial x^i}\right)_p$, $i =, \ldots, m$ (considered as vectors of the space $\mathbb{R}^n$), are related to the vectors $\boldsymbol{e}_j = \left(\frac{\partial}{\partial t^j}\right)_p$, $j = 1, \ldots, n$, by the relations

$$\left(\frac{\partial}{\partial x^i}\right)_p = \left(\frac{\partial t^j}{\partial x^i}\right)_p \cdot \boldsymbol{e}_j, \quad i = 1, \ldots, m, \ j = 1, \ldots, n.$$

On the other hand, according to relations (5.45),

$$\left(\frac{\partial t^j}{\partial x^i}\right)_p = \delta_i^j$$

for any $i, j = 1, \ldots, m$, so that

$$\left(\frac{\partial}{\partial x^i}\right)_p = \boldsymbol{e}_j + \sum_{j=m+1}^{n} \left(\frac{\partial t^j}{\partial x^i}\right)_p \cdot \boldsymbol{e}_j,$$

Therefore, the vector $\boldsymbol{u} \in \mathbb{R}^n$ is orthogonal to the manifold M at the point p if and only if

$$(\boldsymbol{e}_i, \boldsymbol{u}) + \sum_{j=m+1}^{n} \left(\frac{\partial t^j}{\partial x^i}\right)_p \cdot (\boldsymbol{e}_j, \boldsymbol{u}) = 0$$

for any $i = 1, \ldots, m$.

To simplify the formulae, let us assume that the basis $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ of the space $\mathbb{R}^n$ under consideration is orthonormal. Then the numbers $(\boldsymbol{e}_1, \boldsymbol{u}), \ldots, (\boldsymbol{e}_n, \boldsymbol{u}))$ are the coordinates $u^1, \ldots, u^n$ of the vector $\boldsymbol{u}$ in this basis and therefore the previous relations will be rewritten as follows:

$$u^i + \sum_{j=m+1}^{n} \left(\frac{\partial t^j}{\partial x^i}\right)_p u^j = 0, \quad i = 1, \ldots, m. \tag{5.46}$$

Now let

$$\xi : U \to U^0$$

be the coordinate diffeomorphism corresponding to the local coordinates (5.45), and let

$$\eta : \mathbb{R}^n \to \mathbb{R}^{n-m}$$

be the natural projection of the space $\mathbb{R}^n$ onto its subspace $\mathbb{R}^{n-m}$ spanned by the vectors $\boldsymbol{e}_{m+1}, \ldots, \boldsymbol{e}_n$. From relations (5.46) it immediately follows that being defined by the formula

$$\xi_*(p, \boldsymbol{u}) = (\xi(p), \eta(\boldsymbol{u}))$$

the map

$$\xi_* : U_* \to \mathbb{R}^n$$

is a bijective map of the set U_* onto an open subset $U^0 \times \mathbb{R}^{n-m}$ of the space $\mathbb{R}^n$. We introduce into the set U_* the identification topology defined by the map ξ_*. By definition, this topology depends on the choice of an orthonormal basis $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ of the space $\mathbb{R}^n$. However, it is easy to see that for any other choice of this basis (subject only to the condition that the neighbourhood U is normal with respect to it), we obtain the same topology on the set U_*. In other words, the topology of the set U_* under consideration does not actually depend on the choice of the basis $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$. Note that in this topology, the map ξ_* is a homeomorphism.

Let N_* be the set of all pairs of the form $(p, \boldsymbol{u})$, where p is an arbitrary point of the manifold M, and $\boldsymbol{u}$ is an arbitrary vector of the space $\mathbb{R}^n$ orthogonal to the manifold M at the point p. This set is the union of subsets of the form U_* corresponding to all possible normal (with respect to some neighbourhood-dependent basis of the space $\mathbb{R}^n$) neighbourhoods U of points of the manifold M. Since the topologies of sets U_* constructed above do not depend on the choice of bases of the space $\mathbb{R}^n$, it follows directly that for any two normal neighbourhoods $U, V \subset M$ the topologies induced by the topologies of sets U_* and V_* on the set $(U \cap V)_* = U_* \cap V_*$ coincide with each other, i.e., the intersection $U_* \cap V_*$ is a subspace of each of the spaces U_* and V_*. Therefore (see §1.1) on the set N_* there exists a unique topology with respect to which this set is the free union of all subspaces of the form U_*.

A real function F defined on the space N_* is considered smooth if, for any normal neighbourhood $U \subset M$, the function $\circ\xi_*^{-1}$ on the product $U^0 \times \mathbb{R}^{n-m}$, where ξ_* is the homeomorphism $U_* \to U^0 \times \mathbb{R}^{n-m}$ constructed above, is a smooth function on $U^0 \times \mathbb{R}^{n-m}$. It is easily verified that the set $\mathscr{O}(N_*)$ of all smooth functions on N_* is a smoothness on N_* and that the space N_* endowed with this smoothness is a smooth manifold. For example, we can take subspaces of U_* as coordinate neighbourhoods in N_*, and maps

$$\xi_* : U_* \to \mathbb{R}^n$$

as coordinate diffeomorphisms. The local coordinates corresponding to this diffeomorphism

$$x_*^1, \ \ldots, \ x_*^n \tag{5.47}$$

are defined by the formula

$$x_*^1(p, \boldsymbol{u}) = x^1(p), \ \ldots, \ x_*^m(p, \boldsymbol{u}) = x^m(p), \ x_*^{m+1}(p, \boldsymbol{u}) = u^{m+1}, \ \ldots, \ x_*^n(p, \boldsymbol{u}) = u^n,$$
$$(p, \boldsymbol{u}) \in U_* \tag{5.48}$$

where $x^1, \ldots, x^m$ are the local coordinates (5.45), i.e., the first m coordinates in the basis $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ of the point $\in M$, considered as a vector in the space $\mathbb{R}^n$, and $u^{m+1}, \ldots, u^n$ are the last $n-m$ coordinates of the vector $\boldsymbol{u}$ (in the same basis $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$).

We will call the constructed n-dimensional smooth manifold N_* the *manifold of vectors orthogonal to the manifold M*. Its points of the form $(p, \mathbf{0})$, $p \in M$, obviously form a submanifold that is naturally diffeomorphic to the manifold M. We will identify this submanifold with the manifold M and, thus, assume that $M \subset N_*$.

Remark 5.49. The construction of the manifold N_* is very similar to the construction of the manifold M_* presented in §4.6. It can be shown that both of these constructions are special cases of a certain general construction related to arbitrary vector bundles. We do not need this construction in its full generality, and therefore we will not present it here.

Consider a map

$$\rho : N_* \to \mathbb{R}^n$$

of the manifold N_* into the space $\mathbb{R}^n$ that associates to each point $(p, \boldsymbol{u}) \in N_*$ a vector

$$\rho(p, \boldsymbol{u}) = p + \boldsymbol{u},$$

i.e., the vector $\boldsymbol{u}$, measured from the point p. Clearly, the map ρ is smooth. On the submanifold $M \subset N_*$, this map is the identity. Let us express the map ρ in local coordinates.

Let $x_*^1, \ldots, x_*^n$ be the coordinates (5.47) of the point $(p, \boldsymbol{u})$. According to the above, the point p, considered as a vector of the space $\mathbb{R}^n$, has (in the basis $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$) coordinates

$$t^1 = x_*^1, \ \ldots, \ t^m = x_*^m, \ t^{m+1} = t^{m+1}(x_*^1, \ \ldots, \ x_*^n), \ \ldots \ t^n = t^n(x_*^1, \ \ldots, \ x_*^n),$$

where $t^{m+1} = t^{m+1}(x_*^1, \ \ldots, \ x_*^n), \ \ldots \ t^n = t^n(x_*^1, \ \ldots, \ x_*^n)$ are some smooth functions, and the vector $\boldsymbol{u}$ are the coordinates

$$u^1 = -\sum_{k=m+1}^{n} \left(\frac{\partial t^k}{\partial x^1}\right)_p u^k, \ \ldots, \ u^m = -\sum_{k=m+1}^{n} \left(\frac{\partial t^k}{\partial x^m}\right)_p u^k, \ u^{m+1} = x_*^{m+1}, \ \ldots, \ u^n = x_*^n.$$

Therefore, the coordinates $v^1, \ \ldots, \ v^n$ of the point $\boldsymbol{v} = \rho(p, \boldsymbol{u})$ are expressed by the formulae

$$\begin{aligned}
v^1 &= x_*^1 - \sum_{k=m+1}^{n} \frac{\partial t^k}{\partial x_*^1}(x_*^1, \ \ldots, \ x_*^m) x_*^k,\\
\cdots &= \cdots\\
v^m &= x_*^m - \sum_{k=m+1}^{n} \frac{\partial t^k}{\partial x_*^m}(x_*^1, \ \ldots, \ x_*^m) x_*^k,\\
v^{m+1} &= t^{m+1}(x_*^1, \ \ldots, \ x_*^m) + x_*^{m+1},\\
\cdots &= \cdots\\
v^n &= t^n(x_*^1, \ \ldots, \ x_*^m) + x_*^n.
\end{aligned}$$

These formulae represent the map ρ in local coordinates. Differentiating them, we immediately obtain that

$$\frac{\partial v^i}{\partial x_*^j} = \begin{cases} \delta^i_j - \sum_{k=m+1}^n \frac{\partial^2 t^k}{\partial x_*^i \partial x_*^j}(x_*^1, \ \ldots, \ x_*^m)x_*^k, & \text{if} \quad 1 \le i, j \le m; \\ -\frac{\partial t^j}{\partial x_*^i}(x_*^1, \ \ldots, \ x_*^m), & \text{if} \quad 1 \le i \le m \text{ and } m+1 \le j \le n; \\ -\frac{\partial t^i}{\partial x_*^j}(x_*^1, \ \ldots, \ x_*^m), & \text{if} \quad m+1 \le i \le n \text{ and } 1 \le j \le m; \\ \delta^i_j, & \text{if} \quad m+1 \le i, j \le n. \end{cases}$$

Thus,

$$\det \left| \frac{\partial v^i}{\partial x_*^j} \right| = \begin{vmatrix} A & -T \\ T' & E_{n-m} \end{vmatrix},$$

where A is a matrix of order $m \times m$ with elements

$$\delta^i_j - \sum_{k=m+1}^n \frac{\partial^2 t^k}{\partial x_*^i \partial x_*^j}(x_*^1, \ \ldots, \ x_*^m)x_*^k,$$

T is a matrix of order $m \times (n-m)$ with elements

$$\frac{\partial t^j}{\partial x_*^i}(x_*^1, \ \ldots, \ x_*^m),$$

T' is the transpose of T and E_{n-m} is the identity matrix of order $n-m$.

Multiplying this determinant on the right by the determinant

$$\begin{vmatrix} E_m & 0 \\ -T' & E_{n-m} \end{vmatrix} = 1,$$

we obtain that

$$\det \left| \frac{\partial v^i}{\partial x_*^j} \right| = \begin{vmatrix} A + TT' & -T \\ 0 & E_{n-m} \end{vmatrix},$$

i.e, that

$$\det \left| \frac{\partial v^i}{\partial x_*^j} \right| = |A + TT'|. \tag{5.50}$$

In particular,

$$\det \left| \left(\frac{\partial v^i}{\partial x_*^j} \right)_p \right| = |E + TT'|.$$

at each point $p = (p, \mathbf{0}) \in M$. But it is easy to see that

$$|E + TT'| > 0$$

(since a quadratic form with matrix $E + TT'$, being the sum of a positive-definite form with matrix E and a non-negative-definite form with matrix TT', is a positive-definite form). Thus

$$\det\left|\left(\frac{\partial v^i}{\partial x_*^j}\right)_p\right| \neq 0.$$

This means that at each point $p \in M$, the map ρ is regular and therefore is a diffeomorphism near this point. We have thus proved that

Proposition 5.51. *the map $\rho : N_* \to \mathbb{R}^n$ is a smooth map that is the identity on the submanifold $M \subset N_*$ and, at each point of this submanifold, is a local diffeomorphism from the manifold N_* to the space $\mathbb{R}^n$.*

We call an open set $V \subset \mathbb{R}^n$ *tubular* if there exists an open set V^0 in the manifold N_* such that the map ρ takes the set V^0 to the set V and, considered as a map $V^0 \to V$, is a diffeomorphism of this set onto the set V. For each such set, a diffeomorphism is defined

$$\sigma : V \to V^0,$$

inverse to the diffeomorphism $\rho|_{V^0}$. It is clear that any open set $V^{'} \subset V$ contained in some tubular set V is also a tubular set, and the corresponding diffeomorphism $\sigma^{'}$ is the restriction of the diffeomorphism σ. The intersection $V_1 \cap V_2$ of two tubular sets V_1 and V_2 is, generally speaking, not a tubular set. However, if the intersection $V_1^0 \cap V_2^0$ of the corresponding subsets V_1^0 and V_2^0 of the manifold N_* is non-empty, then the set $V = \rho(V_1^0 \cap V_2^0)$ (contained in the intersection $V_1 \cap V_2$) is open in $\mathbb{R}^n$ and is a tubular set. The corresponding diffeomorphism $\sigma : V \to V_1^0 \cap V_2^0$ is then the restriction of each of the diffeomorphisms $\sigma_1 : V_1 \to V_1^0$ and $\sigma_2 : V_2 \to V_2^0$ (so that, in particular, the diffeomorphisms σ_1 and σ_2 on the set V coincide).

It is clear that a similar statement holds for any finite number of tubular sets.

According to the assertion proved above, every point $p \in M$ has at least one tubular neighbourhood in $\mathbb{R}^n$. If M is closed (i.e., is a closed subset of $\mathbb{R}^n$), then the tubular neighbourhoods of all its points, together with the set $\mathbb{R}^n \setminus M$, form an open covering of $\mathbb{R}^n$. Since $\mathbb{R}^n$ is paracompact, some locally finite covering Γ can be inscribed in this covering. Since the space $\mathbb{R}^n$ is normal, there exists an open covering Δ of it that has the same index set as the covering Γ and is such that the closure of each element of Δ is contained in the corresponding element of Γ (see §1.3). Let $\{V_\alpha : \alpha \in A\}$ be the subfamily of Δ consisting of all its elements that intersect the manifold M, and let $\{U_\alpha : \alpha \in A\}$ be the family of corresponding elements of Γ. It is clear that all sets U_α are tubular sets. We denote the diffeomorphism $U_\alpha \to U_{\underline{\alpha}}^0 \subset N_*$ corresponding to the set U_α by the symbol σ_α. Since, by hypothesis, $\overline{V}_\alpha \subset U_\alpha$, the set V_α is also a tubular set, and the corresponding diffeomorphism $V_\alpha \to V_\alpha^0$ is the restriction of the diffeomorphism σ_α.

Consider the set W of all points $\boldsymbol{t}$ in the space $\mathbb{R}^n$, belonging to the set

$$V = \cup_{\alpha \in A} V_\alpha$$

and possessing the property that for any indices $\alpha', \alpha'' \in A$ the following equality holds

$$\sigma_{\alpha'}(\boldsymbol{t}) = \sigma_{\alpha''}(\boldsymbol{t})$$

whenever $\boldsymbol{t} \in \overline{V}_{\alpha'} \cap \overline{V}_{\alpha''}$. It is clear that $M \subset W$.

We will show that

Proposition 5.52. *the set W is open in the space $\mathbb{R}^n$.*

Proof. Indeed, let $\boldsymbol{t}_0$ be an arbitrary point of this set. Since the family $\{\overline{V}_\alpha\}$ is locally finite, the point $\boldsymbol{t}_0$ has a neighbourhood that intersects only finitely many elements of this family. Therefore,

firstly there are only finitely many elements $\overline{V}_{\alpha_1}, \ldots, \overline{V}_{\alpha_r}$ of the family $\{\overline{V}_\alpha\}$ containing the point $\boldsymbol{t}_0$, and

secondly the point $\boldsymbol{t}_0$ has a neighbourhood $U(\boldsymbol{t}_0)$ that does not intersect with any other elements of this family.

In this case, we can obviously assume that $U(\boldsymbol{t}_0) \subset V$. Let

$$U^0_{\alpha_1} = \sigma_{\alpha_1}(U_{\alpha_1}), \ \ldots, \ U^0_{\alpha_r} = \sigma_{\alpha_r}(U_{\alpha_r}).$$

Since $\boldsymbol{t}_0 \in W$ and $\boldsymbol{t}_0 \in \overline{V}_{\alpha_1}, \ldots, \overline{V}_{\alpha_r}$, then $\sigma_{\alpha_1}(\boldsymbol{t}_0) = \cdots = \sigma_{\alpha_r}(\boldsymbol{t}_0)$ and therefore the intersection $U^0_{\alpha_1} \cap \cdots \cap U^0_{\alpha_r}$ is not empty. Let

$$V(\boldsymbol{t}_0) = U(\boldsymbol{t}_0) \cap \rho(U^0_{\alpha_1} \cap \cdots \cap U^0_{\alpha_r}).$$

The set $V(\boldsymbol{t}_0)$ is open in $\mathbb{R}^n$ and contains the point $\boldsymbol{t}_0$, i.e., t is a neighbourhood of it. Moreover, it is contained in the set V, intersects only the elements $\overline{V}_{\alpha_1}, \ldots, \overline{V}_{\alpha_r}$ of the family $\{\overline{V}_\alpha\}$, and has the property that all maps $\sigma_{\alpha_1}. \ldots, \sigma_{\alpha_r}$ on this set coincide (since $V(\boldsymbol{t}_0) \subset \rho(U^0_{\alpha_1} \cap \cdots \cap U^0_{\alpha_r})$). Let $\boldsymbol{t}$ be an arbitrary point in the neighbourhood $V(\boldsymbol{t}_0)$ and let $\alpha', \alpha'' \in A$ be indices such that $\boldsymbol{t} \in \overline{V}_{\alpha'} \cap \overline{V}_{\alpha''}$. Since the neighbourhood $V(\boldsymbol{t}_0)$ intersects only the elements $\overline{V}_{\alpha_1}, \ldots, \overline{V}_{\alpha_r}$, the indices α' and α'' must necessarily be contained among the indices $\alpha^0_1 \ \ldots, \alpha_r$. Since all maps $\sigma_{\alpha_1}, \ldots, \sigma_{\alpha_r} A$ onto neighbourhoods $V(\boldsymbol{t}_0)$ coincide, it follows that $\sigma_{\alpha'}(\boldsymbol{t}) = \sigma_{\alpha''}(\boldsymbol{t})$. Consequently, $\boldsymbol{t} \in W$. We have thus proved that $V(\boldsymbol{t}_0) \subset W$. Thus, every point $\boldsymbol{t}_0 \in W$ has a neighbourhood $V(\boldsymbol{t}_0)$ contained in the set W. Therefore, the set W is open. □

It is clear that all maps σ_α, $\alpha \in A$, on the set W are compatible with each other, i.e., they are restrictions of some single map $\sigma : W \to N_*$. This map is obviously a diffeomorphism of the set W to some open subset W^0 of the manifold N_*. In this case, $\rho(W^0) = W$ and the map $\rho|_{W_0}$, considered as a map $W^0 \to W$, is inverse to the diffeomorphism σ. This means that the set W is a tubular set.

Thus, we have proved that

Proposition 5.53. *every closed submanifold M of the Euclidean space $\mathbb{R}^n$ has a tubular neighbourhood.*

Together with the constructed neighbourhood W, any neighbourhood of the manifold M contained in the neighbourhood W will also be a tubular neighbourhood. Consequently,

Proposition 5.54. *the tubular neighbourhoods of a closed submanifold $M \subset \mathbb{R}^n$ constitute a fundamental system of its neighbourhoods.*

Let W be an arbitrary tubular neighbourhood of M in $\mathbb{R}^n$, and let W^0 be a neighbourhood of M in N_* whose image is the neighbourhood W. For any point $p \in M$, we consider the set ${}_pW$ of all points in $\mathbb{R}^n$ that have the form $\rho(p, \boldsymbol{u})$, where $(p, \boldsymbol{u}) \in W^0$. Clearly,

1) the set ${}_pW$ is diffeomorphic to some open set (depending on p) in $\mathbb{R}^{n-m}$ (containing the point $\mathbf{0}$);

2) the sets ${}_pW$ corresponding to distinct points $p \in M$ are pairwise disjoint;

3) the union of the sets ${}_pW$ corresponding to all possible points $pinM$ coincides with the neighbourhood W.

