Elements of Combinatorial and Differential Topology

V. V. Prasolov

Graduate Studies in Mathematics

Volume 74

American Mathematical Society

Elements of Combinatorial and Differential Topology

V. V. Prasolov

Graduate Studies in Mathematics

Volume 74

American Mathematical Society
Providence, Rhode Island

В. В. Прасолов

Элементы комбинаторной и дифференциальной топологии

МЦНМО, Москва, 2004

Translated from the Russian by Olga Sipacheva

2000 *Mathematics Subject Classification.* Primary 57–01;
Secondary 57Mxx, 57Rxx.

For additional information and updates on this book, visit
www.ams.org/bookpages/gsm-74

Library of Congress Cataloging-in-Publication Data

Prasolov, V. V. (Viktor Vasil'evich)
[Elementy kombinatornoi i differentsial'noi topologii. English]
Elements of combinatorial and differential topology / V. V. Prasolov ; [translated from the Russian by Olga Sipacheva].
p. cm. — (Graduate studies in mathematics, ISSN 1065-7339 ; v. 74)
Includes bibliographical references and index.
ISBN 0-8218-3809-1 (acid-free paper)
1. Combinatorial topology. 2. Differential topology. 3. Low-dimensional topology. 4. Topological manifolds. I. Title. II. Series.

QA612.P73 2006
514'.22—dc22 2006042681

♾ The paper used in this book is acid-free and falls within the guidelines
established to ensure permanence and durability.
Visit the AMS home page at `http://www.ams.org/`

10 9 8 7 6 5 4 3 2 1 11 10 09 08 07 06

Contents

Preface

Modern topology uses many different methods. In this book, we largely investigate the methods of combinatorial topology and the methods of differential topology; the former reduce studying topological spaces to investigation of their partitions into elementary sets, such as simplices, or covers by some simple sets, while the latter deal with smooth manifolds and smooth maps. Many topological problems can be solved by using any of the two approaches, combinatorial or differential; in such cases, we discuss both of them.

Topology has its historical origins in the work of Riemann; Riemann's investigation was continued by Betti and Poincaré. While studying multivalued analytic functions of a complex variable, Riemann realized that, rather than in the plane, multivalued functions should be considered on two-dimensional surfaces on which they are single-valued. In these considerations, two-dimensional surfaces arise by themselves and are defined intrinsically, independently of their particular embeddings in $\mathbb{R}^3$; they are obtained by gluing together overlapping plane domains. Then, Riemann introduced the notion of what is known as a (multidimensional) manifold (in the German literature, Riemann's term *Mannigfaltigkeit* is used). A manifold of dimension n, or n-manifold, is obtained by gluing together overlapping domains of the space $\mathbb{R}^n$. Later, it was recognized that to describe continuous maps of manifolds, it suffices to know only the structure of the open subsets of these manifolds. This was one of the most important reasons for introducing the notion of topological space; this is a set endowed with a topology, that is, a system of subsets (called open sets) with certain properties.

Chapter 1 considers the simplest topological objects, graphs (one-dimensional complexes). First, we discuss questions which border on geometry, such as planarity, the Euler formula, and Steinitz' theorem. Then, we consider fundamental groups and coverings, whose basic properties are well seen in graphs. This chapter is concluded with a detailed discussion of the polynomial invariants of graphs; there has been much interest in them recently, after the discovery of their relationship with knot invariants.

Chapter 2 is concerned with another fairly simple topological object, Euclidean space with standard topology. Subsets of Euclidean space may have very complicated topological structure; for this reason, only a few basic statements about the topology of Euclidean space and its subsets are included. One of the fundamental problems in topology is the classification of continuous maps between topological spaces (on the spaces certain constraints may be imposed; the classification is up to some equivalence). The simplest classifications of this kind are related to curves in the plane, i.e., maps of S^1 to $\mathbb{R}^2$. First, we prove the Jordan theorem and the Whitney Graustein classification theorem for smooth closed curves up to regular homotopy. Then, we prove the Brouwer fixed point theorem and Sperner's lemma by several different methods (in addition to the standard statement of Sperner's lemma, we give its refined version, which takes into account the orientations of simplices). We also prove the Kakutani fixed point theorem, which generalizes the theorem of Brouwer. The chapter is concluded by the Tietze theorem on extension of continuous maps, which is derived from Urysohn's lemma, and two theorems of Lebesgue, the open cover theorem, which is used in the rigorous proofs of many theorems from homotopy and homology theories, and the closed cover theorem, on which the definition of topological dimension is based.

Chapter 3 begins with elements of general topology; it gives the minimal necessary information constantly used in algebraic topology. We consider three properties (Hausdorffness, normality, and paracompactness) which substantially facilitate the study of topological spaces. Then, we consider two classes of topological spaces that are most important in algebraic topology (namely, simplicial complexes and CW-complexes), describe techniques for dealing with them (cellular and simplicial approximation), and prove that these spaces have the three properties mentioned above. We also introduce the notion of degree for maps of pseudomanifolds and apply it to prove the Borsuk–Ulam theorem, from which we derive many corollaries. The chapter is concluded with a description of some constructions of topological spaces, including joins, deleted joins, and symmetric products. We apply deleted joins to prove that certain n-dimensional simplicial complexes cannot be embedded in $\mathbb{R}^{2n}$.

Chapter 4 covers very diverse topics, such as two-dimensional surfaces, coverings, local homeomorphisms, graphs on surfaces (including genera of graphs and graph coloring), bundles, and homotopy groups.