Properties 2) and 3) mean that the neighbourhood W is in some sense "foliated" into sets ${}_pW$. Therefore, for each point $p \in M$, we will call the set ${}_pW$ a *fibre* of the tubular neighbourhood W over the point p.

5.6 Embedding compact manifolds into Euclidean space. Sard's theorem

In order for a smooth manifold M to be diffeomorphic to some submanifold of the Euclidean space $\mathbb{R}^n$, it is necessary that this manifold be separable. It turns out that this necessary condition is also sufficient, i.e.

Proposition 5.55. *any separable smooth manifold M is diffeomorphic to a submanifold of the space $\mathbb{R}^n$, where n is some sufficiently large integer (depending on the manifold M).*

We will first prove this embedding theorem for the simplest, but at the same time most important case, when the given manifold M is compact.

To this end, we note that due to the compactness of the manifold M, it can be covered by a finite number of coordinate neighbourhoods $U_1, \ldots U_s$. Let $x_i^1, \ldots, x_i^m$ be local coordinates defined in a neighbourhood U_i, $i = 1, \ldots, s$, and let $\xi_i : U_i \to \mathbb{R}^m$ be the corresponding coordinate diffeomorphism. As is easy to see, we can assume without loss of generality that each diffeomorphism ξ_i maps the neighbourhood U_i onto the unit ball $|\boldsymbol{t}| < 1$ of the space $\mathbb{R}^m$ and is the restriction of some coordinate diffeomorphism (which we will also denote by the symbol ξ_i, and the local coordinates corresponding to it by the symbols $x_i^1, \ldots, x_i^m$), defined in some larger neighbourhood V_i and mapping this neighbourhood onto the ball $|\boldsymbol{t}| < 3$ of the space $\mathbb{R}^m$. (In this section we will call coordinate neighbourhoods U_i that have these properties *special*.)

Let $F(t)$ be a smooth non-negative function on the space $\mathbb{R}^m$, equal to one on the ball $|\boldsymbol{t}| \leq 1$, less than one outside this ball and equal to zero outside the ball $|\boldsymbol{t}| < 2$. Assuming that $x_i^1 = 1$ on U_i, for each $i = 1, \dots, s$, we define numerical functions $f_i^0, f_i^1, \dots, f_i^m$ on the manifold M by setting

$$f_i^k(p) = \begin{cases} x_i^k(p)F(\xi(p)), & \text{if} \quad p \in V_i, \\ 0, & \text{if} \quad p \notin V_i, \end{cases}$$

for any point $p \in M$ and any number $k = 0, 1, \dots, m$. Clearly, these functions are smooth on M.

The functions f_i^k, $i = 1, \dots, s$; $k = 0, 1, \dots, m$, taken in an arbitrary (but fixed) order, define a certain smooth map Φ of the manifold M into the space $\mathbb{R}^n$, where $n = s(m+1)$ is the total number of all functions f_i^k. We will show that this map is a diffeomorphism of the manifold M onto the submanifold $\Phi(M)$ of the space $\mathbb{R}^n$. As we know (see §5.4), for this it is sufficient to show that the map Φ is regular and is a homeomorphism of the manifold M onto the subspace $\Phi(M)$ of the space $\mathbb{R}^n$.

Let p be an arbitrary point of M, and let i be a number such that $p \in U_i$. Consider the map Φ_i of M into the space $\mathbb{R}^n$ defined by the functions $f_i^1, \dots, f_i^m$. Clearly, Φ is regular at p if Φ_i is regular at this point. But on the neighbourhood U_i, Φ_i obviously coincides with the coordinate diffeomorphism ξ_i and is therefore regular. Thus, the regularity of Φ on M is completely proved.

Since the manifold M is compact and the space $\mathbb{R}^n$ is Hausdorff, to prove that the map Φ is homeomorphic it is sufficient to prove that this map is injective.

Let p and q be points of M such that $\Phi(p) = \Phi(q)$. Consider the neighbourhood U_i containing the point p. Since $f_i^0 < 1$ outside U_i, and $f_i^0(q) = f_i^0(p) = 1$, we have $q \in U_i$. Since $\Phi(p) = \Phi(q)$, we have, in particular, $\Phi_i(p) = Phi_i(q)$. But $\Phi_i|_{U_i} = \xi_i$. Consequently, $\xi_i(p) = \xi_i(q)$ and therefore $p = q$.

Thus, the embedding theorem for compact manifolds is completely proved.

When trying to extend this proof to the case of an arbitrary separable manifold M, we immediately encounter the impossibility of finding a finite covering $\{U_i\}$ with the properties we need. However, since any separable manifold is paracompact and has a countable base, it is easy to see that there exists a countable and locally finite covering $\{U_i\}$ consisting of special coordinate neighbourhoods U_i, $i = 1, 2, \dots$. The functions f_i^k, $k = 0, 1, \dots, m$; $i = 1, 2, \dots$, constructed for this covering define (taken in an arbitrary but fixed order) a certain map Φ of the manifold M no longer into the space $\mathbb{R}^n$, but into the space $\mathbb{R}^\infty$, the points of which are infinite sequences $(t^1, t^2, \dots, t^n, \dots)$ of real numbers, only a finite number of which are non-zero. It is clear that this map is

smooth in the sense that for any $i = 1, 2, \dots$, the function φ^i on the manifold M, taking at the point $p \in M$ a value equal to the i-th coordinate t^i of the point $\Phi(p) \in \mathbb{R}^\infty$, is a smooth function on M (since this function is nothing other than one of the functions f_j^k);

regular in the sense that for any point $p \in M$ there exists a neighbourhood

U and a set $i_1, \ldots, i_m$ of indices i, such that the functions $\varphi^{i_1}|_U, \ldots, \varphi^{i_m}$ constitute a system of local coordinates on the neighbourhood U;

coordinate bounded in the sense that for any $i = 1, 2, \ldots$ ... the function φ_i is bounded;

locally finite-dimensional in the sense that for any point $p \in M$ there exists a neighbourhood U on which only finitely many functions φ^i are non-zero.

Furthermore, it is proved verbatim, just as for compact manifolds, that the map Φ is injective. Thus, instead of an embedding theorem, we obtain for now only that

Proposition 5.56. *for any separable manifold M, there exists a smooth, regular, coordinate-bounded, locally finite-dimensional, and injective map of it into the space $\mathbb{R}^\infty$.*

To derive the embedding theorem from this proposition, we need a statement about smooth maps of manifolds into manifolds of higher dimensions, which is of independent interest. On the other hand, in what follows, we will need a similar statement about maps of manifolds of the same number of dimensions. To avoid unnecessary repetition, I have presented both of these statements here.

Let M and N be two smooth manifolds of dimensions m and n, respectively, and let

$$\Phi : M \to N$$

be an arbitrary smooth map of M into N. We will show that

Proposition 5.57. *if $m = n$ and if M is separable, then the image of the set of all its critical points under Φ (see §4.6) is a set of the first category (in the sense of Baire-Hausdorff) in N.*

We will first prove this statement (known as Sard's theorem) for the special case where the manifolds M and N are the Euclidean space $\mathbb{R}^m$. Since the space $\mathbb{R}^m$ can be represented as a union of a countable number of cubes, to prove Sard's theorem in this particular case it suffices to prove that

Proposition 5.58. *for any smooth map $\varphi : \mathbb{R}^m \to \mathbb{R}^m$, the image $\varphi(K)$ of the set K of all its critical points belonging to the unit cube $I^m : 0 \le t^i \le 1$, $i = 1, \ldots, m$, is nowhere dense in the space $\mathbb{R}^m$.*

Moreover, since the set K is obviously closed and therefore compact, its image $\varphi(K)$ is closed in the space $\mathbb{R}^m$. Therefore, it suffices for us to prove that

Proposition 5.59. *the subset $\varphi(K)$ of the space $\mathbb{R}^m$ has no interior point.*

To this end, we note that, due to the smoothness of the map φ and the compactness of the cube I^m, the map φ uniformly satisfies the Lipschitz condition on this cube, i.e., there exists a number $a > 0$ such that

$$|\varphi(\boldsymbol{v} - \varphi(\boldsymbol{u})| \le a|\boldsymbol{v} - \boldsymbol{u}| \tag{5.60}$$

for any points $\boldsymbol{u}, \boldsymbol{v} \in I^m$. Moreover, it follows directly from Taylor's formula that for any point $\boldsymbol{u} \in \mathbb{R}^m$, the differential $d\varphi_{\boldsymbol{u}}$ of the map φ at this point, considered as a linear map $\mathbb{R}^m \to \mathbb{R}^m$, has the property that for every point $\boldsymbol{v} \in \mathbb{R}^m$, the following estimate holds:

$$|\varphi(\boldsymbol{v}) - \varphi(\boldsymbol{u}) - d\varphi_{\boldsymbol{u}}(\boldsymbol{v})| \leq b(|\boldsymbol{v} - \boldsymbol{u}|), \tag{5.61}$$

where $b(\tau)$ is a positive monotone function of the variable $\tau > 0$ such that

$$\frac{b(\tau)}{\tau} \to 0 \quad \text{as} \quad \tau \to 0.$$

In the case when $\boldsymbol{u} \in K$, all vectors $\varphi(\boldsymbol{u}) + d\varphi_{\boldsymbol{u}}(\boldsymbol{v})$, $\boldsymbol{v} \in \mathbb{R}^m$, belong to some $m-1$-dimensional hyperplane L of the space $\mathbb{R}^m$ passing through the point $\varphi(\boldsymbol{u})$, and for each point $\boldsymbol{v} \in \mathbb{R}^m$ the distance of the point $\varphi(\boldsymbol{v})$ from this hyperplane does not exceed the number $b(|v-u|)$. On the other hand, the distance of the point $\varphi(\boldsymbol{v})$ from the point $\varphi(\boldsymbol{u})$ does not exceed the number $a(|v-u|)$. Therefore, if $|v-u|) < \varepsilon$, then the point $\varphi(\boldsymbol{v})$ belongs to a body cut out from a ball of radius $a\varepsilon$ with centre at the point $\varphi(\boldsymbol{u})$ by two hyperplanes parallel to the hyperplane L and separated from it by a distance $b(\varepsilon)$. It is obvious that the volume of this body does not exceed the volume of a cylinder of height $2b(\varepsilon)$, the base of which is an $n-1$-dimensional ball of radius $a\varepsilon$, i.e., it does not exceed the number $c \cdot b(\varepsilon) \cdot \varepsilon^{m-1}$, where c is some constant (depending on a and m).

Let us now partition the cube I^m, m by systems of parallel hyperplanes into h^m cubes with edge $1/h$ and diagonal $\varepsilon = \sqrt{m}/h$. According to what has just been proved, each cube of this partition, containing at least one point of the set K, is mapped to some solid whose volume does not exceed the number $c \cdot b(\varepsilon) \cdot \varepsilon^{m-1}$. Therefore, the image $\varphi(K)$ of the entire set K is certainly contained in a solid whose volume does not exceed the number

$$c \cdot b(\varepsilon) \cdot \varepsilon^{m-1} \cdot h^m = cm^{m/2} \cdot \frac{b(\varepsilon)}{\varepsilon}. \tag{5.62}$$

Since this number tends to zero as $\varepsilon \to 0$, it follows that the set $\varphi(K)$ cannot contain any even arbitrarily small balls of the space $\mathbb{R}^m$, and therefore none of its points can be its interior points.

Therefore, in the special case when $M = N = \mathbb{R}^m$, Sard's theorem is completely proven.

Let us now consider the general case. Let p_0 be an arbitrary point of the manifold M and let V_{p_0} be an arbitrary coordinate neighbourhood of the point $\Phi(p_0)$ in the manifold N. Since the map Φ is continuous, the point p_0 has a coordinate neighbourhood U_{p_0} such that $\Phi(U_{p_0}) \subset V_{p_0}$. Let $\xi : U_{p_0} \to \mathbb{R}^m$ and $\eta : V_{p_0} \to \mathbb{R}^m$ be coordinate diffeomorphisms of the neighbourhoods U_{p_0} and V_{p_0} into the space $\mathbb{R}^m$, respectively. Without loss of generality, we can obviously assume that the diffeomorphism ξ maps the neighbourhood U_{p_0} onto the entire space $\mathbb{R}^m$. Consider the map

$$\varphi = \eta \circ \Phi \circ \xi^- : \mathbb{R}^m \to \mathbb{R}^m.$$

It is clear that a point $p \in U_{p_0}$ is a critical point of the map Φ if and only if the point $\xi(p_0) \in \mathbb{R}^m$ is a critical point of the map φ. Similarly, a point $q \in V_{p_0}$ is an image under the map Φ of some critical point of this map belonging to the neighbourhood U_{p_0} if and only if the point $\eta(q) \in \mathbb{R}^m$ is an image under the map φ of some critical point of the latter map. By the already proven special case of Sard's theorem, it follows that the image $\Phi(K_{p_0})$ of the set K_{p_0} of critical points of the map Phi belonging to the neighbourhood U_{p_0} is a subset of the first category (the neighbourhood V_{p_0}, and therefore of the manifold N). On the other hand, since the manifold is, by assumption, separable, it has a countable base and therefore one can find in it a countable system of points $p_i \in M$ such that the corresponding neighbourhoods U_{p_i} cover the entire manifold M. But then the set of all critical points of the map Φ will be the union of all sets of the form K_{p_i}, and, consequently, its image under the map Φ will be the union of all sets $\Phi(K_{p_i})$. Thus, Sard's theorem is completely proven, since the union of a countable number of subsets of the first category is also a subset of the first category.

Let us now consider the case $m < n$. It turns out that

Proposition 5.63. *if $m < n$ and if the manifold M is separable, then the image $\Phi(M)$ of the manifold M under the map Φ is a subset of the first category of the manifold N.*

To prove this assertion, we must essentially repeat verbatim the proof of Sard's theorem given above. Namely, the reduction to the special case of Euclidean spaces ($M = \mathbb{R}^m$ and $N = \mathbb{R}^n$) proceeds completely without any changes. Inequalities (5.60) and (5.61) remain valid for the maps $\varphi : \mathbb{R}^m \to \mathbb{R}^n$ $(m \neq n)$. The hyperplane L (now, of course, $n - 1$-dimensional) will now exist for any vector $\boldsymbol{u} \in I^m$ and therefore nowhere dense in $\mathbb{R}^n$ will be not only the set $\varphi(K)$, but the entire set $\varphi(I^m)$ (for the volume of which, instead of estimate (5.62), the estimate $c \cdot b(\varepsilon) \cdot \varepsilon^{n-1} \cdot h = cm^{m/2} \cdot b(\varepsilon) \cdot \varepsilon^{n-m-1}$).

Remark 5.64. The statements proved in this section are special cases of a certain general theorem of Sard, which applies to any dimensions m and n. We will not need this general theorem, and therefore we will neither prove nor formulate it here..

5.7 Embedding separable manifolds into the Euclidean space

To prove the embedding theorem formulated in the previous paragraph in the general case, we must more carefully study the structure of the space $\mathbb{R}^\infty$.

This space is naturally defined as an (infinite-dimensional) linear space (with respect to the coordinate-wise operations of addition and multiplication by a real number). We will assume that its finite-dimensional subspaces are equipped with the Euclidean topology and smoothness (note that we do not introduce either the topology or the smoothness into the space $\mathbb{R}^\infty$ itself). Among the finite-dimensional subspaces of $\mathbb{R}^\infty$, of particular importance for us will be the subspaces $\mathbb{R}^h$, $h \geq 1$, spanned by the first h vectors of the natural basis $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n, \ldots$, and the subspaces $\mathbb{R}^{h+1}_k$, $h > k$, generated by the subspace $\mathbb{R}^h$ and the vector $\boldsymbol{e}_k$, (Thus, $\mathbb{R}^{h+1}_{h+1} = \mathbb{R}^{h+1}$ for any $h \geq 1$.) Among the infinite-dimensional subspaces of $\mathbb{R}^\infty$, we especially note the subspaces $\mathbb{R}^\infty_{h,k}$, $k \geq h$, $h \geq 0$ spanned by all vectors of the natural basis, with the exception of the vectors $\boldsymbol{e}_{h+1}, \ldots, \boldsymbol{e}_k$. (Thus, in particular, $\mathbb{R}^\infty_{h,h} = \mathbb{R}^\infty$ and $\mathbb{R}^\infty = \mathbb{R}^h \times \mathbb{R}^\infty_{0,h}$.) It is clear that each of the subspaces $\mathbb{R}^\infty_{h,k}$ is naturally isomorphic to the entire space $\mathbb{R}^\infty$.

In what follows, we will consider some integer n fixed, no less than $2m + 1$, where m is the dimension of the manifold M under consideration. A unit vector $\boldsymbol{u}$ of the space $\mathbb{R}^{n+1}_k$, $k > n$, will be called a *projector* if its last coordinate u^k is non-zero. Each such vector defines a linear map $\pi_{\boldsymbol{u}}$ of the space $\mathbb{R}^\infty_{n,k-1}$ onto the space $\mathbb{R}^\infty_{n,k}$ (*projection along* the vector $\boldsymbol{u}$), which maps an arbitrary vector $\boldsymbol{t} \in \mathbb{R}^\infty_{n,k-1}$ to the vector

$$\pi_{\boldsymbol{u}} \boldsymbol{t} = \boldsymbol{t} - \frac{t^k}{u^k} \boldsymbol{u}.$$

Let Φ be an arbitrary smooth and locally finite-dimensional map of the manifold M into the space $\mathbb{R}^\infty$ and let p be an arbitrary point of the manifold M. To each vector $A \in M_p$ we associate a vector $d\Phi_p(A)$ of the space $\mathbb{R}^\infty$, setting

$$d\Phi_p(A) = (A\varphi^1, A\varphi^2, \ldots, A\varphi^i, \ldots),$$

where φ^i, $i = 1, 2, \ldots$, are functions on the manifold M that define the map Φ (at the point $q \in M$, the function φ^i, $i = 1, 2, \ldots$, takes a value equal to the i-th coordinate of the point $\Phi(q) \in \mathbb{R}^m$). It is clear that the map $d\Phi_p$ is linear, and therefore its image $d\Phi_p(M_p)$ is an (at most m-dimensional) linear subspace of $\mathbb{R}^\infty$. We will say that the vectors of this subspace are *tangent to* the set $\Phi(M)$ at the point $\Phi(p)$. Clearly, if the map Φ is regular, then the map $d\Phi_p$ is injective. It is easy to see that, conversely,

Proposition 5.65. *if for any point $p \in M$ the map $d\Phi_p : M_p \to \mathbb{R}^\infty$ is injective, then the map Φ is regular.*

Proof. Indeed, by hypothesis, the point p has a neighbourhood U in which only a finite number of functions φ^i are nonzero. Let these be the functions $\varphi^{i_1}, \ldots, \varphi^{i_s}$. Then, on the neighbourhood U, the map Φ can be viewed as a map of this neighbourhood into the finite-dimensional subspace $\overline{\mathbb{R}}^s$ spanned by the vectors $\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_s}$, and, accordingly, the map $d\Phi_p$ can be viewed as a map $M_p \to \overline{\mathbb{R}}^s$. Thus, the map Φ under consideration will be a smooth map of the smooth manifold U into the smooth manifold $\overline{\mathbb{R}}^s$, and the map $d\Phi_p : M_p \to \overline{\mathbb{R}}^s$ will (via the usual identifications) be its differential in the sense of §4.6. The injectivity of this latter map means that the map $\Phi : U \to \overline{\mathbb{R}}^s$ is regular at the point p. But then, as we know, from the functions $\varphi^{i_1}, \ldots, \varphi^{i_s}$ we can choose m functions that constitute a system of local coordinates at this point. To complete the proof, it remains to note that, by definition, this means that the map $\Phi : M \to \mathbb{R}^\infty$ is regular. □