Chapter 5 turns to differential topology. We consider smooth manifolds and the application of smooth maps to topology. First, we introduce some basic tools (namely, smooth partitions of unity and Sard's theorem) and consider an example, the Grassmann manifolds, which plays an important role everywhere in topology. Then, we discuss notions related to tangent spaces, namely, vector fields and differential forms. After this, we prove existence theorems for embeddings and immersions (including closed embeddings of noncompact manifolds), which play an important role in the study of smooth manifolds. Moreover, we prove that a closed nonorientable n-manifold cannot be embedded in $\mathbb{R}^{n+1}$ and determine what two-dimensional surfaces can be embedded in $\mathbb{R}P^3$. Further, we introduce a homotopy invariant, the degree of a smooth map, and apply it to define the index of a singular point of a vector field. We prove the Hopf theorem, which gives a homotopy classification of maps $M^n \to S^n$. We also describe a construction of Pontryagin which interprets $\pi_{n+k}(S^n)$ as the set of classes of cobordant framed k-manifolds in $\mathbb{R}^{n+k}$. We conclude this chapter with Morse theory, which relates the topological structure of a manifold to local properties of singular points of a nondegenerate function on this manifold. We give explicit examples of Morse functions on some manifolds, including Grassmann manifolds.

Chapter 6 is devoted to explicit calculations of fundamental groups for some spaces and to applications of fundamental groups. First, we prove a theorem about generators and relations determining the fundamental group of a CW-complex and give some applications of this theorem. Sometimes, it is more convenient to calculate fundamental groups by using exact sequences of bundles. Such is the case for, e.g., the fundamental group of $SO(n)$. In many situations, the van Kampen theorem about the structure of the fundamental group of a union of two open sets is helpful. For example, it can be used to calculate the fundamental group of a knot complement. At the end of the chapter, we give another theorem of van Kampen, which gives a method for calculating the fundamental group of the complement of an algebraic curve in $\mathbb{C}P^2$. The corresponding calculations for particular curves are fairly complicated; plenty of interesting results have been obtained, but many things are not yet fully understood.

One of the main purposes of this book is to advance in the study of the properties of topological spaces (especially manifolds) as far as possible without employing complicated techniques. This distinguishes it from the majority of topology books.

The book is intended for readers familiar with the basic notions of geometry, linear algebra, and analysis. In particular, some knowledge of open, closed, and compact sets in Euclidean space is assumed.

The book contains many problems, which the reader is invited to think about. They are divided into three groups: (1) *exercises*; solving them should not cause any difficulties, so their solutions are not included; (2) *problems*; they are not so easy, and the solutions to most of them are given at the end of the book; (3) *challenging problems* (marked with an asterisk); each of these problems is the content of a whole scientific paper. They are formulated as problems not to overburden the main text of the book. The solutions to most of these problems are also given at the end of the book. The problems are based on the first- and second-year graduate topology courses taught by the author at the Independent University of Moscow in 2002.

This work was financially supported by the Russian Foundation for Basic Research (project no. 05–01–01012a).

Notation

- $X \approx Y$ means that the topological spaces X and Y are homeomorphic;
- $X \sim Y$ means that the topological spaces X and Y are homotopy equivalent;
- $f \simeq g$ means that the map f is homotopic to g;
- $|A|$ denotes the cardinality of the set A;
- $\operatorname{int} A$ denotes the interior of A;
- $\bar{A}$ denotes the closure of A;
- ∂A denotes the boundary of A;
- id_A denotes the identity map on A;
- K_n denotes the complete graph on n vertices;
- $K_{n,m}$, see p. 7;
- D^n denotes the n-disk (or n-ball);
- S^n denotes the n-sphere;
- Δ^n denotes the n-simplex;
- I^n denotes the n-cube;
- P^2 denotes the projective plane;
- T^2 denotes the two-dimensional torus;
- $S^2 \# nP^2$ and nP^2 denote the connected sum of n projective planes;
- $S^2 \# nT^2$ and nT^2 denote the connected sum of n 2-tori (the sphere with n handles);
- K^2 denotes the Klein bottle;

- $\|x - y\|$ denotes the distance between points x, $y \in \mathbb{R}^n$;
- $\|v\|$ denotes the length of the vector $v \in \mathbb{R}^n$;
- $d(x, y)$ denotes the distance between points x and y;
- inf denotes the greatest lower bound;
- $X \sqcup Y$ denotes the disjoint union of X and Y;
- $\operatorname{supp} f = \overline{\{x : f(x) \neq 0\}}$ denotes the support of the function f;
- $X * Y$ denotes the join of the spaces X and Y;
- $\mathrm{SP}^n(X)$ denotes the n-fold symmetric product of X;
- $f\colon (X, Y) \to (X_1, Y_1)$ denotes the map of pairs which takes $Y \subset X$ to $Y_1 \subset X_1$;
- $\pi_1(X, x_0)$ denotes the fundamental group of the space X with base point $x_0 \in X$;
- $\pi_n(X, x_0)$ denotes the n-dimensional homotopy group of the space X with base point $x_0 \in X$;
- $\deg f$ denotes the degree of a map f;
- $\operatorname{rank} f(x)$ denotes the rank of f at the point x;
- $G(n, k)$ denotes the Grassmann manifold;
- $\mathrm{GL}_k(\mathbb{R})$ denotes the group of $k \times k$ nonsingular matrices with real entries;
- $\mathrm{U}(n)$ denotes the group of unitary matrices of order n;
- $\mathrm{SU}(n)$ denotes the group of unitary matrices of order n with determinant 1;
- $\mathrm{O}(n)$ denotes the group of orthogonal matrices of order n;
- $\mathrm{SO}(n)$ denotes the group of orthogonal matrices of order n with determinant 1;
- T_xM^n denotes the tangent space at the point $x \in M^n$;
- TM^n denotes the tangent bundle;
- $\Omega^k_{\mathrm{fr}}(n+k)$ denotes the set of classes of framed cobordant k-manifolds in $\mathbb{R}^{n+k}$.