Let us now assume that the image $\Phi(M)$ of the map Φ belongs to the subspace $\mathbb{R}^\infty_{n,k-1}$. Then, for any projection $\boldsymbol{u} \in \mathbb{R}^{n+1}_k$, the map $\Psi = \pi_{\boldsymbol{u}} \circ \Phi$ of the manifold M into the subspace $\mathbb{R}^\infty_{n,k}$ is defined. This map is defined by functions

$$\psi^i = \varphi^i - \varphi^k \cdot \frac{u^i}{u^k}, \quad i = 1, 2, \ \ldots,$$

and therefore is also smooth and locally finite-dimensional. Moreover, since for any vector $A \in M_p$ the following equalities hold:

$$A\psi^i = A\varphi^i - A\varphi^k \cdot \frac{u^i}{u^k}, \quad i = 1, 2, \ \ldots,$$

then

$$d\Psi_p = \pi_{\boldsymbol{u}} \circ d\Phi_p.$$

By the proposition proved above, it immediately follows that

Proposition 5.66. *if the map Φ is regular, and the projector $\boldsymbol{u}$ does not touch the set $\Phi(M)$ at any of its points, then the map $\Psi = \pi_{\boldsymbol{u}} \circ d\Phi$ is also regular.*

A non-zero vector $\boldsymbol{u} \in \mathbb{R}^\infty$ will be called a *secant* vector for the set $\Phi(M)$ if there exist two distinct points p and q of the manifold M such that the vector $\boldsymbol{u}$ is proportional to the vector $\Phi(q) - \Phi(p)$. Clearly,

Proposition 5.67. *if the map $\Phi : M \to \mathbb{R}^\infty_{n,k-1}$ is injective, then for any projection $\boldsymbol{u} \in \mathbb{R}^{n+1}_k$ that is not a secant vector for the set $\Phi(M)$, the map $\Psi = \pi_{\boldsymbol{u}} \circ d\Phi$ is also injective.*

Let Φ be an arbitrary smooth, locally finite-dimensional, regular, and injective map of a manifold M into $\mathbb{R}^\infty$, and let $\overline{\mathbb{R}}^{n+1}$ be an arbitrary $n+1$-dimensional linear subspace of $\mathbb{R}^\infty$. Let, furthermore, $\overline{\mathbb{S}}^n$ be the unit sphere of $\overline{\mathbb{R}}^{n+1}$. It is easy to see that

Proposition 5.68. *the set D of vectors $\boldsymbol{u} \in \overline{\mathbb{S}}^n$, either tangent to $\Phi(M)$ at some point or intersecting this set, is of first category in $\overline{\mathbb{S}}^n$.*

Proof. Indeed, let M_* be the manifold of vectors tangent to M. We define a map Φ_* of this manifold into $\mathbb{R}^\infty$ by setting

$$\Phi_*(A) = d\Phi_p(A),$$

if $A \in M_p$. Clearly, this map is smooth. Let π be an arbitrary projection of the space $\mathbb{R}^\infty$ onto the subspace $\overline{\mathbb{R}}^{n+1}$. Consider the map $\pi \circ \Phi_* : M_* \to \overline{\mathbb{R}}^{n+1}$. This map is smooth and therefore continuous. Therefore, the set W of all vectors $A \in M_*$ for which $(\pi \circ \Phi_*)(A) \neq 0$ is open in M_* and therefore is either empty or a $2m$-dimensional manifold. We define the map

$$\Psi := W \to \overline{\mathbb{S}}^n,$$

by setting

$$\Psi(A) = \frac{(\pi \circ \Phi_*(A)}{|(\pi \circ \Phi_*(A)|}$$

for any vector $A \in W$. Clearly, this map is smooth and its image $\Psi(W)$ contains all unit vectors of the space $\overline{\mathbb{R}}^{n+1}$ that are tangent to the set $\Phi(M)$ at at least one of its points. On the other hand, since, by hypothesis, $n \geq 2m + 1 > \dim W$, then, by the results of the previous section, the set $\Psi(W)$, and therefore the set D_1 of all unit vectors of the space $\overline{\mathbb{R}}^{n+1}$ that are tangent to the set $\Phi(M)$ at least at one of its points, has the first category in the sphere $\overline{\mathbb{S}}^n$.

Let us now consider the product $M \times M$ of the manifold M with itself and its map $\Phi^{'}$ into the space $\overline{\mathbb{S}}^n$, defined by the formula

$$\Phi^{'}(p, q) = \pi(\Phi(q) - \Phi(p)), \quad (p, q) \in M \times M,$$

where, as above, π is the projection of the space $\mathbb{R}^\infty$ onto the space $\overline{\mathbb{S}}^n$. This map is smooth and therefore continuous. Therefore, the set $W^{'}$ of all points $(p, q) \in M \times M$ for which $\Phi^{'} \neq 0$ is open in $M \times M$ and therefore is either empty or is a $2m$-dimensional smooth manifold. We define the map

$$\Psi^{'} := W^{'} \to \overline{\mathbb{S}}^n,$$

by setting

$$\Psi^{'}(p, q) = \frac{\Psi(p, q)}{|\Psi(p, q)|}$$

for any vector $(p, q) \in W^{'}$. It is clear that this map is smooth and its image $\Psi^{'}(W)$ contains the set D_2 of all vectors $\boldsymbol{u} \in \overline{\mathbb{S}}^n$ secants for the set $\Phi(M)$.

To complete the proof, it remains to note that the set D is the union of the sets D_1 and D_2 and, therefore, is also a set of the first category. □

Since the sphere $\mathbb{S}^n$ is a Hausdorff locally compact space, we, by comparing the proved propositions, immediately obtain that

Proposition 5.69. *for any smooth, locally finite-dimensional, regular, and injective map Φ of the manifold M into the space $\mathbb{R}^\infty_{n,k-1}$, the set of all projections of the space $\mathbb{R}^{n+1}_k$ for which the map $\pi_{\boldsymbol{u}} \circ \Phi$ is regular and injective is everywhere dense in the unit sphere of the space $\mathbb{R}^{n+1}_k$.*

We will call the map Φ of the manifold M into the space $\mathbb{R}^\infty$, defined by the functions $\varphi^1, \varphi^2, \ldots,$ *bounded with respect to the k-th coordinate* if the function φ^k is bounded (in absolute value).

Let Φ be an arbitrary map of M into the space $\mathbb{R}^\infty_{n,k-1}$ bounded by the k-th coordinate, $\boldsymbol{u} \in \mathbb{R}^{n+1}_k$ be a projection, and

$$\Psi = \pi_{\boldsymbol{u}} \circ \Phi : M \to \mathbb{R}^\infty_{n,k}$$

be the composition of Φ and the projection $\pi_{\boldsymbol{u}}$. Further, let $\varphi^1, \ldots, \varphi^n, \ldots, 0, \varphi^k, \ldots$ be functions defining the map Φ. Consider the maps

$$\pi \circ \Phi, \pi \circ \Psi : M \to \mathbb{R}^n,$$

where π is the natural projection of the space $\mathbb{R}^\infty$ onto the space $\mathbb{R}^n$. These maps are defined, respectively, by the functions $\varphi^1, \ \ldots, \varphi^n$ and $\varphi^1 - \varphi^k \frac{u^1}{u^k}, \ \ldots, \varphi^n - \varphi^k \frac{u^n}{u^k}$. Therefore, for any point $p \in M$, the relation

$$(\pi \circ \Psi)(p) - (\pi \circ \Phi))(p) = -\varphi^k(p) \cdot \frac{\boldsymbol{u}'}{u^k},$$

where $\boldsymbol{u}' = \pi(\boldsymbol{u})$ is a vector of the space $\mathbb{R}^n$ with coordinates $(u^1, \ \ldots, u^n)$. Denoting by K the maximum modulus of the function φ^k on the manifold N and setting $|u^k|| = \cos\alpha$, we immediately obtain from this that

$$|(\pi \circ \Psi)(p) - (\pi \circ \Phi))(p)| \leq K \tan\alpha$$

for any point $p \in M$.

We will call maps Φ_1 and Ψ_1 of a manifold M into the space $\mathbb{R}^n$ *ε-close*, where ε is some positive number, if $|\Psi_1(p) - \Phi_1(p)| < \varepsilon$ for any point $p \in M$. It follows directly from the estimate just proved that

Proposition 5.70. *for any map Φ of a manifold M into the space $\mathbb{R}^\infty_{n,k-1}$ bounded in the k-th coordinate and any $\varepsilon > 0$, in every everywhere dense subset of the unit sphere of $\mathbb{R}^{n+1}_k$ there exists a projector $\boldsymbol{u}$ such that the maps $\pi \circ \Phi$ and $\pi \circ \Psi$, where $\Psi = \pi_{\boldsymbol{u}} \circ \Phi$, are ε-close maps of the manifold M into the space $\mathbb{R}^n$.*

Comparing this assertion with the previous one, we obtain, in particular, that

Proposition 5.71. *for any smooth, locally finite-dimensional, regular, injective, and bounded in the k-th coordinate map Φ of the manifold M into the space $\mathbb{R}^\infty_{n,k-1}$ and any $\varepsilon > 0$, there exists a smooth, locally finite-dimensional, regular, and injective map Ψ of the manifold M into the space $\mathbb{R}^\infty_{n,k}$ of the form $\pi_{\boldsymbol{u}} \circ \Phi$, where $\boldsymbol{u} \in \mathbb{R}^{n+1}_k$ is some projection, such that the map $pi \circ \Phi$ and $\pi \circ \Psi$ are ε-close maps of the manifold M into the space $\mathbb{R}^n$.*

Let us now prove that

Proposition 5.72. *For any smooth, locally finite-dimensional, regular, injective, and bounded in all k-th ($k > n$) coordinates map Φ of the manifold M into the space $\mathbb{R}^\infty$, any $\varepsilon > 0$, and any integer $i \geq 0$, there exists a smooth, locally finite-dimensional, regular, injective, and bounded in all k-th ($k > n$) coordinates map Φ_i of the manifold M into the subspace $\mathbb{R}^\infty_{n,n+i}$ that*

1) *if a point $p \in M$ has the property that $\Phi(p) \in \mathbb{R}^{n+1}$, then $\Phi_j(p) \in \mathbb{R}^n$ and $\Phi_j(p) = \Phi_i(p)$ for any $j \geq i$;*

2) *the composition $\pi \circ \Phi_i$ of the map Φ_i and the natural projection $\pi : \mathbb{R}^\infty \to \mathbb{R}^n$ is a map $\left(1 - \frac{1}{2^i}\right) \varepsilon$-close to the map $\pi \circ \Phi$.*

We will construct the map Φ_i by induction on the number i, assuming that $\Phi_0 = \Phi$. Let the map Φ_i have already been constructed. We set $\Phi_{i+1} = \pi_{\boldsymbol{u}} \circ \Phi_i$, where $\boldsymbol{u} \in \mathbb{R}^{n+1}_{n+i+1}$ is some projection. According to the proposition just proved, the projection $\boldsymbol{u}$ can be chosen such that the map Φ_{i+1} is regular and injective and such that the map $\pi \circ \Phi_{i+1}$ is $\frac{\varepsilon}{2^{i+1}}$-close to the map $\pi \circ \Phi_i$. It is easy to see that the map Φ_{i+1} constructed in this way will have all the required properties. (Note that $\mathbb{R}^{n+i+1} \cap \mathbb{R}^{\infty}_{n,n+i+1} = \mathbb{R}^n$ and $\pi_{\boldsymbol{u}}(t) = \boldsymbol{t}$ for any vector $\boldsymbol{t} \in \mathbb{R}^{\infty}_{n,n+i+1}$.)

From this proposition it follows easily that

Proposition 5.73. *for any smooth map $\varphi_0 : M \to \mathbb{R}^n$ and any $\varepsilon > 0$, there exists a smooth, regular, injective, and ε-close to φ_0 map $\varphi : M \to \mathbb{R}^n$.*

Proof. Indeed, in the previous section we proved that for the manifold M there exists a smooth, locally finite-dimensional, regular, injective, and coordinately bounded map into the space $\mathbb{R}^\infty$. Since the space $\mathbb{R}^\infty$ is isomorphic to the space $\mathbb{R}^\infty_{0,n}$, we can assume that this kind of map exists not into the space $\mathbb{R}^\infty$but into the space $\mathbb{R}^\infty_{0,n}$. Multiplying this map by the given map φ_0 (and using the fact that $\mathbb{R}^n \times \mathbb{R}^\infty_{0,n} = \mathbb{R}^\infty$), we obviously obtain a smooth, locally finite-dimensional, regular, injective and bounded in all k-th ($k > n$) coordinates map Φ of the manifold M into the space $\mathbb{R}^\infty$, which has the property that $\pi \circ \Phi = \varphi_0$. Having constructed for this map the maps Φ_i envisaged by the previous proposition, we define a map φ of the manifold M into the space $\mathbb{R}^n$ by the formula

$$\varphi(p) = \Phi_i(p), \quad p \in M,$$

where i is a number such that $\Phi_i(p) \in \mathbb{R}^{n+i}$. According to property 1) of the maps Phi_i, this definition is correct, and according to 2) the map φ constructed in this way is ε-close to the map $\varphi_0 = \pi \circ \Phi$. We will show that it is smooth, regular, and injective.

Since the map Φ is locally finite-dimensional, for any point $p \in M$ there exists a neighbourhood U and an integer $i \geq 0$ such that $\Phi(U) \subset \mathbb{R}^{n+i}$. Consequently, on the neighbourhood U the map φ coincides with Φ_i and is therefore smooth and regular. Similarly, for any two points $p, q \in M$ there exists an integer $i \geq 0$ such that $\Phi(p), \Phi(q) \in \mathbb{R}^{n+i}$, and therefore $\varphi(p) = \Phi_i(p), \varphi(q) = \Phi_i(q)$. Consequently, if $\varphi(p) = \varphi(q)$, then $\Phi_i(p) = \Phi_i(q)$ and therefore $p = q$. Thus, the map φ is injective.

Thus, the above statement (known as the approximation theorem) is completely proven. □

The approximation theorem does not cover the embedding theorem, because the map φ it implies may not be a homeomorphism. Nevertheless, deriving the embedding theorem from it is no longer difficult. To do this, we first prove that

Proposition 5.74. *for any separable manifold M, there exists a smooth map of compact character to the real axis $\mathbb{R}$, and therefore to the space $\mathbb{R}^n$.*

For this purpose, we will again consider the functions f_i^k, $k = 0, 1, \ldots, m$; $i = 1, 2, \ldots$ constructed at the beginning of the previous section. We will now need not all of these functions, but only the functions

$$f_1^0, f_2^0, \ \dots, f_i^0, \ \dots$$

These functions are smooth on M, have compact supports, take values from the interval $[0, 1]$, and have the property that in the neighbourhood of each point $p \in M$, only a finite number of them are non-zero. Therefore, the formula

$$f(p) = f_1^0(p) + 2f_2^0(p) + \cdots + f_i^0(p) + \cdots, \quad p \in M,$$

defines a smooth non-negative function f on M, with the property that for any integer $i > 0$, the set $[f \le i] = f^{-1}([0, i])$ is compact. But it is clear that any such function realises a smooth map of the compact nature of the manifold M onto the real axis $\mathbb{R}$.

We further note that

Proposition 5.75. *any mapping $\varphi : M \to \mathbb{R}^n$ that is ε-close for some $\varepsilon > 0$ to a map $\varphi_0 : M \to \mathbb{R}^n$ of compact character is also a map of compact character.*

Proof. Indeed, for any set $A \subset \mathbb{R}^n$, the set $\varphi^{-1}(A)$ is obviously contained in the set $\varphi_0^{-1}(\overline{S}_\varepsilon(A))$, where $\overline{S}_\varepsilon(A)$ is the closure of the ε-neighbourhood $S_\varepsilon(A)$ of A. On the other hand, if the set A is compact, then the set $S_\varepsilon(A)$, as is easily seen, is also compact. □

Thus, applying the approximation theorem to an arbitrary smooth map $\varphi_0 : M \to \mathbb{R}^n$ of compact character, we obtain a smooth, regular, injective, and compact map $\varphi : M \to \mathbb{R}^n$. However, since the space $\mathbb{R}^n$ is a compactly generated (or "kaonic") Hausdorff space, every injective map of compact character from an arbitrary Hausdorff space to $\mathbb{R}^n$ is a homeomorphism onto a closed subset. Consequently, in particular, the map φ is also a homeomorphism. Since this map is also regular, it is a diffeomorphism of the manifold M onto some (closed) submanifold of the space $\mathbb{R}^n$.

Thus, the embedding theorem is completely proven.

Remark 5.76. We have proved even more than the embedding theorem asserted. First, we have proven that the dimension n of the Euclidean space into which a given m-dimensional manifold M is embedded can be any integer $n \ge 2m+1$, and in particular, the number $2m+1$. Second, we have proved that M can be realised in the space $\mathbb{R}^n$ (in particular, in the space $\mathbb{R}^{2m+1}$) as a closed submanifold.

It is also worth noting that the embedding theorem directly implies the theorem in §4.10 on the possibility of defining any separable manifold as a Riemannian space. Indeed, the space $\mathbb{R}^n$ is obviously a Riemannian space, and, on the other hand, each submanifold M of a Riemannian space N is obviously a Riemannian space with a metric tensor field $\iota^* g$, where $\iota : M \subset N$ is an inclusion map and g is the metric tensor field of the space N.

Chapter 6

Critical points of smooth functions

We finally turn to an exposition of Morse's results on the relationship between the topological structure of a smooth manifold M and the critical points of smooth functions defined on M. We also present a generalisation of these results (obtained by Bott), as well as some of their simplest applications.

§6.1 introduces the basic concepts of a critical point, a Hessian, a non-degenerate critical manifold, and a non-degenerate critical point. Here we also prove Morse's fundamental lemma on the representation of a smooth function in a neighbourhood of its non-degenerate critical point.

§6.2 proves, in two different ways, that on any separable smooth manifold there exists a smooth function all of whose critical points are non-degenerate.

§6.3 is devoted to Morse's fundamental theorem.

§6.4, based on Morse's theorem, proves that any separable smooth manifold is homotopy equivalent to some countable cellular decomposition.

The complex projective space $\mathbb{C}P^n$ is considered as an example.

In the final section, §6.5, Morse's theorem is generalised to the case where the critical points of a given function are only weakly nondegenerate.

The appendix derives classical Morse inequalities for the Betti numbers of the manifold under consideration. This appendix is the only place in the book where the reader is required to be familiar with homology theory.

However, all the necessary properties of the Betti numbers are explicitly formulated in this appendix.

6.1 Critical points and non-degenerate critical manifolds

In accordance with the general definition (see §4.6), a point p of a smooth manifold M is called a *critical point* of a smooth function $f \in \mathcal{O}(M)$ if $df_p = 0$, i.e., if in some (and therefore in any) system of local coordinates $x^1, \dots, x^m$ at the point p the following equalities hold

$$\left(\frac{\partial f}{\partial x^1}\right)_p = 0, \dots, \left(\frac{\partial f}{\partial x^m}\right)_p = 0.$$

A number $a \in \mathbb{R}$ is called a *critical value* of the function f, if there exists a critical point $p \in M$ such that $a = f(p)$.

Since for every smooth function f on the manifold M and any vector field $X \in \mathcal{O}^1(M)$ the expression $df(X)$ is a smooth function on M, we can talk about its differential $d[df(X)]$. For any two vector fields $X, Y \in \mathcal{O}^1(M)$, we set

$$Hf(X,Y) = d[df(X)](Y) - \frac{1}{2}df([X,Y]).$$

Obviously, the function Hf of the pair of fields $X, Y \in \mathcal{O}^1(M)$ defined in this way is linear in each argument, i.e., it is a bilinear form on the linear space $\mathcal{O}^1(M)$. (However, this function is not $\mathcal{O}^1(M)$-linear, i.e., it is not a tensor field.) In local coordinates $x^1, \dots, x^m$, the form Hf is defined, as is easily seen, by the equality

$$Hf(X,Y) = \frac{\partial^2 f}{\partial x^i \partial x^j} X^i Y^j + \frac{1}{2}\frac{\partial f}{\partial x^j}\left(\frac{\partial Y^j}{\partial x^i} X^i + \frac{\partial X^j}{\partial x^i} Y^i\right).$$

From this, in particular, it follows that the value $Hf(X,Y)(p)$ of the function $Hf(X,Y)$ at an arbitrary critical point $p \in M$ of the function f is expressed by the formula

$$Hf(X,Y)(p) = \left(\frac{\partial^2 f}{\partial x^i \partial x^j}\right)_p X^i_p Y^j_p$$

and therefore depends only on the vectors X_p and Y_p.

Therefore, the formula

$$H_p f(A,B) = Hf(X,Y)(p),$$

where $A, B \in M_p$, and X and Y are vector fields on the manifold M such that $X_p = A$ and $Y_p = B$, uniquely defines a certain (obviously symmetric) bilinear form $H_p f$ on the space M_p. The corresponding quadratic form

$$H_p f(A) = H_p f(A,A)$$

will be called the *Hessian* of the function f at the point p. In an arbitrary system of local coordinates $x^1, \dots, x^m$ on the manifold M at the point p, the Hessian $H_p f$ is expressed by the formula

$$H_p f(A) = \left(\frac{\partial^2 f}{\partial x^i \partial x^j}\right)_p A^i A^j$$

where $A^1, \dots, A^m$ are the components of the vector A in the basis $\left(\partial \partial x^i\right)_p$, $i = 1, \dots, n$.

Let us emphasise that the Hessian $H_p f$ is defined only for critical points p of the function f.

It is clear that the Hessian depends linearly on the function f, i.e.

Proposition 6.1. *For any functions $f, g \in \mathscr{O}(M)$ that have a critical point $p \in M$ and any real numbers $a, b \in \mathbb{R}$, the following equality holds:*

$$H_p(af + bg) = aH_p f + bH_p g.$$

Let $\Phi : N \to M$ be an arbitrary smooth map, and let the point $p \in M$ be a critical point of some function $\in \mathscr{O}(M)$. Clearly, any point $q \in N$ for which $\Phi(q) = p$ is a critical point of the function $\circ\Phi \in \mathscr{O}(N)$ (since $d(f \circ \Phi) = \Phi^*(df)$). Moreover, as is easy to see,

$$H_q(f \circ \Phi) = \Phi^* H_p f.$$

In particular, let N be an arbitrary submanifold of M and let a critical point p of f belong to N. Then this point, considered as a point of N, is a critical point of $|_N$, the restriction of f to N, and

Proposition 6.2. *the Hessian $H_p(f|_N)$ of the restriction $f|_N$ is the restriction $H_p f|_{N_p}$ of the Hessian $H_p f$ of f to the subspace N_p of M_p:*

$$H_p(f|_N) = H_p f|_{N_p}.$$

We will use these simple but important properties of the Hessian repeatedly in the future.

The *corank* of a quadratic form defined on a finite-dimensional linear space is the difference between the dimension of that space and the *rank* of the form in question, i.e., the rank of its matrix in an arbitrary basis. From the elements of linear algebra, it is known that the corank of a quadratic form can also be defined as the dimension of the form's nullspace, i.e., the maximal subspace on which the form is identically equal to zero.

We will call the *index* of a quadratic form its negative index of inertia, i.e., the number of negative terms in its representation as an algebraic sum of squares. This number can be characterized as the dimension of the maximal subspace on which the given form is negative definite.

We will call the corank and index of the Hessian $H_p f$ of a function f at a critical point p the corank and index of p, respectively, and will denote them by the symbols $c()$ (or $c_f(p)$) and $h(p)$ (or $h_f(p)$). Sometimes we will also use the rank $r(p) = r_f(p)$ of a critical point pp, i.e., the rank of the Hessian $H_p f$ at this point.

A system of local coordinates $x^1, \ldots, x^m$ at a critical point p of a function f will be called *associated with the function f* if

1) all coordinates $x^1, \ldots, x^m$ are equal to zero at the point p;

2) in the neighbourhood of the point p, the following equality holds

$$f = f(p) - (x^1)^2 - \cdots - (x^h)^2 + (x^{h+1})^2 + \cdots + (x^r)^2 + \sum_{i,j=r+1}^{m} h_{ij} x^i x^j, \quad (6.3)$$

where $h_{ij} = h_{ji}$ are some smooth functions equal to zero at this point:

$$h_{ij}(p) = 0, \quad i, j = r+1, \ \ldots, \ m.$$

Differentiating equality (6.3) and setting $x^1 = 0, \dots, x^m = 0$, we immediately obtain that

$$\left(\frac{\partial^2 f}{\partial x^i \partial x^j}\right)_p = \begin{cases} -2, & \text{if} \quad 1 \le i = j \le h, \\ +2, & \text{if} \quad h+1 \le i = j \le r, \\ 0, & \text{in all other cases.} \end{cases}$$

Therefore,

Proposition 6.4. *the numbers h and r involved in formula (6.3) coincide with the index $h(p)$ and the rank $r(p)$ of the critical point p.*

The foundation of the entire theory of critical points is the remarkable fact that

Proposition 6.5. *for any smooth function f on a manifold M, at each of its critical points p there exists a system of local coordinates associated with this function.*

Proof. Indeed, let $x^1, \dots, x^m$ be an arbitrary system of local coordinates at the point p, equal to zero at this point. According to formula (4.50), In some neighbourhood of the point p, the following equality holds

$$f = f(p) + g_i x^i,$$

where $g_1, \dots, g_m$ are smooth functions at the point p such that

$$g_i(p) = \frac{\partial f}{\partial x^i}_p = 0, \quad i = 1, \dots, m.$$

For the same reasons, there exist smooth functions $h^0_{ij} x^j$, $i, j = 1, \dots, m$ at p such that

$$g_i = h^0_{ij} x^j$$

for any $i = 1, \dots, m$, and

$$h^0_{ij}(p) = \frac{\partial g}{\partial x^j}_p$$

for all $i, j = 1, \dots, m$. We symmetrise the "coefficients" h^0_{ij}, setting

$$h_{ij} = \frac{1}{2}(h^0_{ij} + h^0_{ji}).$$

Since $h^0_{ij} x^i x^j = h_{ij} x x^i x^j$, then in a sufficiently small neighbourhood of the point p the following equality holds

$$f = f(p) + h_{ij} x^i x^j. \tag{6.6}$$

Let not all functions h_{ij} be zero at point p. After performing a linear coordinate transformation (known from the elementary theory of reducing quadratic forms to a sum of squares), we can assume that

$$h_{11} \ne 0.$$

Therefore, the function

$$g = \sqrt{|h_{11}|}$$

will be smooth at point p. Following the standard algorithm for reducing quadratic forms to sums of squares, we set

$$\begin{aligned} x^{'1} &= \left(x^1 + \sum_{i=2}^{m} \frac{h_{1i}}{h_{11}} x^i\right) g, \\ x^{'2} &= x^2, \\ \cdots &= \cdots \\ x^{'m} &= x^m. \end{aligned} \tag{6.7}$$

Since $\frac{D(x^{'1}, \ldots, x^{'m})}{D(x^1, \ldots, x^m)} = g(p) \neq 0$, then, by the theorem on the replacement of local coordinates, the functions $x^{'1}, \ldots, x^{'m}$ are local coordinates in some neighbourhood of the point p (equal to zero at the point p). On the other hand, it is easily verified that near the point p the equality holds

$$f = f(p) \pm (x^{'1})^2 + \sum_{i,j=2}^{m} h^{'}_{ij} x^i x^j,$$

where $h^{'}_{ij} = h_{ji}, i, j = 2, \ldots, m$, are some smooth functions.

Continuing this process of "splitting off squares," we will eventually obtain at the point p a system of local coordinates associated with the given function f. □

A closed submanifold N of a manifold M is called a *non-degenerate critical manifold* of the function f if

1) the submanifold M is connected;

2) all its points are critical points of the function f;

3) in some neighbourhood of it there are no other critical points of the function f;

4) for any point $p \in N$, the null space of the quadratic form $H_p f$ coincides with the subspace N_p tangent to the submanifold N.

According to this definition, all points $p \in N$ have the same corank $c(p)$, equal to the dimension n of the manifold N. On the other hand, by obvious continuity considerations, the index of a point $p \in N$, as this point moves continuously along the manifold N, can change only at the points where the corank jumps. Therefore, all points $p \in N$ also have the same index $h(p)$. We will call it the *index* of the non-degenerate critical manifold N and denote it by the symbol $h(N)$.

We will call critical points belonging to non-degenerate critical manifolds *weakly non-degenerate* critical points of the function f.

A non-degenerate critical manifold N containing a weakly non-degenerate critical point p is defined in some neighbourhood of this point by the equations

$$\frac{\partial f}{\partial x^1} = 0, \ \ldots, \ \frac{\partial f}{\partial x^m} = 0,$$

where $x^1, \ \ldots, \ x^m$ is an arbitrary system of local coordinates at the point p. In the system of local coordinates $x^1, \ \ldots, \ x^m$ associated with the function /, these equations have the form

$$-2x^1 + \sum_{i,j=r+1}^{m} \frac{\partial h_{ij}}{\partial x^1} x^i x^j = 0, \ \cdots, \ -2x^h + \sum_{i,j=r+1}^{m} \frac{\partial h_{ij}}{\partial x^h} x^i x^j = 0,$$
$$2x^{h+1} + \sum_{i,j=r+1}^{m} \frac{\partial h_{ij}}{\partial x^{h+1}} x^i x^j = 0, \ \cdots, \ 2x^r + \sum_{i,j=r+1}^{m} \frac{\partial h_{ij}}{\partial x^r} x^i x^j = 0,$$
$$2\sum_{j=r+1}^{m} h_{r+1,j} x^j + \sum_{i,j=r+1}^{m} \frac{\partial h_{ij}}{\partial x^{r+1}} x^i x^j = 0, \ \cdots, \ 2\sum_{j=r+1}^{m} h_{m,j} x^j + \sum_{i,j=r+1}^{m} \frac{\partial h_{ij}}{\partial x^m} x^i x^j = 0.$$

Let's put it this way

$$\begin{gathered} y^1 = -x^1 + \frac{1}{2}\sum_{i,j=r+1}^{m} \frac{\partial h_{ij}}{\partial x^1} x^i x^j, \ \cdots, \ y^h = -x^h + \frac{1}{2}\sum_{i,j=r+1}^{m} \frac{\partial h_{ij}}{\partial x^h} x^i x^j, \\ y^{h+1} = x^{h+1} + \frac{1}{2}\sum_{i,j=r+1}^{m} \frac{\partial h_{ij}}{\partial x^{h+1}} x^i x^j, \ \cdots, \ y^r = x^r + \frac{1}{2}\sum_{i,j=r+1}^{m} \frac{\partial h_{ij}}{\partial x^r} x^i x^j, \\ y^{r+1} = x^{r+1}, \ \cdots, \ y^m = x^m. \end{gathered} \tag{6.8}$$

The theorem on the change of local coordinates immediately implies that the functions $y^1, \ \ldots, \ y^m$ on the manifold M are local coordinates at the point p, equal to zero at this point. Therefore, the equations

$$y^1 = 0, \ \ldots, \ y^r = 0 \tag{6.9}$$

define, in the corresponding coordinate neighbourhood U, a certain $m-r$-dimensional submanifold N' of the manifold M containing the point p. Clearly, this submanifold has the property that

$$N \cap U \subset N'.$$

Since $m - r = n$, it follows that near the point p he submanifolds N and N' coincide, i.e., near the point p the submanifold NN is defined by equations (6.9).

Furthermore, as is easy to calculate, the system of local coordinates $y^1, \ \ldots, \ y^r$ is still associated with the function f. A system of local coordinates at a point of a non-degenerate critical manifold that has these two properties will be called *rigidly associated with the function* f. Thus, we have shown that

Proposition 6.10. *at any point p of an arbitrary non-degenerate critical manifold N, there exists a system of local coordinates rigidly associated with the given function f.*

In any associated system of local coordinates $x^1, \ldots, x^m$, the function f in some neighbourhood U of the weakly non-degenerate critical point p^m under consideration is given by the formula

$$f = f(p) - (x^1)^2 - \cdots - (x^h)^2 + (x^{h+1})^2 + \cdots + (x^r)^2 + \sum_{i,j=r+1}^{m} h_{ij} x^i x^j,$$

where $h_{ij} = h_{ji}$, $i, j = r+1, \ldots, m$, are some smooth functions at the point p that are equal to zero at this point. If the coordinates $x^1, \ldots, x^m$ are rigidly associated with the function f, then the form of these functions can be specialised. Specifically, it turns out that

Proposition 6.11. *without loss of generality, we can assume that the functions h_{ij} have the form*

$$h_{ij} = \sum_{k,\ell=1}^{r} h_{ijk\ell} x^k x^\ell, \quad i, j = r+1, \ldots, m,$$

where $h_{ijk\ell}$ are certain smooth functions.

Proof. Indeed, using formula (4.50) twice, we can write each of the functions h_{ij} in the following form:

$$h_{ij} = h^0_{ij} + \sum_{k=1}^{r} h^0_{ijk} x^k + \sum_{k,\ell=1}^{r} h_{ijk\ell} x^k x^\ell, \quad i, j = r+1, \ldots, m,$$

where the functions h^0_{ij} and h^0_{ijk} depend only on the coordinates $x^{r+1}, \ldots, x^m$. Without loss of generality, we can assume that the functions h^0_{ij}, h^0_{ijk}, and $h_{ijk\ell}$ are symmetric with respect to the indices i and j, and the functions $h_{ijk\ell}$ are symmetric with respect to the indices k and ℓ. Thus

$$\frac{\partial f}{\partial x^s} = \begin{cases} \pm 2x^s + \sum_{i,j=r+1}^m \left(h^0_{ijs} + 2\sum_{k=1}^r h_{ijks} x^k + \sum_{k,\ell=1}^r \frac{\partial h_{ijk\ell}}{\partial x^s} x^k x^\ell \right) x^i x^j & \text{if} \quad 1 \le s \le r; \\ \sum_{i,j=r+1}^m \left(\frac{\partial h^0_{ij}}{\partial x^s} + \sum_{k=1}^r \frac{\partial h^0_{ijk}}{\partial x^s} x^k + \sum_{k,\ell=1}^r \frac{\partial h_{ijk\ell}}{\partial x^s} x^k x^\ell \right) x^i x^j & \\ +2\sum_{i=r+1}^m \left(h^0_{is} + \sum_{k=1}^r h^0_{ijk} x^k + \sum_{k,\ell=1}^r h_{ijk\ell} x^k x^\ell \right) x^i, & \text{if} \quad r+1 \le s \le m. \end{cases}$$

Setting here $x^1 = 0, \ldots, x^r = 0$, we immediately obtain that on the set $N \cap U$ the following equalities hold:

$$\sum_{i,j=r+1}^{m} h_{ijs} x^i x^j = 0, \quad s = 1, \ldots, r, \tag{6.12}$$

$$\sum_{i,j=r+1}^{m} \frac{\partial h_{ij}^0}{\partial x^s} x^i x^j + 2 \sum_{i,j=r+1}^{m} h_{is}^0 x^i = 0, \quad s = r+1, \dots, m. \tag{6.13}$$

But, by hypothesis, the functions h_{ij}^0 and h_{ijs}^0 depend only on the coordinates $x^{r+1}, \dots, x^m$. Therefore, the left-hand sides of equality (6.12) and (6.13) depend only on these coordinates and, therefore, are identically equal to zero, i.e., on the entire neighbourhood U.

Since

$$\sum_{i,j=r+1}^{m} \frac{\partial h_{ij}^0}{\partial x^s} x^i x^j + 2 \sum_{i,j=r+1}^{m} h_{is}^0 x^i = \frac{\partial}{\partial x^s} \left(\sum_{i,j=r+1}^{m} h_{ij}^0 x^i x^j \right)$$

and since the functions h_{ij}^0 do not depend on the coordinates $x^1, \dots, x^r$ and are equal to zero at the point p, it follows from relations (6.13) that in the neighbourhood U the following equalities hold:

$$\sum_{i,j=r+1}^{m} h_{ij}^0 x^i x^j = 0.$$

In view of relations (6.12) we obtain from here that

$$\sum_{i,j=r+1}^{m} h_{ij} x^i x^j = \sum_{k,\ell=1}^{r} \sum_{i,j=r+1}^{m} h_{ijk\ell} x^k x^\ell x^i x^j.$$

Thus, the above statement is completely proven. □

This proposition immediately implies, in particular, that at all points of the manifold N belonging to a neighbourhood U, the function f takes the same value $f(p)$. Since the manifold N is connected, it follows that $f(p) = f(q)$ for any points $p, q \in N$. In other words,

Proposition 6.14. *on every non-degenerate critical manifold N, the function f is constant.*

A critical point p of a function f is called *non-degenerate* if its corank $c(p)$ is zero. In the neighbourhood of any point $p \in M$, one can define a smooth function f for which the point p is a non-degenerate critical point of any given index h. Such a function can be defined, for example, by the formula

$$f = -(x^1)^2 - \dots - (x^h)^2 + (x^{h+1})^2 + \dots + (x^m)^2 \tag{6.15}$$

where $x^1, \dots, x^m$ is an arbitrary system of local coordinates at the point p that vanish at this point. Up to an arbitrary constant, this is the general form of smooth functions that have p as their non-degenerate critical point. Indeed, any such function f in normal (with respect to this function) coordinates $x^1, \dots, x^m$ has the form (6.15) (up to a constant term $f(p)$). Thus,

Proposition 6.16. *a point $p \in M$ is a non-degenerate critical point of index h of a smooth function f if and only if at p there exists a system of local coordinates $x^1, \dots, x^m$ (equal to zero at this point) such that in some neighbourhood of this point the following equality holds*

$$f = f(p) - (x^1)^2 - \cdots - (x^h)^2 + (x^{h+1})^2 + \cdots + (x^m)^2.$$

From this statement (known as Morse's lemma) it follows directly, in particular, that

Proposition 6.17. *every non-degenerate critical point of the function f is isolated in the set of all critical points of this function, i.e., it has a neighbourhood free of other critical points.*

Thus, each non-degenerate critical point p, considered as a zero-dimensional submanifold of M, satisfies the isolation condition (condition 3) of the definition of a non-degenerate critical manifold. The remaining conditions of this definition are obviously satisfied. Thus,

Proposition 6.18. *any non-degenerate critical point is a weakly non-degenerate point.*

Consider again an arbitrary non-degenerate critical manifold N of the function f. Let p be an arbitrary point of it, and let $N^\#$ be an arbitrary r-dimensional ($r = m - n$) submanifold of M complementary to the submanifold N at the point p (i.e., having only one common point p near the point p with the submanifold N and possessing the property that its tangent space $N^\#$ intersects the space N_p only at zero).

By hypothesis, the subspace N_p is the null space of the quadratic form $H_p f$. Therefore, on the complementary space $N_p^\#$, this quadratic form is non-degenerate. But we know that the restriction $H_p f|_{N_p^\#}$ of the Hessian $H_p f$ to the subspace $N_p^\#$ is nothing other than the Hessian of the restriction $f|_{N^\#}$ of the function f to the submanifold $N^\#$. Thus,

Proposition 6.19. *the function $f|_{N^\#}$ on the submanifold $N^\#$ has the point p as its non-degenerate critical point.*

Remark 6.20. It can be shown that, conversely, any connected submanifold N of M consisting of critical points of f is a non-degenerate critical manifold if, for any point p of N and any submanifold $N^\#$ complementary to the submanifold N at p, the latter point is a non-degenerate critical point of $f|_{N^\#}$. We will not need this fact and therefore we will leave it without proof.

6.2 Functions all of whose critical points are non-degenerate

In this section, we prove that

Proposition 6.21. *on any smooth separable manifold M, there exists a non-negative smooth function f of compact character (i.e., representing a map of the compact character of the manifold M onto the real axis $\mathbb{R}$), all of whose critical points are non-degenerate.*

In proving this statement, by the embedding theorem proved in the previous chapter, we can, without loss of generality, assume that the manifold M under consideration is realised as a closed submanifold of some Euclidean space $\mathbb{R}^n$.