Basic Definitions

To start the book, we only need some basic notions of topology. We present them below.

A *topological space* is a set X with a system τ of distinguished subsets having the following properties:

(i) the empty set and the entire set X belong to τ;

(ii) the intersection of any finite set of elements of τ belongs to τ;

(iii) the union of an arbitrary family of elements of τ belongs to τ.

The sets belonging to τ are said to be *open*. A *neighborhood* of a point $x \in X$ is an arbitrary open set containing x. The sets with open complements are *closed*.

Let A be a subset in a topological space. Its *closure* $\overline{A}$ is defined as the minimal closed set containing A, and its *interior* $\operatorname{int} A$ is the maximal open set contained in A. The closure of A is the intersection of all closed sets containing A, and the interior of A is the union of all open sets contained in A.

The most important example of a topological space is the Euclidean space $\mathbb{R}^n$. The open sets in $\mathbb{R}^n$ are the balls $D^n_{a,\varepsilon} = \{x \in \mathbb{R}^n : \|x - a\| < \varepsilon\}$ and all their unions.

A family $\tau' \subset \tau$ is called a *base* for the topology τ if any element of τ is a union of elements of τ'.

Exercise 1. Prove that a family $\tau' \subset \tau$ is a base for the topology τ if and only if, for any point x and any neighborhood U of x, there exists a $V \in \tau'$ such that $x \in V \subset U$.

Exercise 2. Prove that a family of sets τ' is a base for some topology if and only if, for any $U, V \in \tau'$ and any $x \in U \cap V$, there exists a $W \in \tau'$ such that $x \in W \subset U \cap V$.

A topological space X is called *second countable* if it has a countable base. For example, the open balls $D^n_{a,\varepsilon}$, where the numbers ε and all coordinates of the points a are rational, form a countable base of the space $\mathbb{R}^n$.

Problem 1. Let X be a second countable topological space. Prove that any cover of X by open sets U_α has a countable subcover.

If X is a topological space and Y is a subset of X, then Y can be endowed with the *induced topology*, which consists of the intersections of Y with open subsets of X. This turns the sphere $S^n = \{x \in \mathbb{R}^{n+1} : \|x\| = 1\}$ into a topological space.

A map of topological spaces is said to be *continuous* if the preimage of any open set is open, or, equivalently, the preimage of any closed set is closed.

In proving the continuity of a map f, the following continuity criterion is often convenient: *A map $f: X \to Y$ is continuous if and only if, for any point $x \in X$ and any neighborhood U of $f(x)$, there exists a neighborhood $V(x)$ of x such that its image is entirely contained in U.* Indeed, if this condition holds, then the preimage $f^{-1}(U)$ of any open set U can be represented as the union $\bigcup_{x\in f^{-1}(U)} V(x)$ of open sets and is therefore open. The converse assertion is obvious: for $V(x)$ we can take $f^{-1}(U)$.

Exercise 3. Prove that a map $f: \mathbb{R}^n \to \mathbb{R}^m$ is continuous if and only if, for any $x \in \mathbb{R}^n$ and any $\varepsilon > 0$, there exists a $\delta > 0$ such that $\|f(x) - f(a)\| < \varepsilon$ whenever $\|x - a\| < \delta$.

The following gluing theorem for continuous maps is used fairly often in topology.

Theorem 0.1. *Suppose that $X = X_1 \cup \dots \cup X_n$ and the sets $X_1, \dots, X_n$ are closed. A map $f: X \to Y$ is continuous if and only if the restrictions $f_i = f \upharpoonright X_i$ are continuous.*

Proof. Clearly, if a map f is continuous, then all of the maps f_i are continuous as well. Suppose that the maps f_i are continuous and $C \subset Y$ is an arbitrary closed set. Then, for each i, the set $C_i = f_i^{-1}(C) = f^{-1}(C) \cap X_i$ is closed in X_i, i.e., there exists a closed subset C_i' of X such that $C_i = C_i' \cap X_i$. Both sets C_i' and X_i are closed in X; therefore, C_i is closed in X also. Hence the set $f^{-1}(C) = C_1 \cup \dots \cup C_n$ is closed in X. □

A map $f: X \to Y$ is called a *homeomorphism* if it is one-to-one and both of the maps f and f^{-1} are continuous. In this case, we say that the topological spaces X and Y are *homeomorphic* and write $X \approx Y$.

Exercise 4. Prove that the spaces $\mathbb{R}^n$ and $S^n \setminus \{x_0\}$ are homeomorphic.

Problem 2. Prove that $S^{n+m-1} \setminus S^{n-1} \approx \mathbb{R}^n \times S^{m-1}$. (We assume that the sphere S^{n-1} is standardly embedded in S^{n+m-1}.)