Let $\boldsymbol{u}$ be an arbitrary vector $\mathbb{R}^n$. Consider on the manifold M the function f defined by the formula

$$f(p) = \frac{1}{2}|\boldsymbol{u} - p|^2, \quad p \in M. \tag{6.22}$$

Clearly, this function is smooth, non-negative, and has the property that for any $a \in \mathbb{R}^n$, the set $[f \leq a]$ (being a closed and bounded subset of the space $\mathbb{R}^n$) is compact. On the other hand, it is obvious that a non-negative function on a manifold M has the latter property if and only if it maps the compact character of the manifold M onto the real axis $\mathbb{R}$. Thus, to prove the above statement, it suffices to show that for at least one vector $\boldsymbol{u} \in \mathbb{R}^n$, function (6.22) has only non-degenerate critical points.

Let p be an arbitrary point of M. Formula (6.22) immediately implies that for any system $x^1, \ldots, x^m$ of local coordinates on the manifold M and the point p, the following equalities hold:

$$\frac{\partial f}{\partial x^i} = -\left(\boldsymbol{u} - p, \left(\frac{\partial}{\partial x^i}\right)_p\right), \quad i = 1, \ldots, m,$$

where the vectors $\left(\frac{\partial}{\partial x^i}\right)_p \in M_p$, $i = 1, \ldots, m$, are considered as vectors of the space $\mathbb{R}^n$. Therefore,

Proposition 6.23. *a point $p \in M$ is a critical point of function (6.22) if and only if the vector $\boldsymbol{u} - p$ is orthogonal to the manifold M, i.e., when the pair $(p, \boldsymbol{u} - p)$ belongs to the manifold N_* constructed in §5.5.*

Further, if the local coordinates $x^1, \ldots, x^m$ are defined by formulae (5.45), then

$$\frac{\partial}{\partial x^i} = \boldsymbol{e}_i + \sum_{k=m+1}^{n} \frac{\partial t^k}{\partial x^i} \boldsymbol{e}_k$$

(To simplify the formulae, we omit the reference to the point p here) and therefore

$$\frac{\partial f}{\partial x^i} = -\left[u^i - x^i + \sum_{k=m+1}^{n} \frac{\partial t^k}{\partial x^i} \cdot (u^k - t^k)\right]$$

(We use the notation and conventions introduced in §5.5; in particular, we assume that the basis $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ of the space $\mathbb{R}^n$ is orthonormal). Therefore,

$$\frac{\partial^2 f}{\partial x^i \partial x^j} = \delta^i_j - \sum_{k=m+1}^{n} \frac{\partial^2 t^k}{\partial x^i \partial x^j} \cdot (u^k - t^k) + \sum_{k=m+1}^{n} \frac{\partial t^k}{\partial x^i} \cdot \frac{\partial t^k}{\partial x^j}.$$

Using the matrices A and T introduced in §5.5, we can, as is easily seen, rewrite these relations as a single matrix equality

$$\frac{\partial^2 f}{\partial x^i \partial x^j} = A + TT'$$

(meaning that the matrices A and T are calculated at the point $(p, \boldsymbol{u} - p) \in N_*$). Comparing this equality with formula (5.50), we immediately obtain that

Proposition 6.24. *the determinant*

$$\det \left| \frac{\partial^2 f}{\partial x^i \partial x^j} \right|$$

of the Hessian of function (6.22) at the critical point $p \in M$ coincides with the Jacobian

$$\det \left| \frac{\partial v^i}{\partial x_*^j} \right|$$

of the map $\rho : N_* \to \mathbb{R}^n$ at the point $(p, \boldsymbol{u} - p) \in N_*$.

Therefore

Proposition 6.25. *A critical point $p \in M$ of function (6.22) is degenerate if and only if the point $(p, \boldsymbol{u} - p) \in N_*$ is a critical point of the map ρ.*

Since $\rho(p, \boldsymbol{u} - p) = \boldsymbol{u}$, it follows that

Proposition 6.26. *if the vector $\boldsymbol{u} \in \mathbb{R}^n$ is not the image under the map ρ of any critical point of this map, then all critical points of the corresponding vector and function (6.22) are non-degenerate.*

Since the set of vectors $\boldsymbol{u} \in \mathbb{R}^n$ possessing the latter property, being - by Sard's theorem - the complement of some set of the first category, is not empty, the existence of functions (6.22) having only critical points is thereby completely proven.

Although functions (6.22) are analytically quite simple, they do not have a simple, visually understandable geometric interpretation. We will now present another method for constructing functions with only non-degenerate critical points that avoids this drawback. However, the functions obtained by this method are, generally speaking, unbounded below and are not compact maps onto the real axis. Nevertheless, for the most interesting case of compact manifolds, this circumstance is unimportant, since any function on a compact manifold is automatically a compact map and bounded below (and therefore can be converted into a non-negative function by adding some constant).

Let L^{n-1} be an arbitrary $n-1$-dimensional linear subspace (hyperplane) of $\mathbb{R}^n$. (We still assume that $M \subset \mathbb{R}^n$). Consider on the manifold M a numerical function f whose value $f(p)$ at an arbitrary point $p \in M$ is (taken with a certain sign) the distance of the point p from the hyperplane L^{n-1}. Considering the hyperplane L^{n-1} as horizontal, we can visualize the number $f(p)$ as the "height" of the point p above the hyperplane L^{n-1}.

Let L^1 be a one-dimensional linear space (line) orthogonal to the hyperplane L^{n-1}, and let π be the projection of the space $\mathbb{R}^n$ onto the line L^1. Then we can define the function f by the formula

$$\pi(p) = f(p)\boldsymbol{u}, \quad p \in M,$$

where $\boldsymbol{u}$ is a unit vector located on the line L^1, i.e., a unit vector orthogonal to the hyperplane L^{n-1}. It is clear that

$$f = \xi \circ \pi \circ \iota, \tag{6.27}$$

where ι is the inclusion map $M \subset \mathbb{R}$, and $\xi : L^1 \to \mathbb{R}$ is the coordinate diffeomorphism corresponding to the vector $\boldsymbol{u}$ (which is a basis for the space L^1). Therefore,

$$d_p = d\xi_{\pi_p} \circ d\pi_p \circ d\iota_p$$

for any point $p \in M$. Since the map $d\xi_{\pi_p}$ is isomorphic, this formula immediately implies that $df_p = 0$ if and only if $d\pi_p \circ d\iota_p = 0$, i.e., when $d\pi_p(M_p) = 0$. On the other hand, it is clear that, by standard identifications, the linear map $\pi : \mathbb{R}^n \to L^1$ coincides with its differential $d\pi_p : (\mathbb{R}^n)_p \to (L^1)_{\pi(p)}$. Therefore, $d\pi_p(M_p) = 0$ if and only if $\pi_p(M_p) = 0$, i.e., when the vector $\boldsymbol{u}$ is orthogonal to the subspace $M_p \subset \mathbb{R}^n$.

Let φ be a smooth function on the manifold N_* that associates with each point $(p, \boldsymbol{u}) \in N_*$ the length $|\boldsymbol{u}|$ of the vector $\boldsymbol{u}$. Consider the level surface $N = [\varphi = 1]$ of φ, i.e., the set of all pairs $(p, \boldsymbol{u}) \in N_*$ for which $|\boldsymbol{u}| = 1$. Obviously, this level surface is regular and therefore represents an $n-1$-dimensional submanifold of the manifold N_*. We will call this *submanifold the manifold of unit vectors orthogonal to the manifold N*.

Let $(p_0, \boldsymbol{u}_0)$ be an arbitrary point of the manifold N. According to what was said in §5.5, in order to construct local coordinates on the manifold N_* at the point $(p_0, \boldsymbol{u}_0)$, we should choose in the space $\mathbb{R}^n$ an orthonormal basis $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ with the property that its first m vectors $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$ form a basis of the space M_{p_0}. Moreover, since the vector $\boldsymbol{u}_0$ is orthogonal, by hypothesis, to the space M_{p_0}, we can, without loss of generality, assume that $\boldsymbol{e}_n = \boldsymbol{u}_0$.

Let $x_*^1, \ldots, x_*^n$ be the local coordinates at the point $(p, \boldsymbol{u})$ of the manifold N_* corresponding to the basis $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{n-1}, \boldsymbol{e}_n = \boldsymbol{u}_0$ (see formula (5.48)). Consider the subset of the neighbourhood U_* consisting of all its points $(p, \boldsymbol{u})$ for which $x_*^n(p, \boldsymbol{u}) > 0$, i.e., for which $u^n > 0$. Clearly, this subset is open and contains the point $(p_0, \boldsymbol{u}_0)$. Let V be its intersection with the submanifold N and let

$$y^1 = x_*^1|_N, \ \ldots, \ y^m = x_*^m|_N, \ y^{m+1} = \frac{x_*^{m+1}}{x_*^n}|_N, \ \ldots, \ y^{n-1} = \frac{x_*^{n-1}}{x_*^n}|_N. \tag{6.28}$$

It is easy to verify that functions (6.28) are local coordinates of the manifold N at the point $(p_0, \boldsymbol{u}_0)$ with coordinate neighbourhood V.

In explicit form, coordinates (6.28) are defined by the formulae

$$y^1(p,\boldsymbol{u}) = x^1(p), \ \ldots, \ y^m(p,\boldsymbol{u}) = x^m(p),$$
$$y^{m+1}(p,\boldsymbol{u}) = \frac{u^{m+1}}{u^n}, \ \ldots, \ y^{n-1}(p,\boldsymbol{u}) = \frac{u^{n-1}}{u^n}, \quad (p,\boldsymbol{u}) \in V,$$

where, in particular, $u^{m+1}, \ \ldots, \ u^n$ are the last $n-m$ coordinates of the vector $\boldsymbol{u}$ in the basis $\boldsymbol{e}_1, \ \ldots, \ \boldsymbol{e}_m$. From this and from relations (5.46) it immediately follows that for each point $(p,\boldsymbol{u}) \in V$ the coordinates $u^1, \ \ldots, \ u^n$ of the vector $\boldsymbol{u}$ are connected with the coordinates $y^1, \ \ldots, \ y^{n-1}$ of the point $(p,\boldsymbol{u})$ by the formulae:

$$\begin{aligned}
&\overline{u}^1 = -\sum_{j=m+1}^{n-1}\left(\frac{\partial t^j}{\partial x^1}\right)_p y^j - \left(\frac{\partial t^n}{\partial x^1}\right)_p, \ \ldots, \ \overline{u}^m = -\sum_{j=m+1}^{n-1}\left(\frac{\partial t^j}{\partial x^m}\right)_p y^j - \left(\frac{\partial t^n}{\partial x^m}\right)_p, \\
&\overline{u}^m = y^{m+1}, \ \ldots, \ \overline{u}^{n-1} = y^{n-1},
\end{aligned} \tag{6.29}$$

where $\overline{u}^1 = \frac{u^1}{u^n}, \ \ldots, \ \overline{u}^{n-1} = \frac{u^{n-1}}{u^n}$. In these formulas, the functions $\frac{\partial t^j}{\partial x^i}$, $i = 1, \ \ldots, \ m$; $j = m+1, \ \ldots, \ n$, should be considered as functions on the manifold N and, accordingly, considered as functions of the local coordinates $y^1, \ \ldots, \ y^{n-1}$ (in fact, only the coordinates $y^1, \ \ldots, \ y^m$).

Let us now consider the map

$$\nu : N \longrightarrow \mathbb{S}^{n-1}$$

of the manifold N into the unit sphere $\mathbb{S}^{n-1}$ of the space $\mathbb{R}^n$, defined by the formula

$$\nu(p,\boldsymbol{u}) = \boldsymbol{u}.$$

Since the quantities $\overline{u}^1, \ \ldots, \ \overline{u}^{n-1}$ can obviously be taken as local coordinates on the sphere $\mathbb{S}^{n-1}$ at the point $\boldsymbol{u}_0 = \boldsymbol{e}_m$, then formulae (6.29) represent nothing more than expressions of the map ν in the local coordinates $y^1, \ \ldots, \ y^{n-1}$ and $\overline{u}^1, \ \ldots, \ \overline{u}^{n-1}$. Differentiating these formulae and taking into account the identities $y^1(p,\boldsymbol{u}) = x^1(p), \ \ldots, \ y^m(p,\boldsymbol{u}) = x^m(p)$ and the fact that the functions $\frac{\partial t^j}{\partial x^i}$, $i = 1, \ \ldots, m$; $j = m+1, \ \ldots, \ n$, do not depend on the coordinates $y^1, \ \ldots, \ y^{n-1}$, we immediately obtain that

$$\frac{\partial \overline{u}^i}{\partial y^k} = \begin{cases} -\sum_{j=m+1}^{n-1}\left(\frac{\partial^2 t^j}{\partial x^i \partial x^k}\right)_p y^i - -\left(\frac{\partial^2 t^n}{\partial x^i \partial x^k}\right)_p, & \text{if} \quad 1 \le i,k \le m, \\ -\left(\frac{\partial t^k}{\partial x^i}\right)_p, & \text{if} \quad 1 \le i \le m, \ m+1 \le k \le n-1, \\ \delta^i_k, & \text{in other cases.} \end{cases}$$

Since at the point $(p_0,\boldsymbol{u}_0)$ the coordinates $y^{m+1}, \ \ldots, \ y^{n-1}$ are equal to zero, then, in particular,

$$\left(\frac{\partial \overline{u}^i}{\partial y^k}\right)_{(p_0,\boldsymbol{u}_0)} = \begin{cases} -\left(\frac{\partial^2 t^n}{\partial x^i \partial x^k}\right)_p, & \text{if} \quad 1 \le i,k \le m, \\ -\left(\frac{\partial t^k}{\partial x^i}\right)_p, & \text{if} \quad 1 \le i \le m, \ m+1 \le k \le n-1, \\ \delta^i_k, & \text{in other cases.} \end{cases}$$

It follows that the point $(p_0, \boldsymbol{u}_0)$ is a critical point of the map ν if and only if

$$\det\left|\left(\frac{\partial^2 t^n}{\partial x^i \partial x^k}\right)_p\right| = 0. \tag{6.30}$$

On the other hand, it is clear that the function t^n is nothing other than function (6.27) corresponding to a hyperplane orthogonal to the vector $\boldsymbol{u}_0$, so equality (6.30) represents a condition for the degeneracy of the critical point p_0 of this function. We have thus proved (we drop the index 0 in the notation of the point p_0 and the vector $\boldsymbol{u}_0$) that

Proposition 6.31. *a critical point $p \in M$ of function (6.27) is non-degenerate if and only if the pair $(p, \boldsymbol{u})$ (which is a point of the manifold N) is not a critical point of the map $\nu : N \to \mathbb{S}^{n-1}$.*

Therefore,

Proposition 6.32. *if the vector $\boldsymbol{u} \in \mathbb{S}^{n-1}$ is not the image under the map ν of any critical point of this map, then all critical points of function (6.27) are non-degenerate.*

Thus, according to Sard's theorem, we can always choose the hyperplane L^{n-1} (or, equivalently, the vector $\boldsymbol{u} \in \mathbb{S}^{n-1}$) so that function (6.27) has only non-degenerate critical points.

6.3 Morse's theorem

Let M be an arbitrary separable smooth manifold and f an arbitrary smooth function on M. In this section, we describe the evolution of the sets $[f \le a]$ as the parameter a increases. Although this evolution can, generally speaking, be traced up to a diffeomorphism, for simplicity we restrict ourselves to the more elementary problem of studying the evolution of the sets $[f \le a]$ up to homotopy equivalence.

First, we consider the case where the parameter a does not pass through any critical value of the function f. It turns out that in this case, the set $[f \le a]$ is essentially unchanged. Specifically, as we will now show,

Proposition 6.33. *if the numbers a and $b > a$ have the property that the set $[a \le f \le b]$ is compact and does not contain any critical point of the function f, then there exists a diffeomorphism Φ of the manifold M onto itself that maps the set $[f \le a]$ onto the set $[f \le b]$.*

Moreover,

Proposition 6.34. *in this case, the set $[f \le a]$ is a deformation retract of the set $[f \le b]$.*

Since the manifold M is, by assumption, separable, we can define it as a Riemannian space (see §4.10) and therefore introduce the gradient field $\operatorname{grad} f$ of the function f. This field, and hence the function $(\operatorname{grad} f, \operatorname{grad} f)$, is zero at the critical points of f and only at these points. In particular, on the set $[a \le f \le b]$, which, by assumption, does not contain critical points of f, the function $(\operatorname{grad} f, \operatorname{grad} f)$ is positive. Therefore, the function $1/(\operatorname{grad} f, \operatorname{grad} f)$ is smooth on the set $[a \le f \le b]$. Since the set $[a \le f \le b]$ is, by assumption, compact, there exists a positive smooth function h with compact support on the manifold M, which is an extension from the set $[a \le f \le b]$ of the function $1/(\operatorname{grad} f, \operatorname{grad} f)$ (see §4.4). Let

$$X = h \cdot \operatorname{grad} f.$$

It is clear that the support of this vector field, being, obviously, a subset of the support of the function h, is compact. Therefore, on the manifold M there exists a one-parameter group of diffeomorphisms

$$\varphi : \mathbb{R} \times M \to M$$

generating the field X (see §5.3).

Let p be an arbitrary point of M and let

$$g(t) = f(\varphi_p(t)), \quad t \in \mathbb{R},$$

where, as in §5.3,

$$\varphi_p(t) = \varphi(t, p)$$

be the trajectory of the group φ passing through the point p. Since this trajectory is an integral curve of the field X, then

$$\dot{g}(t) = X_{\varphi_p(t)} f = (X, \operatorname{grad} f)(\varphi_p(t)) = (h \cdot (\operatorname{grad} f, \operatorname{grad} f))(\varphi_p(t)).$$

Since $h = 1/(\operatorname{grad} f, \operatorname{grad} f)$ on $[a \le f \le b]$, it follows that $\dot{g}(t) = 1$ if $\varphi_p(t) \in [a \le f \le b]$, i.e., $[a \le g(t) \le b]$. Furthermore, the function g is obviously monotone. Therefore, denoting by t_0 the value of the parameter t for which $g(t_0) = a$ (if it exists), and setting $d = b - a$, we obtain that

$$g(t) \begin{cases} < a, & \text{if} \quad t < t_0, \\ = t - t + a, & \text{if} \quad t_0 \le t \le t_0 + d, \\ > b, & \text{if} \quad t_0 + d < t. \end{cases}$$

Now let $p \in [f \le a]$, i.e., let $g(0) \le a$. Then

either $g(t) < a$ for any $t \in \mathbb{R}$ and, in particular, for $t = d$,

or there exists $t \in \mathbb{R}$ such that $a \le g(t) \le b$.

In the latter case, the function $g(t)$ has the form described above, and the corresponding number t_0 is obviously non-negative. Therefore, $d \le t_0 + d$ and, consequently, either $g(d) < a$ (if $d < t_0$), or $g(d) = d - t_0 + a = b - t_0 \le b$ (if $t_0 \le d$). Thus, in all cases $g(d) \le b$, i.e., $\varphi_p(d) \in [f \le b]$.

Conversely, let a point $p \in M$ have the property that $\varphi_p(d) \in [f \le b]$, i.e., that $g(d) \le b$. Then either $g(t) < a$ for any $t \in \mathbb{R}$ and, in particular, for $t = 0$, or the function $g(t)$ has the form described above, and the corresponding number t_0 is non-negative (since if $g(d) < a$, then $d < t_0$, and if $a \le g(d) \le b$, then $d - t_0 + a = b - t_0 \le b_0$). Therefore, in this case too, $g(0) \le a$, i.e., $\in [f \le a]$.

Let us now consider the diffeomorphism

$$\varphi_d : M \to M,$$

defined by the formula

$$\varphi_d(p) = \varphi(d, p), \quad p \in M.$$

Since for any point $p \in M$ the following equality holds

$$\varphi_d(p) = \varphi_p(d)$$

then, according to what has just been proved, this diffeomorphism maps the set $[\le a]$ onto the set $[\le b]$. Therefore, we can take it for the desired diffeomorphism Φ.

Next, for any point $p \in [f \le b]$ and any number $t \in I$, we set

$$r_t(p) = \begin{cases} p, & \text{if} \quad p \in [f \le a], \\ \varphi_p(t(a - f(p))), & \text{if} \quad p \in [a \le f \le b]. \end{cases}$$

It is easy to see that this gives us a homotopy

$$r_t : [f \le b] \to [f \le b],$$

for which the map r_0 is the identity map of the space $[f \le b]$, and the map r_1 is a retraction map of this space onto the subspace $[f \le a]$.

The homotopy r_t can be visually described as a "shift" of the set $[f \le b]$ onto the set $[f \le a]$ along the "flow lines" of the function f.

Thus, the above statements are completely proven.

Let us now consider an arbitrary critical value c of the function f. Let this critical value

1) be isolated in the sense that for some $\varepsilon_0 > 0$, the interval $[c - \varepsilon_0, c + \varepsilon_0]$ does not contain any other critical values of f;

2) be compact in the sense that for some $\varepsilon_0 > 0$, the set $[c - \varepsilon_0 \le f \le c + \varepsilon_0]$ is compact (clearly, under both conditions, we can consider the number $\varepsilon_0 > 0$ to be the same);

3) be non-degenerate in the sense that all critical points $p_1, \ldots, p_n$ of f located on the level surface $[f = c]$ are non-degenerate. (By condition 2, the number of these points is finite.)

Remark 6.35. It is easy to see that condition 1) follows from conditions 2) and 3). Perhaps it is also worth noting that, generally speaking, it does not follow from condition 3) *alone* (even if the number of critical points of the function f on the level surface $[f = c]$ is finite).

It is clear that the points $p_1, \ldots, p_n$ have pairwise disjoint coordinate neighbourhoods $U_1, \ldots, U_n$ such that

a) for any $i = 1, \ldots, n$, the local coordinates on the neighbourhood U_i are restrictions of some smooth functions $x^1, \ldots, x^m$ on M (and independent of i);

b) on each neighbourhood U_i, $i = 1, \ldots, n$, the coordinates $x^1, \ldots, x^m$ satisfy the conditions of the Morse lemma, i.e., in this neighbourhood the following equality holds

$$f = c - (x^1)^2 - \cdots - (x^{h_i})^2 + (x^{h_i+1})^2 + \cdots + (x^m)^2,$$

where h_i is the index of the critical point p_i;

c) the coordinate homeomorphism of the neighbourhood U_i, $i = 1, \ldots, n$, into the space $\mathbb{R}^m$, corresponding to the local coordinates $x^1, \ldots, x^m$, maps the neighbourhood U_i onto the open ball $|\boldsymbol{t}|^2 < 2\varepsilon_0$, where ε_0 is the positive number stipulated by conditions 1) and 2) above.

To simplify further calculations, it is convenient to introduce on the set

$$U = U_1 \cup \cdots \cup U_n$$

smooth functions defined on U_i, $i = 1, \ldots, n$, by the formulae

$$\alpha = (x^1)^2 + \cdots + (x^{h_i})^2, \quad \beta = (x^{h_i+1})^2 + \cdots + (x^m)^2.$$

According to the above, the function f on the set U_i is expressed by the formula:

$$f = c - \alpha + \beta.$$

Let ε be an arbitrary positive number less than ε_0. We will henceforth assume this number to be fixed. Since $\varepsilon < \varepsilon_0$, the inequality

$$\alpha + \beta \leq 2\varepsilon$$

obviously defines, in each neighbourhood U_i, a subset homeomorphic to a closed n-dimensional ball.

Let us now introduce the function g, defined by the formulae:

$$g(p) = \begin{cases} f(p) - \mu(\alpha(p) + \beta(p)), & \text{if} \quad p \in U, \\ f(p), & \text{if} \quad p \notin U, \end{cases}$$

where μ is a smooth numerical function such that

$$\mu(0) > \varepsilon, \quad \mu(t) = 0 \text{ for } t \leq \varepsilon, \quad -1 < \mu^{'}(t) \leq 0 \text{ for all } t \in \mathbb{R}.$$

For example, we can assume that

$$\mu(t) = \begin{cases} \varepsilon\nu\left(\frac{\varepsilon - t}{\varepsilon}\right), & \text{if} \quad t \leq 2\varepsilon, \\ 0, & \text{if} \quad t \geq 2\varepsilon, \end{cases}$$

where $\nu(t) = \frac{t}{2}\exp\left(\frac{q}{3} - \frac{1}{t^2}\right)$. It is clear that the function g constructed in this way is a smooth function on M (since for $\alpha(p) + \beta(p) \geq 2\varepsilon$ it is necessary that $\alpha(p) + 2\beta(p) \geq 2\varepsilon$ and therefore $\mu(\alpha(p) + \beta(p)) = 0$). Let us show that

Proposition 6.36. *the set $[f \leq c + \varepsilon]$ coincides with the set $[g \leq c + \varepsilon]$.*

Proof. Indeed, by definition, we have $g \leq f$. Therefore, $[f \leq c+\varepsilon] \subset [g \leq c+\varepsilon]$. Conversely, let $p \in [g \leq c + \varepsilon]$, i.e., let $g(p) \leq c + \varepsilon$. If $\alpha(p) + 2\beta(p) \leq 2 \leq \varepsilon$, then

$$f(p) = c - \alpha(p) + \beta(p) \leq c + \frac{\alpha(p)}{2} + \beta(p) \leq c + \varepsilon,$$

i.e.,

$$p \in [f \leq c + \varepsilon].$$

If $\alpha(p) + 2\beta(p) \geq 2\varepsilon$, then $f(p) = g(p)$ and therefore again $p \in [f \leq c + \varepsilon]$. □

Further, it is easy to see that

Proposition 6.37. *the critical points of the function g coincide with the critical points of the function f.*

Proof. Indeed, since $f = g$ outside the set U, to prove this statement it is sufficient to show that the points $p_1, \ldots, p_n$ are critical points of the function g and that in U_i for any $i = 1, \ldots, n$ this function has no other critical points. But in this neighbourhood

$$\begin{aligned} dg &= d - \mu^{'}(\alpha + 2\beta)(d\alpha + 2d\beta) \\ &= (-1 - \mu^{'}(\alpha + 2\beta))d\alpha + (1 - 2\mu^{'}(\alpha + 2\beta))d\beta \\ &= 2(-1 - \mu^{'}(\alpha + 2\beta))(x^1 dx^1 + \cdots + (x^{h_i} dx^{h_i}) \\ &+ 2(1 - 2\mu^{'}(\alpha + 2\beta))(x^{h_i+1} dx^{h_i+1} + \cdots + (x^m dx^m), \end{aligned}$$

Moreover, due to the conditions imposed above on the function μ, the expressions $-1 - \mu^{'}(\alpha + 2\beta)$ and $1 - 2\mu^{'}(\alpha + 2\beta)$ are everywhere nonzero. Therefore, the point $p \in U_i$ is a critical point of the function g if and only if $x^1(p) = \cdots = x^m(p) = 0$, i.e., when $p = p_i$. □

Now consider the set $[c - \varepsilon \leq g \leq c + \varepsilon]$. Since $f \leq c + \varepsilon] = [g \leq c + \varepsilon]$ and $g \leq$, we have

$$[c - \varepsilon \leq g \leq c + \varepsilon] \subset [c - \varepsilon \leq f \leq c + \varepsilon].$$

Therefore, the set $[c-\varepsilon \le g \le c+\varepsilon]$ is compact. Furthermore, since $g(p_i) = c - \mu(0) < c-\varepsilon$, this set does not contain the points $p_1, \ldots, p_n$, and, consequently, does not contain any critical points of the function f, and hence of the function g. Therefore, according to what was proved above,

Proposition 6.38. *the set $[g \le \varepsilon]$ is a deformation retract of the set $[g \le c+\varepsilon] = [f \le c+\varepsilon]$.*

Remark 6.39. By similar considerations, the set $[g \le c]$ is also a deformation retract of the set $[f \le c+\varepsilon]$.

This is also true in the case when conditions 1) and 2) are satisfied only "from above", i.e., if there exists $\varepsilon_0 > 0$ such that the segment $[c, c+\varepsilon_0]$ does not contain critical values of the function f different from x, and the set $[c \le f \le c+\varepsilon_0]$ is compact.

Let Q be the set of points $p \in U$ for which

$$\alpha(p) < \varepsilon, \quad \beta(p) = 0.$$

It is clear that for any $i = 1, \ldots, n$ the intersection $Q \cap U_i$ is homeomorphic to the unit open ball E^{h_i} of the space $\mathbb{R}^{h_i}$. Its closure $\overline{Q \cap U_i}$ consists of points $p \in U_i$ for which

$$\alpha(p) \le \varepsilon, \quad \beta(p) = 0.$$

and is therefore homeomorphic to the closed ball $\mathbb{E}^{h_i}$. Thus, the set Q is the union of n disjoint cells

$$e^{h_1} = Q \cap U_1, \ \ldots, \ e^{h_n} = Q \cap U_n$$

of dimensions $h_1, \ldots, h_n$, respectively. For any $i = 1, \ldots, n$, the boundary $\dot{e}^{h_i}$ of the cell e^{h_i} consists of points $p \in U_i$ for which

$$\alpha(p) = \varepsilon, \quad \beta(p) = 0,$$

and therefore coincides with the intersection $[f \le c-\varepsilon] \cap \overline{e}^{h_i}$. This means that

Proposition 6.40. *the set $[f \le c-\varepsilon] \cup Q$ can be viewed as the result of gluing n closed balls of dimensions $h_1, \ldots, h_n$ to the set $[f \le c-\varepsilon]$ by means of some continuous maps of their boundaries into the set $[f \le c-\varepsilon]$.*

Since $\beta = 0$ on Q, then

$$g = c - \alpha - \mu(\alpha) \text{ on } Q, \text{ where } 0 \le \alpha < \varepsilon.$$

On the other hand,

$$\alpha + \mu(\alpha) = \mu(0) + \int_0^\alpha (1 + \mu'(t))dt > \mu(0) > \varepsilon,$$

since, by hypothesis, $1 + \mu'(t) > 0$. Therefore,

$$g \le c - \varepsilon \text{ on } Q,$$

i.e., $Q \subset [g \leq c - \varepsilon]$. Therefore,

$$[f \leq c - \varepsilon] \cup Q \subset [g \leq c - \varepsilon].$$

We now show that

Proposition 6.41. *the set* $[f \leq c - \varepsilon] \cup Q$ *is a deformation retract of the set* $[g \leq c - \varepsilon]$.

Proof. To this end, we consider the intersection F of the set $[g \leq c - \varepsilon]$ with the set $\alpha + \beta \leq 2\varepsilon$ (which, as we know, is the union of n disjoint closed m-dimensional balls). This intersection is contained in the set U and is defined by the inequalities

$$\alpha + \beta \leq 2\varepsilon, \quad \alpha - \beta + \mu(\alpha + 2\beta) \geq \varepsilon.$$

We can represent the set F as the union of three sets F_1, F_2, and F_3, the first of which is defined by the inequalities

$$\alpha \leq \varepsilon, \quad \alpha - \beta + \mu(\alpha + 2\beta) \geq \varepsilon,$$

the second by the inequalities

$$\varepsilon \leq \alpha \leq \beta + \varepsilon, \quad \alpha - \beta + \mu(\alpha + 2\beta) \geq \varepsilon,$$

and the third by the inequalities

$$\beta + \varepsilon \leq \alpha, \quad \alpha + \beta \leq 2\varepsilon$$

(see Fig. 6.1).

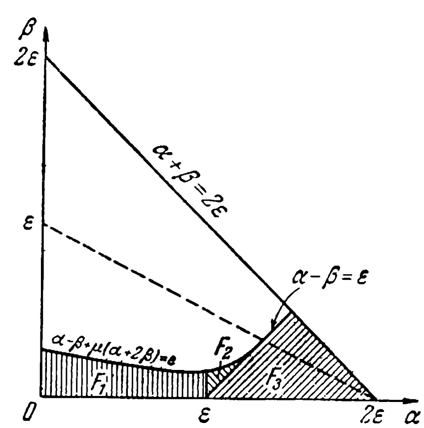


Figure 6.1:

Let $0 \le t \le 1$. To each point $p \in F$ we assign a point $q = r_t(p) \in U$, located in the same neighbourhood U_i as the point p, whose coordinates $x^1(q), \dots, x^m(q)$ are related to the coordinates $x^1(p), \dots, x^m(p)$ of the point p by the relations:

$$
\begin{gathered}
x^1(q) = x^1(p), \\
\dots \\
x^h(q) = x^h(p), \\
x^{h+1}(q) = s_t x^{h+1}(p), \\
\dots \\
x^m(q) = s_t x^m(p),
\end{gathered}
$$

where

$$
s_t = \begin{cases} 1 - t, & \text{if} \quad p \in F_1, \\ 1 - t + t\sqrt{\frac{\alpha(p) - \varepsilon}{\beta(p)}}, & \text{if} \quad p \in F_2, \\ 1, & \text{if} \quad p \in F_3. \end{cases}
$$

As is easy to see, despite the fact that for $\alpha = \varepsilon$ and $\beta = 0$ the function s_t has a discontinuity, the point $r_t(p)$ depends continuously on t and p, i.e., the family of maps

$$
r_t : F \longrightarrow U
$$

is a homotopy.

Since for any point p F the following equalities hold:

$$
\alpha(r_t(p)) = \alpha(p), \quad \beta(r_t(p)) = s_t^2 \beta(p),
$$

then

$$
\begin{aligned}
&\alpha(r_t(p)) + \beta(r_t(p)) = \alpha(p) + s_t^2\beta(p) \le= \alpha(p) + \beta(p) \le 2\varepsilon, \\
&\alpha(r_t(p)) - \beta(r_t(p)) + \mu(\alpha(r_t(p)) + 2\beta(r_t(p))) = \\
&\alpha(p) - s_t^2\beta(p) + \mu(\alpha(r_t(p)) + 2s_t^2\beta(p)\beta(r_t(p)) \ge \\
&\alpha(p) - \beta(p) + \mu(\alpha(p) + 2\beta(p)) \ge \varepsilon,
\end{aligned}
$$

since $0 \le s_t^2 \le 1$. Thus, $r_t(p) \in F$ for any $t \in I$, i.e., in other words, the family r_t is a homotopy of maps of the set F into itself. Since $r_t(p) = p$ for $\alpha(p) + \beta(p) = 2\varepsilon$, we can extend this homotopy to a homotopy

$$
r_t : [g \le c - \varepsilon] \longrightarrow [g \le c - \varepsilon],
$$

assuming that $r_t(p) = p$ for $p \notin F$.

It is clear that for $t = 0$ the map r_t is the identity map of the set $[g \le c - \varepsilon]$, i.e. the homotopy r_t is a deformation of the set $[g \le c - \varepsilon]$. Consider the map r_t.

Let $p \in F$. If $p \in F_1$ then $s_1 = 0$, and therefore $\beta(r_1(p)) = 0$. Since, in addition, $\alpha(r_1(p)) = \alpha(p) \le \varepsilon$, then, therefore, $r_1(p) \in Q$. If $p \in F_2$, then $s_1^2 = \frac{\alpha(p)-\varepsilon}{\beta(p)}$ and therefore $\beta(r_1(p))) = \alpha(p) - \varepsilon$, i.e. $f(r_1(p)) = c - \alpha(r_1(p)) + \beta(r_1(p)) = c - \varepsilon$. Finally, if $p \in F_3$, then $s_1 = 1$ and therefore $f(r_1(p)) = c - \alpha(p) + \beta(p) < c - \varepsilon$. Thus, for any point $p \in F$, the inclusion $r_1(p) \in [f \leq c - \varepsilon] \cup Q$ holds. Since the sets $[g \leq c - \varepsilon]$ and $[f \leq c - \varepsilon]$ coincide outside the set F, this inclusion remains valid for $p \notin F$. Moreover, if $p \in [f \leq c - \varepsilon] \cup Q$, i.e., if either $f(p) \leq c - \varepsilon$ or $\alpha(p) < \varepsilon$ and $\beta(p) = 0$, then, as is easy to see, $r_1(p) = p$. Consequently, the map r_1 is a retraction of the set $[g \leq c - \varepsilon]$ onto the set $[f \leq c - \varepsilon] \cup Q$. Thus, the statement formulated above is completely proven. □

Remark 6.42. Similarly, one can prove that the set $[f \leq c]$ is a deformation retract of the set $[g \leq c]$.

Indeed, it suffices to set $\varepsilon = 0$ in the proof given above. In this case, the set F_1 can be ignored (since it will consist of only one point). The set F_2 will be defined by the inequalities

$$\alpha \leq \beta, \quad \alpha - \beta + \mu(\alpha + 2\beta) \geq 0,$$

the set $F3$ by the inequalities

$$\beta \leq \alpha, \quad \alpha - \beta + \mu(\alpha + 2\beta) \geq 0,$$

and the numbers s_t by the formula

$$s_t = \begin{cases} 1 - t + t\sqrt{\frac{\alpha(p)}{\beta(p)}}, & \text{if} \quad p \in F_2, \\ 1, & \text{if} \quad p \in F_3. \end{cases}$$

Comparing the assertions proved above, we immediately obtain that

Proposition 6.43. *under assumptions 1), 2), and 3), the set $[f \leq c + \varepsilon]$ contains a subset $[f \leq c - \varepsilon] \cup Q$, which is obtained from the set $[f \leq c - \varepsilon]$ by gluing n balls of dimensions $h_1, \ldots, h_n$ (via certain maps of their boundaries) and is a deformation retract of the set $[f \leq c + \varepsilon]$.*

In particular,

Proposition 6.44. *under assumptions 1), 2), and 3), the set $[f \leq c + \varepsilon]$ is homotopically equivalent to the set $[f \leq c - \varepsilon]$, to which n balls of dimensions $h_1, \ldots, h_n$, respectively, are glued.*

This statement was first proved by Morse and is known as Morse's theorem.

Remark 6.45. Combining the above remarks, we similarly obtain that under assumptions 1), 2), and 3), the set $[f \leq c]$ is a deformation retract of the set $[f \leq c + \varepsilon]$.

Moreover, in accordance with the above, we can assume that conditions 1) and 2) are satisfied only from above.

By Morse's theorem, it follows directly that

Proposition 6.46. *under assumptions 1), 2), and 3), the set* $[f \le c]$ *is homotopy equivalent to the set* $[f \le c - \varepsilon] \cup Q$.

Proof. Indeed, as proved above, the set $[f \le c]$ is homotopy equivalent to the set $[f \le c-\varepsilon]$, which in turn is homotopy equivalent to the set $[f \le c-\varepsilon] \cup Q$. □

6.4 Smooth manifolds and cellular decompositions

It follows from the results of the previous subsection that

Proposition 6.47. *if a smooth function* f *on a separable smooth manifold* M *and a number* $a \in \mathbb{R}^n$ *have the property that*

1) for some $\varepsilon > 0$, *the set* $[f \le a + \varepsilon]$ *is compact;*

2) the set $[a < f \le a + \varepsilon]$ *does not contain critical points of* f;

3) all critical points of f *belonging to the set* $[f \le a]$ *are non-degenerate,*

then the set $[f \le a]$ *is homotopy equivalent to some cellular decomposition* K *whose cells are in one-to-one correspondence with the critical points of* f *belonging to the set* $[f \le a]$, *where the dimension of each cell is equal to the index of the corresponding critical point.*

Proof. Indeed, it is clear that the function f has only a finite number of critical values not exceeding a given number a, and all of these critical values

$$c_0 < c_1 < \cdots < c_n$$

are isolated, non-degenerate, and compact (see §6.3). Therefore, according to Morse's theorem, there exist numbers

$$\varepsilon_0 > 0, \varepsilon_1 > 0, \ \ldots, \ \varepsilon_n > 0,$$

such that for any $i = 0, 1, \ \ldots, \ n$, the set $[f \le c_i + \varepsilon_i]$ is homotopy equivalent to the set $[f \le c_i - \varepsilon_i]$, to which some number of closed balls are attached. Moreover, without loss of generality, we can obviously assume that for any $i = 0, 1, \ \ldots, \ n-1$, the inequality

$$c_i + \varepsilon_i < c_{i+1} - \varepsilon_{i+1},$$

and for $i = n$, the inequality

$$c_n + \varepsilon_n < a + \varepsilon_0$$

Then, by what was proved in the previous paragraph, the set $[f \le c_i + \varepsilon_i]$ for $i = 0, 1, \ \ldots, n-1$ will be a deformation retract of the set $[c_{i+1} - \varepsilon_{i+1}]$, and the set $[f \le c_n + \varepsilon_n]$ will be homotopy equivalent to the set $f \le a]$.