A topological space X is said to be *discrete* if all of its subsets are open (or, equivalently, if all of its subsets are closed). The topology of a discrete topological space is called the *discrete topology.* If X is a discrete topological space and Y is an arbitrary topological space, then any map $f: X \to Y$ is continuous.

A topological space X is called *connected* if it contains no proper subsets which are both open and closed. In other words, if a set $A \subset X$ is both open and closed, then either $A = \varnothing$ or $A = X$.

Exercise 5. Prove that the space $\mathbb{R}^n$ is connected.

Exercise 6. Prove that if X is a connected topological space and Y is a discrete topological space, then any continuous map $f: X \to Y$ is constant, i.e., $f(X)$ consists of one point.

Problem 3. Prove that if A and B are connected subsets of a topological space X that are not disjoint (i.e., have a nonempty intersection), then $A \cup B$ is connected.

Problem 3 shows that the following relation on the set of points of a topological space X is an equivalence: two points are considered equivalent if there exists a connected set containing both of them. The equivalence class of a point $x \in X$ is the maximal connected set containing x. It is called a *connected component*.

A *metric space* is a set X such that, for any two points $x, y \in X$, a number $d(x, y) \geq 0$ is defined and satisfies the following conditions:

(1) $d(x, y) = d(y, x)$;

(2) $d(x, y) + d(y, z) \geq d(x, z)$ (the *triangle inequality*);

(3) $d(x, y) = 0$ if and only if $x = y$.

The number $d(x, y)$ is called the *distance* between the points x and y.

In any metric space X, the open balls $D^n_{a,\varepsilon} = \{x \in X : d(x, a) < \varepsilon\}$ form a base for some topology. This topology is said to be *induced by the metric d*. If X is a topological space and its topology is induced by some metric, then X is *metrizable.*

A topological space is said to be *compact* if any cover of this space by open sets has a finite subcover.

Exercise 7. Prove that the sphere S^n is compact and the space $\mathbb{R}^n$ is noncompact.

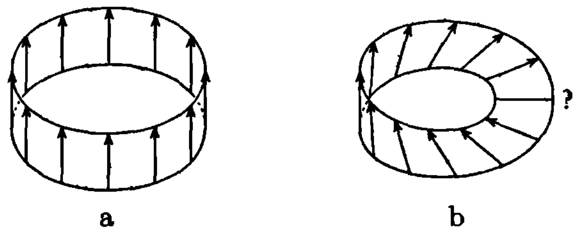

Figure 1. The cylinder and the Möbius band

Exercise 8. Prove that any continuous image of a compact space is compact.

Problem 4. Let K be a compact metric space with metric ρ. Suppose that $f\colon K \to K$ is a continuous map such that $\rho(f(x), f(y)) < \rho(x,y)$ for any distinct $x, y \in K$. Prove that f has a fixed point.

On the Cartesian product $X \times Y$ of topological spaces X and Y, the *product topology* is defined. The open subsets of $X \times Y$ in this topology are the Cartesian products of open subsets of X and Y and their unions.

The product topology arises from the natural requirement that the projections $p_X(x,y) = x$ and $p_Y(x,y) = y$ be continuous. Indeed, for these maps to be continuous, the sets $U \times Y$ and $X \times V$ must be open for every open $U \subset X$ and $V \subset Y$. The minimal topology on $X \times Y$ containing all these sets coincides with the product topology.

Note that the Cartesian product $S^1 \times I$, where I is the interval $[0,1]$, is the cylinder (Figure 1a) rather than the Möbius band (Figure 1b). The point is that, although the Möbius band admits a natural projection onto S^1, is does not admit any natural projection onto I.

For any subset Y of a topological space X, the *quotient space* X/Y is obtained by identifying all the points of Y. Thus, the points of the space X/Y are all the points of the set $X \setminus Y$ and one point Y. A subset of X/Y is open if and only if its preimage under the natural projection $p\colon X \to X/Y$ is open.

A quotient space can also be defined for a space X on which an equivalence relation $\sim$ is given. The points of the *quotient space* $X/\sim$ are the equivalence classes; a subset of $X/\sim$ is open if its preimage under the natural projection $p\colon X \to X/\sim$ is open. (Declaring that $x_1 \sim x_2$ if and only if $x_1 = x_2$ or $x_1, x_2 \in Y \subset X$, we obtain the preceding construction.)

Graphs

We consider graph theory in much more detail than it is usually done in topology courses. Sections 1 and 3 refer to graph theory proper and are not used in what follows; so those readers who are not interested in graph theory may skip them (except the definition of a graph in Section 1).

1. Topological and Geometric Properties of Graphs

Let us take several points $A_1, \dots, A_n$ in the space $\mathbb{R}^3$ and join some of them by pairwise disjoint finite-sided polygonal curves. The set thus obtained with the topology induced from $\mathbb{R}^3$ is called a *graph*, or a *one-dimensional complex*. The points $A_1, \dots, A_n$ are its *vertices*, or 0-*cells* (zero-dimensional cells), and the joining curves are its *edges*, or 1-*cells*. If an edge e contains a vertex v, then e and v are said to be *incident*. The number of edges incident to a vertex of a graph is called the *degree* of this vertex. If any vertex of a graph can be reached from any other vertex by a path along edges, then the graph is said to be *connected*.

A graph may have *loops* (a loop is an edge with coinciding end vertices) and *multiple edges* (that is, several edges may join the same pair of vertices).

A sequence of pairwise different vertices $v_1, \dots, v_n$ joined by edges v_1v_2, $v_2v_3, \dots, v_nv_1$, is called a *cycle*.