We now show that

Proposition 6.48. *there exists an increasing sequence*

$$K_0 \subset K_1 \subset \cdots \subset K_n$$

of finite cellular decompositions K_i, $i = 0, 1, \ldots, n$, *each of which is a subdecomposition of the next decomposition, and homotopy equivalences*

$$\varphi_i : [f \le c_i + \varepsilon_i] \to K_i, \quad i = 0, 1, \ldots, n,$$

such that for any $i = 0, 1, \ldots, n-1$, *the equivalence* φ_i *is the restriction of the equivalence* φ_{i+1}*:*

$$\varphi_i = \varphi_{i+1}|_{[f \le c_i + \varepsilon_i]}.$$

To this end, we note that since the number c_0 is obviously the minimum of the function f, the set $[f \le c_0 - \varepsilon_0]$ is empty. Therefore, the set $[f \le c_0 + \varepsilon_0]$ is homotopically equivalent to the topological sum K_0 of some (finite) number of closed balls. Let

$$\varphi_0 : [f \le c_0 + \varepsilon_0] \to K_0$$

be the corresponding homotopy equivalence.

Arguing by induction, suppose that for some /=^{1}, ..., n the partition Kui and the homotopy equivalence

$$\varphi_{i-1} : [f \le c_{i-1} + \varepsilon_{i-1}] \to K_{i-1}$$

have already been constructed.

Since, according to what was said above, the set $[f \le c_{i-1} + \varepsilon_{i-1}]$ is a deformation retract of the set $[f \le c_i - \varepsilon_i]$, we can extend the equivalence φ_{i-1} to some homotopy equivalence

$$\psi_i : [f \le c_i - \varepsilon_i] \to K_{i-1}$$

On the other hand, according to Morse's theorem, the set $[f \le c_i + \varepsilon_i]$ contains a subset of the form $[f \le c_i - \varepsilon_i] \cup Q$, where Q is the topological sum of some number of closed balls, which is its deformation retract. Let

$$r_i : [f \le c_i + \varepsilon_i] \to [f \le c_i - \varepsilon_i] \cup Q$$

be the corresponding retracting map.

Consider the union $\dot{Q}$ of the boundary spheres of all balls whose sum is the set Q. Let

$$\dot{\psi}_i : \dot{Q} \to K_{i-1}$$

be the restriction of the map ψ_i to the subspace $\dot{Q} \subset [f \le c_i - \varepsilon_i]$. According to the results of §2.3, we can extend the homotopy equivalence ψ_i to some homotopy equivalence

$$\overline{\psi}_i : [f \le c_i - \varepsilon_i] \cup Q \to K_{i-1} \cup_{\dot{\psi}_i} Q.$$

On the other hand, according to the results of §3.4, the space $K_{i-1} \cup_{\psi_i} Q$ is homotopy equivalent to some cellular decomposition K_i, containing the decomposition K_{i-1} as a subdivision. Moreover, we can choose the homotopy equivalence

$$\omega_i : K_{i-1} \cup_{\psi_i} Q \to K_i.$$

such that it is the identity map on the decomposition K_{i-1}.

We set

$$\varphi_i : \omega_i \circ \overline{\psi} \circ r_i : [f \leq c_i + \varepsilon_i] \to K_i.$$

Clearly, this map is a homotopy equivalence, coinciding on the sub-decomposition K_{i-1} with the equivalence φ_{i-1}. Thus, the existence of decompositions K_i and equivalences φ_i is proven for all $i = 0, 1, \ldots, n$ and, in particular, for $i = n$. To prove the proposition formulated at the beginning of this section, it remains to set $K = K_n$. □

From the proved proposition it easily follows that

Proposition 6.49. *every smooth function of compact character bounded below on a smooth separable manifold M, having only non-degenerate critical points, defines a homotopy equivalence of the manifold M with some cellular decomposition K, whose cells are in one-to-one correspondence with the critical points of the function f, where the dimension of each cell is equal to the index of the corresponding critical point.*

Proof. Indeed, it is clear that every number $a \in \mathbb{R}$ satisfies (together with the function f) the conditions of the previous proposition. Therefore, for any $n \geq 0$, there exists some homotopy equivalence

$$\varphi_i : [f \leq n] \to K_n,$$

where K_n is a finite cellular decomposition whose cells are in one-to-one correspondence with the critical points of the function f belonging to the set $[f \leq n]$. Moreover, it is clear that the decompositions K_n can be chosen so that

$$K_0 \subset K_0 \subset \cdots \subset K_n \subset K_{n+1} \subset \cdots$$

and so that for any $n \geq 0$ the decomposition K_n is a sub-decomposition of the decomposition K_{n+1}, and the homotopy equivalence φ_n coincides with the restriction $\varphi_n|_{K_n}$ of the homotopy equivalence φ_{n+1} to the decomposition K_n.

Let K be the free union of decompositions K_n, $n \geq 0$. By the results of §3.2, this union is a cellular decomposition containing the decompositions K_n as sub-decompositions. We define the map

$$\varphi : M \to K,$$

assuming that

$$\varphi|_{[f \leq n]} = \varphi_n$$

for any $n \geq 0$. Clearly, this map is uniquely defined and continuous.

Since the manifold M is paracompact and is the union of the interiors $[f < n]$ of sets $[f \leq n]$, it is a homotopy limit of sets $[f \leq n]$. The decomposition K is also a homotopy limit of the decompositions K_i. Since for any $n \geq 0$ the map is, by construction, a homotopy equivalence, it follows from the results of §2.5 2 that the map φ is also a homotopy equivalence. □

Since, according to §6.2, a function f satisfying the conditions of this proposition exists on any separable smooth manifold, we, in particular, obtain from this that

Proposition 6.50. *every separable smooth manifold is homotopy equivalent to some countable cellular decomposition.*

In the case where the function f is defined in a more or less observable way on the manifold under consideration M, one can, from the proved proposition, obtain quite complete information about the homotopy type of the manifold.

Example 6.51. Example. Let $\mathbb{C}^{n+1}$ be a complex $n+1$-dimensional arithmetic space whose points are n+1-tuple $(z^0, z^1, \ldots, z^n)$ of complex numbers. Consider the unit sphere $\mathbb{S}^{2n+1}$ of this space, i.e., the set of all points $(z^0, z^1, \ldots, z^n) \in \mathbb{C}^{n+1}$ for which

$$|z^0|^2 + |z^1|^2 + \cdots + |z^n|^2 = 1.$$

On this sphere, we define an equivalence relation by considering two points $(z^0, z^1, \ldots, z^n)$ and $(\tilde{z}^0, z^1, \ldots, \tilde{z}^n)$ are equivalent if and only if there exists a complex number $\lambda \neq 0$ (whose modulus is necessarily equal to unity) such that

$$\tilde{z}^i = \lambda z^i$$

for all $i = 0, 1, \ldots, n$. The corresponding quotient space of the sphere $\mathbb{S}^{2n+1}$ is called the *complex projective space of complex dimension n*. We will denote it by the symbol $\mathbb{C}P^n$. The point of this space containing the point $(z^0, z^1, \ldots, z^n) \in \mathbb{S}^{2n+1}$ will be denoted by the symbol $(z^0 : z^1 : \cdots : z^n)$. Clearly, the space $\mathbb{C}P^n$ is a compact Hausdorff space.

Let U_i, $i = 0, 1, \ldots, n$, be the set of all points $(z^0, z^1, \ldots, z^n) \in \mathbb{C}^{n+1}$ for which $z^i \neq 0$. (Clearly, this condition is invariant, i.e., it does not depend on the choice of the representative $(z^0, z^1, \ldots, z^n) \in \mathbb{S}^{2n+1}$ of the point $(z^0 : z^1 : \cdots : z^n)$.) Each point $(z^0 : z^1 : \cdots : z^n) \in U_i$ has one and only one representative $(z^0, z^1, \ldots, z^n) \in \mathbb{S}^{2n+1}$ for which $z^i = 1$. This establishes a one-to-one correspondence between the points of the set U_i and the points of the space $\mathbb{C}^n$. It is easy to verify that this correspondence is a homeomorphism. Since the space $\mathbb{C}^n$ is naturally homeomorphic to the space $\mathbb{R}^{2n}$ (the corresponding homeomorphism maps each point $(z^1, \ldots, z^n) \in \mathbb{C}^n$ to the point $(x^1, y^1, \ldots, x^n, y^n) \in \mathbb{R}^{2n}$, where x^1 is the real part and y^1 is the imaginary part of the complex number z^i, $i = 1, \ldots, n$), thereby a homeomorphism is defined

$$\xi_i : U_i \to \mathbb{R}^{2n}, \quad i = 0, 1, \ldots, n.$$

For example, the homeomorphism $\xi_0 : U_0 \to \mathbb{R}^{2n}$ maps the point $(z^0 : z^1 : \cdots : z^n) \in U_i$, to the point $x^1,^1, \ldots, x^n, y^n \in \mathbb{R}^{2n}$, for which

$$x^1 + iy^1 = \frac{z^1}{z^0}, \ \ldots, \ x^n + iy^n = \frac{z^n}{z^0}.$$

Let

$$\pi : \mathbb{S}^{2n+1} \to \mathbb{C}P^n$$

be the natural projection (associating the point $(z^0, z^1, \ldots, z^n) \in \mathbb{S}^{2n+1}$ with the point $(z^0 : z^1 : \cdots : z^n) \in\in \mathbb{C}P^n$). We call a function f on the space $\mathbb{C}P^n$ *smooth* if $\circ\pi$ is a smooth function on the sphere $\mathbb{S}^{2n+1}$. It is easily verified that the set $\mathcal{O}(\mathbb{C}P^n)$ of all functions smooth in this sense is a smoothness on the space $\mathbb{C}P^n$ and that the space $\mathbb{C}P^n$ equipped with this smoothness is a smooth $2n$-dimensional manifold. (The homeomorphisms ξ_i, $i = 0, 1, \ldots, n$ constructed above are diffeomorphisms with respect to this smoothness and can therefore be taken as coordinate diffeomorphisms.)

Consider the function f on the manifold $\mathbb{C}P^n$ defined by the formula

$$f(z^0 : z^1 : \cdots : z^n) = c_0|z^0|^2 + c_1|z^1|^2 + \ldots + |z^n|^2,$$

where $c_0,^1, \ldots, c_n$, are some pairwise distinct real numbers. Clearly, this function is uniquely defined and is smooth on the manifold $\mathbb{C}P^n$.

On the coordinate neighbourhood U_0, this function is expressed through the local coordinates constructed above

$$x^1, y^1, \ \ldots, \ x^n, y^n$$

according to the formula

$$f(z^0 : z^1 : \cdots : z^n) = c_0 + (c_1 - c_0)((x^1)^2 + (y^1)^2) + \cdots + (c_n - c_0)((x^n)^2 + (y^n)^2).$$

Consequently, in the neighbourhood U_0, the function f has one and only one critical point

$$p_0 = (1 : 0 : \cdots : 0).$$

This point is non-degenerate, and its index is equal to twice the number of coefficients c_i, with the property that $c_i < c_j$.

Similarly, in the coordinate neighbourhood U_j, $j = 1, \ldots, n$, the function f has one and only one critical point

$$p_j = (0 : \cdots : 0 : 1 : 0 : \cdots : 0), \quad i = 1, \ \ldots, \ n.$$

This point is non-degenerate, and its index is equal to twice the number of coefficients c_i with the property that $c_i < c_j$.

Thus, the function f has $n+1$ critical points, all of which are non-degenerate. Moreover, among the indices of these points, each even number from 0 to $2n$ occurs (exactly once).

By the above remark, it is thus proved that

Proposition 6.52. *the manifold $\mathbb{C}P^n$ is homotopy equivalent to some cellular partition containing $n+1$ cells*

$$e^0, e^2, \ \dots, \ e^{2n}$$

of dimensions $0, 2, \ \dots, \ 2n$, respectively.

Remark 6.53. A closer examination of the structure of the manifold $\mathbb{C}P^n$ reveals that it is not only homotopically equivalent, but even homeomorphic to the indicated cellular decomposition. In general, by way of refining the general proposition proved in this section, it can be shown that

Proposition 6.54. *any smooth separable manifold is the body of some countable cellular decomposition.*

However, as long as we are only interested in the homotopy properties of manifolds, this refinement is useless. For this reason, and also because its proof is quite difficult, we will not dwell on it.

6.5 Bott's theorem

In all the preceding discussions, the assumption of non-degeneracy of the critical points of the function f under consideration played a significant role. In this section, we will show that, in a somewhat weakened form (more sufficient for most applications), the main result of §6.3 remains valid even under the assumption that the critical points of f are only weakly non-degenerate. Namely, we will show that

Proposition 6.55. *if a critical value c of f is isolated, compact, and has the property that all critical points of f located on a level surface $[f = c]$ are weakly non-degenerate, then for some $\varepsilon > 0$, the set $[f \leq c + \varepsilon]$ contains a subset of the form $[f \leq c - \varepsilon] \cup Q$, where Q is the topological sum of a finite number of closed balls, which is its deformation retract; moreover, the dimension of the balls that make up the set Q is not less than the smallest of the indices of the critical points $p \in [f = c]$.*

Clearly, to prove this assertion (first proved by Bott), it suffices to prove that

Proposition 6.56. *on the manifold M there exists a smooth function g such that*

1) for some $\varepsilon > 0$, the following equalities hold:

$$[g \leq c - \varepsilon] = [f \leq c - \varepsilon], \quad [g \leq c + \varepsilon] = [f \leq c + \varepsilon];$$

2) the set

$$[c - \varepsilon \leq g \leq c + \varepsilon] = [c - \varepsilon \leq f \leq c + \varepsilon]$$

is compact and contains only non-degenerate critical points of g;

3) *all these critical points are located on the level surface* $[f = c]$, *and the index of each of them is not less than the smallest index of the critical points* $p \in [f = c]$ *of* f.

The condition of weak non-degeneracy of critical points $p \in [f = c]$ of the function f means that each of these critical points belongs to some non-degenerate critical manifold. Since the set $[f = c]$ is compact, each of these manifolds is compact and their number is finite. Let

$$N_1, \ \ldots, \ N_s$$

be all non-degenerate critical manifolds of the function f located on the level surface $[f = c]$, and let

$$h_1, \ \ldots, \ h_s$$

be their indices. We set

$$h = \min(h_1, \ \ldots, \ h_s).$$

According to the Whitney embedding theorem (see §5.6 and §5.7), we can assume without loss of generality that the manifold M, and therefore the manifold

$$N = N_1 \cup \cdots \cup N_s$$

are closed submanifolds of some Euclidean space $\mathbb{R}^t$.

By hypothesis, for the function f there exists a number $\varepsilon > 0$ such that the set $[c - \varepsilon \leq fc + \varepsilon]$ is compact and does not contain any critical points of this function that are different from the points of the manifold N. Let W be an arbitrary tubular neighbourhood of N in the space $\mathbb{R}^t$ with the property that

$$W \cap M \subset [c - \frac{\varepsilon}{2} \leq f \leq c + \frac{\varepsilon}{2}].$$

Furthermore, let W_1 and W_2 be (necessarily tubular) neighbourhoods of N in the space $\mathbb{R}^t$ with compact closures $\overline{W}_1$ and $\overline{W}_2$ such that $\overline{W}_1 \subset W_2$ and $\overline{W}_2 \subset W$.

According to the definition of non-degenerate critical manifolds, the manifold N has a neighbourhood U in the manifold M such that no point of this neighbourhood, distinct from points of the manifold N, is a critical point of the function f. Consider the set

$$V = U \cap W_1 \cap M.$$

This set

1) contains the submanifold N;

2) is open in the manifold M;

3) is contained in the compact set $[c - \frac{\varepsilon}{2} \leq f \leq c + \frac{\varepsilon}{2}]$;

4) is the union of sets
$${}_pV = U \cap {}_pW_1 \cap M \quad p \in N,$$
where ${}_pW_1$ are the fibres of the tubular neighbourhood W_1.

We will call the sets ${}_pV$ the *fibres* of the neighbourhood V. Clearly, each fibre ${}_pV$ is a submanifold of M complementary to the submanifold N at the point p.

By construction, for each point $q \in V \setminus N$, the differential df_q of the function f is nonzero. Furthermore, it is easy to see that (reducing the neighbourhood U if necessary) we can, without loss of generality, assume that this differential is non-zero already in the subspace $({}_pV)$ of M_q, where ${}_p$ is the fibre of the neighbourhood V containing the point q.

Assuming that the manifold M is defined as a Riemannian space, consider the gradient $\operatorname{grad} f$ of the function f. By hypothesis, this gradient is non-zero at every point of the set $[c - \varepsilon \le f \le c + \varepsilon] \setminus V$. Therefore, since the set $[c - \varepsilon \le f \le c + \varepsilon] \setminus V$ is compact, there exists a number $K > 0$ such that

$$|\operatorname{grad} f| \ge K \quad \text{on } [c - \varepsilon \le f \le c + \varepsilon] \setminus V.$$

According to the theorem proved in §6.3, on the manifold N there exists a smooth function $\overline{\varphi}$ with only non-degenerate critical points. We extend this function to the neighbourhood W, assuming that on each fibre of this neighbourhood the function $\overline{\varphi}$ is constant. Since the set $\overline{W_1}$ is compact and is contained in the neighbourhood $W_2 \subset W$, there exists a smooth function $\overline{\overline{\varphi}}$ on the space $\mathbb{R}^t$ that coincides with the function $\overline{\varphi}$ on the neighbourhood W_1 and vanishes outside the neighbourhood W_2. The restriction

$$\varphi = \overline{\overline{\varphi}}|_M$$

of the function $\overline{\overline{\varphi}}$ on the manifold M is a smooth function on M with the following properties:

a) the function $\varphi|_N$ has (on N) only non-degenerate critical points;

b) on each fibre ${}_pV$, $p \in N$, of a neighbourhood V, the function φ is constant;

c) the support $\overline{[\varphi \neq 0]}$ of the function φ is compact and contained in the set $[c - \frac{\varepsilon}{2} \le f \le c + \frac{\varepsilon}{2}]$;

From property b) it follows directly that

d) at any point $q \in {}_pV$, the differential $d\varphi_p$ of the function φ is equal to zero on the subspace $({}_pV)_q$ of M_q;

e) the point $p \in N$ is a critical point of the function φ if and only if this point is a critical point of the function $\varphi|_N$ (on the manifold N).

Moreover, in view of property a)

f) for any critical point $p \in N$ of the function φ, the maximal subspace of M_p on which the Hessian $H_p\varphi$ is identically zero coincides with the space $({}_pV)$ tangent to the fibre ${}_pV$ at the point p.