1.1. Planar Graphs. A graph G is said to be *planar* if it can be drawn in the plane in such a way that its edges are pairwise disjoint (except at the endpoints).

The nonplanarity of graphs is usually proved by applying the simplest version of the Jordan curve theorem, which deals with finite-sided polygonal curves.

Theorem 1.1 (piecewise linear Jordan theorem). *Let C be a closed self-avoiding finite-sided polygonal curve in the plane $\mathbb{R}^2$. Then $\mathbb{R}^2 \setminus C$ consists of precisely two connected domains, and the boundary of each domain is C.*

Proof. Choose a disk D that intersects C in a straight line segment. From any point of the set $\mathbb{R}^2 \setminus C$, we can approach arbitrarily closely the line C without intersecting it. Then, going along C, we can enter the disk D. The polygonal curve C divides D into two parts; hence the number of the connected domains under consideration is at most two.

It remains to prove that the set $\mathbb{R}^2 \setminus C$ is not connected. Take a point $x \in \mathbb{R}^2 \setminus C$ and let l be any ray emanating from x. The intersection of l with C consists of several points and line segments. To each of these points and segments we assign 0 or 1, depending on the arrangement of the edges of C with respect to l: if the endpoints of the edge or edges containing the point (or adjacent to the segment) under consideration are on the same side of l, then this point or segment is assigned 0; otherwise, it is assigned 1. The parity (remainder after division by 2) of the sum of all assigned numbers does not change under rotation of the ray. It is also clear that the parity must be the same at all points of a connected domain in $\mathbb{R}^2 \setminus C$. On the other hand, if some straight line segment intersects C in precisely one point, then the parity takes different values at its endpoints. □

Corollary. *Let a, b, c, and d be points of a closed self-avoiding polygonal curve C arranged on C in the order in which they are listed. Suppose that the points a and c are joined by a polygonal curve L_1, the points b and d are joined by a polygonal curve L_2, and both of these polygonal curves are contained in the same of the two domains determined by C. Then L_1 and L_2 intersect.*

Proof. The points a and c divide the polygonal curve C into two parts. The polygonal curves C and L_1 divide the plane into three domains, one of which is bounded by C and the other two are bounded by L_1 and segments of C (to see this, consider the endpoints of a line segment which intersects L_1 in one point and does not intersect C). By condition, the polygonal curve L_2 is contained in the same domain of $\mathbb{R}^2 \setminus C$ as L_1. Therefore, the domains that are determined by the polygonal curves C and L_1 and contain the points of L_2 close to b and d are different. □

The simplest examples of nonplanar graphs are the graphs $K_{3,3}$ and K_5 shown in Figure 1 (the vertices of these graphs are represented by heavy

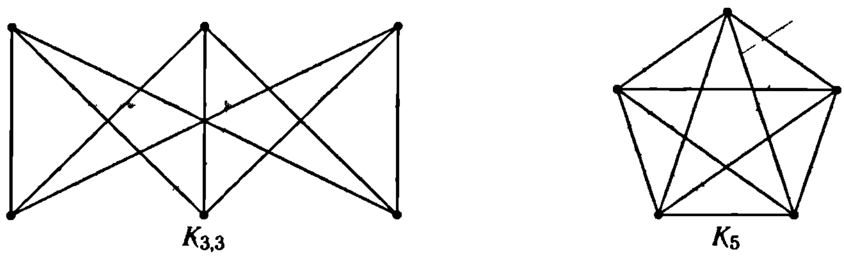

Figure 1. The graphs $K_{3,3}$ and K_5

dots: the graph $K_{3,3}$ has six vertices, and the graph K_5 has five vertices). Similar graphs K_n and $K_{n,m}$ can be defined for any n and m. The graph K_n (the *complete graph on n vertices*) consists of n vertices pairwise joined by edges. The graph $K_{n,m}$ has $n+m$ vertices grouped into two sets, one of which contains n vertices and the other contains m vertices; all pairs of vertices belonging to different sets are joined by edges.

Theorem 1.2. *The graphs $K_{3,3}$ and K_5 are nonplanar.*

Proof. The vertices of the graph $K_{3,3}$ can be numbered in such a way that the consecutive edges form a closed polygonal curve $x_1x_2x_3x_4x_5x_6$; the graph contains also the edges x_1x_4, x_2x_5, and x_3x_6. If $K_{3,3}$ were planar, then this polygonal curve would partition the plane into two domains, and two of the three edges would be contained in one of these domains. These edges must intersect.

The proof of the nonplanarity of K_5 is similar. The closed polygonal curve $x_1x_2x_3x_4x_5$ divides the plane into two domains. Three of the five other edges are contained in one of these domains, and two of them have no common vertices. □

Problem 5. (a) Is it possible to embed the graph $K_{3,3}$ in the Möbius band (i.e., draw it on the Möbius band in such a way that its edges are pairwise disjoint)?

(b) Is it possible to embed the graph K_5 in the torus?

Yet another approach to the proof of the nonplanarity of $K_{3,3}$ and K_5 is as follows. Suppose that the graphs are drawn in the plane (their edges may intersect; the number of intersection points is finite, and none of the edges passes through a vertex). We define the *intersection number* of two graphs as the total number of intersection points of the edges of one graph with the edges of the other graph modulo 2.

Problem 6. Prove that the intersection number of any two cycles is zero.

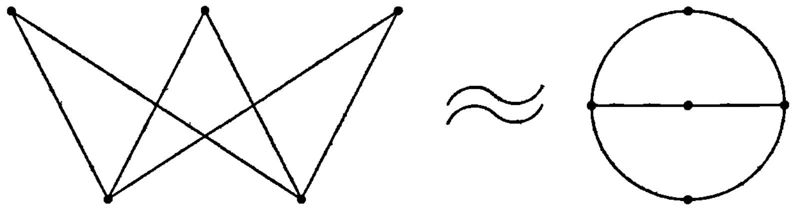

Figure 2. The graph $K_{3,2}$

The *self-intersection number* of a graph in the plane is defined as the sum of the intersection numbers of all (unordered) pairs of its nonadjacent edges modulo 2.

Problem 7. (a) Suppose that for each edge of some graph, the edges not adjacent to it form a cycle. Prove that the self-intersection number of such a graph depends only on the graph itself and does not depend on its embedding in the plane.

(b) Prove that the graphs $K_{3,3}$ and K_5 are nonplanar.

A graph G' is called a *subgraph* of a graph G if any vertex of G' is a vertex of G and any edge of G' is an edge of G.

Clearly, if a graph contains a subgraph homeomorphic[1] to $K_{3,3}$ or K_5, then it is nonplanar. In 1930, Kuratowski [72] proved the converse.

Theorem 1.3 (Kuratowski). *A graph is nonplanar if and only if it contains a subgraph homeomorphic to $K_{3,3}$ or K_5.*

Proof (see [80]). It remains to prove the difficult part of the Kuratowski theorem, namely, that if a graph is nonplanar, then it contains a subgraph homeomorphic to $K_{3,3}$ or K_5. Suppose that there exist nonplanar graphs without subgraphs homeomorphic to $K_{3,3}$ or K_5. Among all such graphs we choose a graph G with minimum number of edges.

Step 1. Let xy be an edge of G. Removing the vertices x and y (and the edges incident to them) from G, we obtain a graph $G - x - y$ containing no subgraphs homeomorphic to $K_{3,2}$ (see Figure 2).

Indeed, suppose that the graph $G' = G - x - y$ contains a subgraph homeomorphic to $K_{3,2}$.

The graph G/xy obtained from G by contracting the edge xy to a point has no subgraphs homeomorphic to $K_{3,3}$ or K_5. Indeed, if G/xy had a subgraph homeomorphic to $K_{3,3}$, then G would have a subgraph homeomorphic

[1] A graph can be treated as a topological space. We mean here a homeomorphism between such spaces.

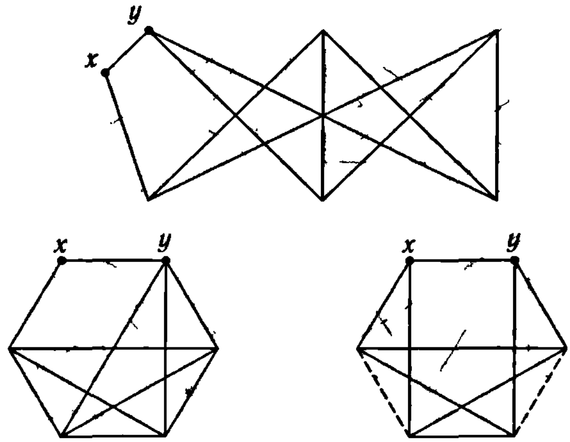

Figure 3. The graph G contains $K_{3,3}$ or K_5

to $K_{3,3}$, and if G/xy had a subgraph homeomorphic to K_5, then G would have either a subgraph homeomorphic to K_5 or a subgraph homeomorphic to $K_{3,3}$ (see Figure 3).

The graph G/xy is planar because G is minimal. Therefore, the graph $G' = G - x - y = (G/xy) - (xy)$ is planar also. Consider an embedding of G/xy in the plane and the induced embedding of G' in the plane. One of the domains into which the plane is divided by the edges of G' contains the vertex xy of the graph G/xy. Let F be the subgraph of G' consisting of the edges which bound this domain. The graph F contains no subgraphs homeomorphic to $K_{3,2}$ (otherwise, we could take a point inside the domain, join it by edges with the three vertices on the three arcs of the graph $K_{3,2}$, and obtain an embedding of $K_{3,3}$ in the plane). By assumption, the graph G' contains a subgraph homeomorphic to $K_{3,2}$. Therefore, G' has an edge e different from the edges of F. This means, in particular, that the graph F divides the plane into two parts. Hence it contains a cycle C. The removed vertex xy and the edge e are contained in different parts of the plane determined by the cycle C. Suppose for definiteness that the vertex xy is inside C and the edge e is outside C.

To obtain a contradiction, we construct an embedding of G in the plane. For $\operatorname{ext} C$, i.e., for the part of the graph $G' \subset G$ that is exterior to the cycle C, we use the already existing embedding of G'. We denote the remaining part of G by $G - \operatorname{ext} C$. It has strictly fewer edges than G because it does not contain the edge e. Since G is minimal, the graph $G - \operatorname{ext} C$ is planar. In this graph, the vertices x and y are joined by an edge; therefore, under any embedding of $G - \operatorname{ext} C$ in the plane, the vertices x and y are either both inside C or both outside C. We can assume that they are inside C.