Finally, since, in view of property c) the function φ on the manifold M is bounded, we can, without loss of generality, assume (multiplying, if necessary, the function φ by some constant) that

g) for any point $p \in M$ the following inequality holds

$$|\varphi(p)| \leq \frac{\varepsilon}{2};$$

Furthermore, by similar considerations, we can assume that

h) on the entire manifold M the following inequality holds

$$|\operatorname{grad} \varphi| \leq \frac{K}{2}.$$

Let us now consider the function

$$g = f + \varphi.$$

Let p be a point of the manifold M such that

$$g(p) \leq c + \varepsilon.$$

If $p \notin [c - \frac{\varepsilon}{2} \leq f \leq c + \frac{\varepsilon}{2}]$, then $g(p) = f(p)$ and therefore

$$f(p) \leq c + \varepsilon.$$

If $c - \frac{\varepsilon}{2} \leq f \leq c + \frac{\varepsilon}{2}$, then even more so

$$f(p) \leq c + \varepsilon.$$

Conversely, let

$$f(p) \leq c + \varepsilon.$$

If $c - \frac{\varepsilon}{2} \leq f \leq c + \frac{\varepsilon}{2}$, then in view of property g)

$$g(p) \leq c + \varepsilon.$$

If $p \notin [c - \frac{\varepsilon}{2} \leq f \leq c + \frac{\varepsilon}{2}$, then in view of property c) $\varphi(p) = 0$ and therefore again

$$g(p) \leq c + \varepsilon.$$

Thus,

$$[g \leq c + \varepsilon] = [f \leq c + \varepsilon].$$

Similarly, it is proved that

$$[g \leq c - \varepsilon] = [f \leq c - \varepsilon].$$

By construction, for any point $q \in V \setminus N$, the differential df_q of df is no-nzero on the subspace $({}_pV)_q$ of M_q, where ${}_pV$ is the fibre of the neighbourhood V containing the point q. By property c), it immediately follows that the differential dg_q of g is also non-zero on the subspace $({}_pV)_q$. Therefore, it is non-zero on the entire space M_q. Thus, no point $q \in V \setminus N$ is a critical point of g.

Similarly, by the inequality h), no point $q \in [c - \varepsilon \leq f \leq c + \varepsilon] \setminus V$ is also a critical point of the function g. Consequently, all critical points of the function g belonging to the set $[c-\varepsilon \leq f \leq c+\varepsilon]$ are located on the manifold $N \subset [f = c]$.

On the other hand, since $df_p = 0$ for any point $p \in N$, this point is a critical point of g if and only if it is a critical point of φ. Moreover, at any such point, the following equality holds

$$H_p g = H_p f + H_p \varphi.$$

Since the maximal subspaces N_p and $({}_pV_p)$ of M_p on which the quadratic forms $H_p f$ and $H_p \varphi$ are identically zero are mutually complementary subspaces of M_p, it follows immediately that the quadratic form $H_p g$ is non-degenerate and that its index is equal to the sum of the indices $h_f(p)$ and $h_\varphi(p)$ of $H_p f$ and $H_p \varphi$ and is no less than the number $h \leq h_f(p)$.

This proves that the function g satisfies all the conditions listed above.

Remark 6.57. The theorem just proved admits of a refinement that is completely analogous to Morse's theorem and transforms into this theorem when the critical value c is non-degenerate. However, in practice, this refinement is essentially useless. Therefore, we will not present it here.

Bott's theorem applies to the problem of describing the homotopy type of smooth manifolds in exactly the same way as Morse's theorem (see §6.4). For example, we immediately obtain from it that

Proposition 6.58. *every smooth function f of compact character bounded below on a smooth separable manifold M, having only weakly non-degenerate critical points, defines a homotopy equivalence of M with some space obtained from the set N of all minimum points of f (which, by assumption, is a submanifold - generally, not connected - of M) by successively gluing together a countable number of closed balls whose dimensions are no less than the smallest of the indices of the critical points of f that are not minimum points.*

Appendix

Morse inequalities

Strictly speaking, the fundamental theorem of §6.3 was only implicitly contained in Morse's research. In fact, Morse described the relationship between the topology of a smooth manifold M and the critical points of a smooth function f on M using so-called Morse inequalities. Although these inequalities convey significantly less information than the theorem of §6.3, we will nevertheless present them here, since it is often sufficient to use only these inequalities without resorting to the more complex theorem of §6.3. Furthermore, these inequalities have a certain historical interest.

Morse inequalities involve the so-called Betti numbers of the manifold M (the ranks of its homology groups). Referring the definition and proof of the basic properties of these numbers to any algebraic topology textbook (see the historical and literary commentary at the end of the book), we will only recall here that

1) The definition of Betti numbers involves a certain field, called the *coefficient field*; all results of this section remain valid regardless of the choice of this field;

2) Betti numbers are defined for certain pairs (X, A) consisting of a topological space X and its subspace A and are non-negative integers; we will denote them by the symbols
$$b_n(X, A),$$
where $n = 0, 1, \dots$; for $A = \varnothing$, we will write $b_n(X)$ instead of $b_n(X; \varnothing)$;

3) if the number $b_n(X, A)$ is defined for the pair (X, A), then it is also defined for any pair (X', A') that is homototically equivalent to the pair (X, A), and
$$b_n(X', A') = b_n(X, A);$$

4) the number $b_n(X, A)$ is a priori defined if the pair (X, A) is cellular and the number of n-dimensional cells of the decomposition X that do not belong to the sub-decomposition A is finite;

5) if the number $b_n(X, A)$ is defined for a cellular pair (X, A), then it is also defined for any pair of the form $(X^m \cup A, A)$, where X_m is the n-dimensional

skeleton of the decomposition X, and

$$b_n(X^m \cup A, A) = b_n(X, A),$$

if $n < m$;

6) if the space X is obtained from the space A by gluing k closed m-dimensional balls (via some maps of their boundaries), then

$$b_n(X, A) = \begin{cases} k, & \text{if} \quad n = m, \\ 0, & \text{if} \quad n \neq m. \end{cases}$$

We'll need another property of Betti numbers, usually not specifically mentioned in textbooks.

Let $s(X, A)$ be an arbitrary numerical invariant defined for some pairs (X, A) of topological spaces. We will call this invariant *subadditive* if, for any three nested spaces $X \supset A \supset B$, the relation

$$s(X, B) \leq s(X, A) + s(A, B) \tag{6.59}$$

(provided, of course, that all numbers involved in this relation are defined). In the case where equality always holds in relation (6.59, we will call the invariant $s(X, A)$ *additive*.

Using the so-called "exact homotopy sequences of pairs", one can easily prove that

7) for any $n \geq 0$ the number $b_n(X, A)$ is subadditive

and, moreover, that

8) for any $n \geq 0$ the following number is subadditive

$$c_n(X, A) = b_n(X, A) - b_{n-1}(X, A) + \cdots + (-1)^n b_0(X, A).$$

For every pair (X, A) with the property that $b_n(X, A) = 0$ for all sufficiently large n, there is a number

$$\chi(X, A) = b_0(X, A) - b_1(X, A) + \cdots + (-1)^n b_n(X, A) + \cdots,$$

called the *Euler characteristic* of the pair (X, A). Similarly to property 8), it can be easily proved that

9) the Euler characteristic $\chi(X, A)$ is additive.

Let K be an arbitrary cellular decomposition such that for any $n \geq 0$, the number $a_n(K)$ of its n-dimensional cells is finite. Then (see Property 4) for any $n \geq 0$ and $m \geq 0$, the number $b_n(K^m, k^{m-1})$ and also the number $b_n(K)$ are defined. Moreover, according to Property 6), for any $n \geq 0$ and $m \geq 0$, the equality holds

$$b_n(K^m, k^{m-1}) = \begin{cases} a_n(K), & \text{if} \quad m = n, \\ 0, & \text{if} \quad m \neq n. \end{cases}$$

Summing these equalities over m, we obtain that for any $n \geq 0$, the equality

$$\sum_{m=0}^{n+1} b_n(K^m, k^{m-1}) = a_n(K). \tag{6.60}$$

Applying the subadditivity relation to this equality the required number of times and taking into account that $b_n(K^{n+1}, K^{-1}) = b_n(K)$, we immediately obtain that

$$b_n(K) \leq a_n(K). \tag{6.61}$$

Similarly, taking the alternating sum of relations (6.60) and using the additivity property of the Euler characteristic, we obtain (assuming that $\dim K = m$, i.e., that $a_n(K) = 0$ for $n > m$) the relation

$$b_0(K) - b_1(K) + \cdots + (-1)^m b_m(K) = a_0(K) - a_1(K) + \cdots + (-1)^m a_m(K) \tag{6.62}$$

More precise relations can be obtained using the subadditivity of the invariants $c_n(X, A)$. Indeed, taking the alternating sum of the first n relations in (6.60), we immediately obtain that

Proposition 6.63. *for any $n \geq$ the following inequalities hold:*

$$b_n(K) - b_{n-1}(K) + \cdots + (-1)^n b_0(K) \leq a_n(K) - a_{n-1}(K) + \cdots + (-1)^n a_0(K). \tag{6.64}$$

It is clear that inequalities (6.61) are immediate consequences of inequalities (6.64). (To obtain inequality (6.61) for $n = k$, it suffices to add together inequalities (6.64) corresponding to the values $n = k$ and $n = k - 1$.) Equality (6.62) also follows from inequalities (6.64) (it suffices to compare inequalities (6.64) corresponding to the values $n = m$ and $n = m + 1$, and to use the fact that $a_{m+1}(K) = b_{m+1}(K) = 0$).

Remark 6.65. Inequalities (6.61) follow directly from the definition of Betti numbers as the ranks of homology groups. Equality (6.62) is known as the "topological invariance theorem for the Euler characteristic," and its proof (essentially identical to the above) can be found in any algebraic topology textbook. In contrast, inequalities (6.64) are much more profound and are usually not presented in topology textbooks.

Let M be an arbitrary smooth separable manifold, and f an arbitrary smooth, lower-bounded function of compact character on M with only nondegenerate critical points. By §6.4, this function defines a homotopy equivalence of M with some cellular decomposition K, whose cells are in one-to-one correspondence with the critical points of f, where the dimension of each cell is equal to the index of the corresponding critical point. Suppose that the number of critical points of f is finite (this condition is necessarily satisfied, for example, if M is compact). Then the decomposition K has the above property, and therefore all the results obtained above are applicable to it. Given that the numbers $b_n(K)$ are equal to the numbers $b_n(M)$ (by Property 3) of the Betti numbers) and that for any $n = 0, 1, \ldots, m$, where $m = \dim M$, the number $a_n(K)$ is equal to the number $a_n(f)$ of critical points of the function f with index n, we immediately deduce from this that

Proposition 6.66. *for any smooth separable manifold M, any smooth function f bounded below of compact character on M, having only a finite number of critical points, all of which are non-degenerate, and any $n \geq 0$, the following equalities hold*

$$b_n(M) - b_{n-1}(M) + \cdots + (-1)^n b_0(M) \leq a_n(f) - a_{n-1}(f) + \cdots + (-1)^n a_0(f).$$

where $b_n(M)$ are the Betti numbers of M, and $a_n(f)$ is the number of critical points of f with index n.

In particular,

Proposition 6.67. *for any $n \geq 0$, the following inequalities hold*

$$b_n(M) \leq a_n(f).$$

Furthermore, the following equality holds

$$b_0(M) - b_1(M) + \cdots + (-1)^m b_m(M) = a_0(f) - a_1(f) + \cdots + (-1)^m a_m(f).$$

These relations sometimes allow us to quite accurately estimate the Betti numbers of a given manifold M (or, conversely, to obtain information about the number of critical points of the function f from knowledge of its Betti numbers). For example, it easily follows from them that

Proposition 6.68. *if $a_{n-1}(f) = a_{n+1}(f) = 0$, then $b_{n-1}(M) = b_{n+1}(N) = 0$ and $b_n(M) = a_n(f)$.*

Remark 6.69. We emphasise that all these results are valid only for functions of a compact nature that are bounded below.

Chapter 7

Elements of Riemannian Geometry

This chapter contains a strictly selected minimum of information from the theory of Riemannian spaces (and spaces with affine connections) necessary for the subsequent chapters. Therefore, this chapter does not contain any specific theory. However, the entire presentation is conducted with complete rigor, with particular attention to details that are usually overlooked in classical expositions.

The first six sections are devoted to spaces with affine connections.

§7.1 introduces the basic concepts of the theory of spaces with affine connections.

§7.2 - §7.4 construct the formal apparatus of covariant (absolute) differentiation.

§7.5, after a few simple remarks on geodesics, proves the existence of neighbourhoods normal with respect to each of their points (Whitehead's theorem).

Section 6 derives the well-known Cartan equations for differential connection forms.

The remaining sections (sections 7.7 - 7.12) consider proper Riemannian spaces.

§7.7 proves the existence and uniqueness of a Riemannian connection.

§7.8 constructs and studies the Riemannian curvature tensor. In this same section, the metric tensor is expressed in terms of differential connection forms of the corresponding Riemannian connection.

In §7.9, the intrinsic metric of a Riemannian space is introduced and its simplest properties are considered.

In §7.10, the existence of normal convex neighbourhoods is proved. A convergence lemma, needed in the next section, is also proved here.

In §7.11, it is proved that in a complete Riemannian space, any two points can be connected by a shortest arc (the Hopf-Rinow theorem).

Finally, in §7.12, it is proved that a Riemannian space is complete if and only if it is metrically complete.

7.1 Affine connections

7.2 Covariant differentiation of vector fields

7.3 Parallel translation along a curve

7.4 Covariant differentiation of tensor fields

7.5 Geodesic. Normal neighbourhoods

7.6 Differential forms of connection

7.7 Riemannian connections

7.8 Riemannian curvature tensor and connection forms

7.9 Curve lengths and intrinsic metrics

7.10 Geodesic. Normal neighbourhoods

7.11 Complete Riemannian spaces

Remark 7.1. It is impossible to assert that the function $\rho^2(p,q)$ is smooth at all points $(p,q) \in M \times M$: this is not true even for such simple manifolds as spheres.

7.12 Conditions for the completeness of Riemannian spaces

Quite unexpectedly, complete Riemannian spaces, in the class of all Riemannian spaces, can be characterised by extremely simple, essentially purely topological properties. Specifically, it turns out that

Proposition 7.2.

From the proven statement, in particular, it follows that any compact connected Riemannian space is complete.

Chapter 8

Variational theory of geodesics

This chapter presents the fundamental analytical-geometric Morse index theorems, which reduce the calculation of the number of conjugate points located on an arbitrary interval to the calculation of the negative index of inertia of a certain quadratic form.

After the introductory §8.1 - 8.2, §8.3 - 8.4 present the basic concepts of a Jacobi field and a conjugate point and prove their simplest properties.

In sections §8.5 - 8.7, it is established that Jacobi fields are nothing more than fields that minimize a certain quadratic functional.

In sections §8.8 - 8.10, the fundamental Morse form is constructed.

In sections §8.11 - 8.12, the fundamental Morse theorems are proved.

In section §8.13, Morse's theorems are reformulated in a more convenient form.

In the appendix, the main results of the chapter are extended to the case of a problem with a moving end.

8.1 Geodesics as lines of fixed length

8.2 The second variation of the length of the geodesic arc

8.3 Jacobian variations and Jacobian fields

8.4 Conjugate points

8.5 Piecewise smooth and discontinuous vectors

8.6 Minimal vector fields

8.7 Existence of minimal fields

8.8 Jacobi's broken fields

8.9 Isomorphism Theorem

8.10 Quadratic Morse form

8.11 Calculating the index of a point using Morse form

8.12 Calculating the interval index using Morse form

8.13 Bott's Quadratic Form: A Final Statement of Index Theorems

Appendix

Focal Points

8.A The second quadratic form of a submanifold

8.B Focal points

8.C Calculating the index λ_N

$$I_a^b(X(t)) > 0. \tag{8.1}$$

8.D Proof of inequality (8.1)

Chapter 9

Path Space Exploration. Applications

This chapter presents the fundamental theorems of Morse theory concerning the structure of path spaces on smooth manifolds. They are obtained by a fairly simple synthesis of the results of the previous chapters. The last two sections of the chapter present the simplest applications of Morse theory to the topology and geometry of Riemannian spaces.

§9.1, which is purely topological in nature, studies the space of all paths of a given path-connected space X. The main result is that the homotopy type of the space of paths connecting two given points of X does not depend on the choice of these points.

§9.2 compares the space of piecewise smooth curves with the space of paths for an arbitrary Riemannian space. The main result is that the space of piecewise smooth curves connecting two given points of a Riemannian space is homotopy equivalent to the corresponding space of paths.

Section 3, based on the results of Chapter 8 proves the fundamental reduction theorem, which reduces the study of the space of curves to the study of certain finite-dimensional manifolds.

In Section 5, the general results of Section 4 are applied to the study of loop spaces on spheres. As an application, Freudenthal's theorem on homotopy groups of spheres is proved.

In §9.6, Morse inequalities are proved that relate the numbers of geodesics of a given index to the Betti numbers of the loop space. In particular, Morse's theorem on geodesics on manifolds homotopically equivalent to a sphere is proved. In this same section, as a fundamentally different application of Morse theory, Cartan's theorem on simply connected complete Riemannian spaces of non-positive curvature is proved.

9.1 Path spaces

9.2 Spaces of piecewise smooth curves

9.3 Reduction theorem

9.4 Homotopy type of spaces $\widehat{\Omega}(p, q)$

9.5 Applications to Topology. Freudenthal's Theorem

Remark 9.1. It can be shown that the generalized Freudenthal theorem is valid even without the above-mentioned conditions (ensuring weak nondegeneracy of geodesics). However, such a strengthening of this theorem is of little interest, since these conditions are satisfied in all known applications of it to date.

9.6 Applications to the Geometry of Riemannian Spaces. Theorems of Morse and Cartan

Remark 9.2. Using powerful methods of modern algebraic topology, Serre proved that for any manifold M not contractible to a point, the Betti numbers b_n of its loop space are nonzero for an infinite number of values of n. Consequently, Morse's theorem holds not only for spaces homotopy equivalent to a sphere, but also for any Riemannian complete space not contractible to a point. A proof of Serre's theorem is beyond the scope of this book.

This remarkable theorem is due to E. Cartan. It is one of many theorems linking the topological properties of Riemannian spaces with the properties of their curvature tensor. The reader can find other theorems of this type in the relevant literature.

History and literary commentary

Bibliography

[1] Aleksandroff P., *Über die Metrisation der inkleinen kompakten topologlschen Räume*, Math. Ann. **92** (1924), 294-301.

[2] H. Seifert and V. Threlfall, *A Textbook of Topology*, Academic Press, New York, 1980.

[3] R. H. Crowell and R. H. Fox, *Introduction to Knot Theory*, Springer, New York, 1963(1977).

[4] W. S. Massey, *Algebraic Topology: An Introduction*, Springer, New York, 1977.

[5] J. W. Minor, *Topology from the Differentiable Viewpoint*, Princeton University Press, Princeton, 1997.

[6] A. Fomenko and A. S. Mishchenko, *A course of differential geometry and topology*, Mir Publishers, Moscow, 1988.

[7] L. S. Pontrjagin, *Smooth manifolds and their applications in homotopy theory, in Topological Library*, World Scientific, Singapore, 2007.

[8] Postnikov M. M., *Introduction to Morse theory*, Nauka, Moscow.

[9] M. M. Postnikov, *Variational theory of geodesics*, Nauka, Moscow, 1965.

[10] D. V. Fuks and V. A. Rokhlin, *Beginner's Course in Topology: Geometric Chapters*, Springer, New York, 1984.

[11] R. M. Switzer, *Algebraic topology. Homotopy and homology*, Springer, Berlin, 1975.

[12] S. Eilenberg and N. Steenrod, *Foundations of Algebraic Topology*, Princeton University Press, Princeton, 1952.

[13] E. H. Spanier, *Algebraic Topology*, Springer (was McGraw-Hill), New York, 1966.

[14] P. J. Hilton and S. Wylie, *Homology Theory: An Introduction to Algebraic Topology*, Cambridge University Press, Cambridge, 1960.

[15] M. Hirsch, *Differential Topology*, Springer, New York, 1976.

[16] S-T. Hu, *Homotopy theory*, Academic Press, Cambridge, Massachusetts, 1959.

[17] A. Fomenko and D. Fuchs, *Homotopical Topology*, Springer, New York, 2016.

[18] T. tom Dieck, K. H. Kammps, and D. Puppe, *Homotopietheorie. - Lect. Notes Math., 157*, Springer, Berlin, 1970.

[19] G. W. Whitehead, *Elements of homotopy theory*, Springer, Berlin, 1978.