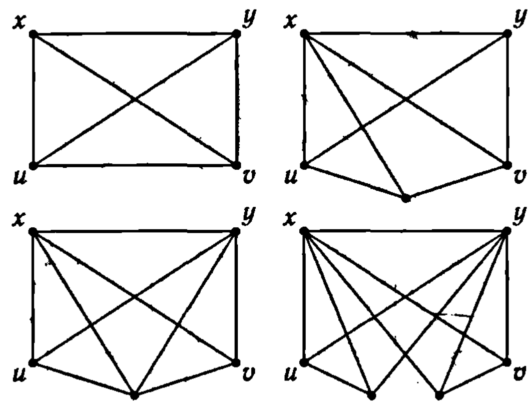

Figure 4. The four possible cases for the graph G

Thus, the planar graph $G - \operatorname{ext} C$ admits an embedding such that the cycle C is the boundary of the embedded graph. Combining the embeddings of the graphs $\operatorname{ext} C$ and $G - \operatorname{ext} C$, we obtain an embedding of the graph G.

Step 2. If xy is an edge of G, then the graph $G - x - y$ cannot have two vertices of degree 1.

Suppose that u and v are vertices of $G - x - y$ and each of them is incident to a single edge. Since the graph G is minimal, it has no vertices incident to fewer than three edges. Therefore, the vertices u and v are joined by edges to the vertices x and y. The vertices x and y are joined by an edge to each other; hence the subgraph of G generated by the vertices x, y, u, and v contains $K_{3,2}$. According to Step 1, the graph G cannot have edges that are not incident to any of the vertices x, y, u, and v. Let w be a vertex of G different from x, y, u, and v. The vertex w is incident to at least three edges; therefore, it is joined by an edge to one of the vertices u and v. By assumption, each of the vertices u and v is joined by an edge to no more than one vertex different from x and y. Hence the graph G contains no more than two vertices different from x, y, u, and v. Thus, G is one of the four graphs shown in Figure 4.

Step 3. If xy is an edge of G, then the graph $G - x - y$ is a cycle.

Let $G' = G - x - y$. The graph G' has no isolated vertices because each isolated vertex of G' corresponds to a vertex of G incident to at most two edges, which contradicts the minimality of G. As shown at Steps 1 and 2, the graph G' consists of one or several "trees" with vertex-cycles (Figure 5), and G' cannot have more than one pendant edge.

The graph G' contains a pendant cycle C which has precisely one vertex v joined by an edge to some vertex w not belonging to C (see Figure 6).

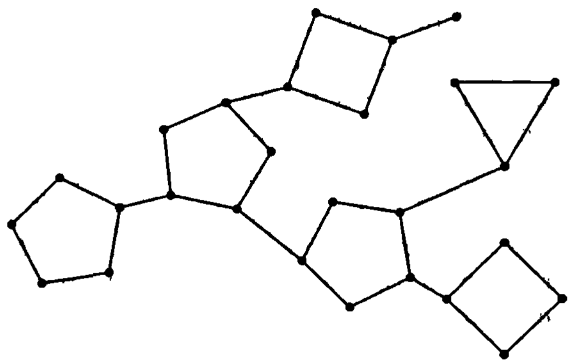

Figure 5. A connected component of the graph G'

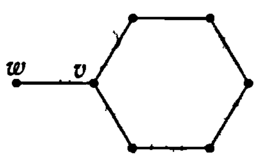

Figure 6. A subgraph of G'

In the graph G, any vertex of C except v is joined by an edge to x or y. There are at least two such vertices; therefore, the vertices of the cycle C together with x and y generate a graph containing a subgraph homeomorphic to $K_{3,2}$. It follows that each edge of G' has common vertices with C. But the only vertex which can be shared by the cycle C and an edge not included in C is v, and v is incident to only one edge not contained in C. Therefore, the graph G consists of the subgraph shown in Figure 6 and edges with vertices x and y.

Now, the proof of Kuratowski's theorem is easy to complete. Let x and y be adjacent vertices of G. Then the graph $G - x - y$ is a cycle (we denote it by C) each of whose vertices is incident to x or y in the graph G. Suppose that a vertex u of C is joined to x and not joined to y. Then the vertex v neighboring u in C is not joined to x. Indeed, if the graph G contains the edge xv, then, removing this edge, we obtain a planar graph. This planar graph does not contain the edge yu; therefore, we can join the points x and v in the plane and obtain an embedding of the graph G, which does not exist.

Thus, either all vertices of the cycle C are joined to both vertices x and y or they are joined to x and y alternately. In the former case, the graph G contains a subgraph homeomorphic to K_5, and in the latter case, the graph G contains a subgraph homeomorphic to $K_{3,3}$. □

Different proofs of Kuratowski's theorem are contained in [**124**].

Kuratowski's theorem is not the only criterion for graph planarity; see, e.g., [**8**, **22**, **81**]. One of the most interesting criteria belongs to Whitney [**142**]; it is based on the duality of graphs.

Graphs G and G^* are said to be *dual* if there exists a one-to-one correspondence between their edges under which the cycles of one graph correspond to the cut sets of the other, and *vice versa.* A *cut set* is a minimal set of edges such that removing these edges increases the number of connected components of the graph.

Theorem 1.4 (Whitney). *A graph is planar if and only if there exists a graph dual to it.*

Proof (see [**96**]). First, we show that any planar graph has a dual. Consider an embedding of a planar graph G in the plane and choose one point in each of the domains into which G divides the plane. These points are vertices of the dual graph G^*. Two vertices of G^* are joined by an edge if the corresponding parts of the plane are adjacent along some edge of G. Clearly, the graphs G and G^* are dual to each other.

Now, we prove the difficult part of Whitney's theorem: if a graph is nonplanar, then it has no dual. We use Kuratowski's theorem; namely, we shall prove that if a graph has a subgraph homeomorphic to $K_{3,3}$ or K_5, then it has no dual.

Step 1. The graphs $K_{3,3}$ and K_5 have no duals.

Suppose that G is a graph dual to $K_{3,3}$. The graph $K_{3,3}$ has no cut sets consisting of fewer than three edges, and it has cycles only of lengths 4 and 6. Therefore, the graph G has no cycles of length 1 or 2, and each of its vertices is incident to at least four edges. These two observations imply that G has at least five vertices. Each vertex is incident to at least four edges. Thus, G has at least $5 \cdot 4/2 = 10$ edges. This contradicts the fact that $K_{3,3}$ has only nine edges.

Now, suppose that G is a graph dual to K_5. The graph K_5 has no multiple edges, and it has only cut sets consisting of four and six edges. Therefore, the graph G has no vertices of degree less than 3, and it has only cycles of lengths 4 and 6. The graph G has ten edges and does not coincide with K_5; hence it has at least six vertices. If G has precisely six vertices, then it must be of the form shown in Figure 7. Such a graph has nine edges, while K_5 has ten edges. If G has seven or more vertices (which must be of degree at least 3), then the number of its edges is at least $7 \cdot 3/2 > 10$.

Step 2. If a graph G has a dual, then each of its subgraphs has a dual.

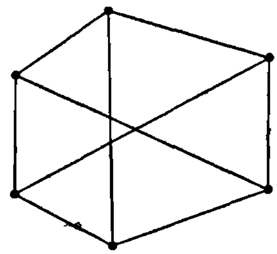

Figure 7. The structure of the graph G on six vertices

It suffices to prove that if a graph G has a dual G^* and e is an edge of G, then the graph H obtained from G by removing the edge e has a dual H^*. It is easy to verify that if e^* is the edge of G^* corresponding to an edge e of G, then the graph H^* obtained from G^* by contracting the edge e^* to a point is dual to H.

Step 3. If a graph G has a dual, then any graph H homeomorphic to G has a dual.

It suffices to prove that if a graph G has a dual G^* and H is obtained from G by adding a vertex of degree 2 on an edge e of G, then H has a dual H^*. It is easy to verify that the graph H^* obtained from G^* by adding one edge with the same end vertices as e^* is dual to H. □

The edges of a planar graph drawn in the plane are generally arbitrary curves; however, Wagner [136] and Fary [36] independently proved the following theorem.

Theorem 1.5. *Any planar graph without loops and multiple edges can be embedded in the plane in such a way that all of its edges are straight line segments.*

Proof. It suffices to prove the theorem for maximal planar graphs. (A planar graph is *maximal* if it ceases to be planar after the addition of any edge.) Clearly, each of the *faces* (i.e., the regions into which the graph divides the plane) of a maximal planar graph contains precisely three edges. Let G be a maximal planar graph on $v \geq 4$ vertices (for $v < 4$, the assertion is obvious). Take any vertex V_1 of G different from the vertices of the curvilinear triangle bounding G. Let G_1 be the graph obtained from G by removing the vertex V_1 and the edges incident to it. In the graph G_1, all faces except the face F_1 that contained the removed vertex V_1 are triangular. The face F_1 is bounded by a cycle C_1. Choose a vertex V_2 in C_1 different from the vertices of the triangle bounding G and consider the graph G_2 obtained from G_1 by removing the vertex V_2 and the edges incident to it.

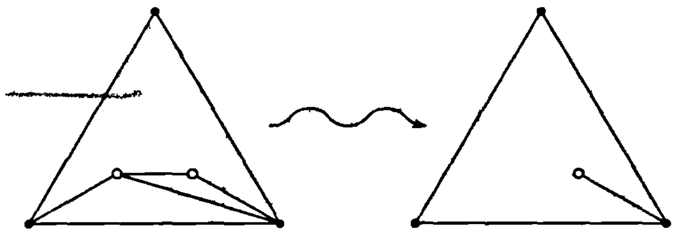

Figure 8. The boundary of a face is not necessarily a a cycle

In the graph G_2, the face F_2 that contained V_2 is not necessarily bounded by a cycle (an example is given in Figure 8).

To ensure that the face F_2 is bounded by some cycle C_2, we must choose the vertex V_2 in a special way. Namely, suppose that the cycle C_1 contains a vertex V of degree 2 (in the graph G_1) and V is different from the vertices of the triangle bounding G. Then for V_2 we take this vertex V. The endpoints of the edges incident to V are joined by an edge; therefore, removing V, we obtain a cycle C_2. If the degrees of all the vertices of C_1 that are not vertices of the triangle bounding G are different from 2, then any such vertex can be taken for V_2.

Continuing this procedure, we obtain a sequence of graphs G, G_1, G_2, $\dots$, G_{v-3}, where G_{v-3} is the graph consisting of three vertices pairwise joined by edges; the boundary of each face F_i is a cycle. Now, we can construct the required embedding of G step by step, starting with the graph G_{v-3}. The embedded graph G_{v-3} is an arbitrary triangle. For the vertex V_{v-3} we take any point inside this triangle. We join the point V_{v-3} to two or three vertices of the triangle, depending on the graph G_{v-4}. This divides the triangle into three triangles or into a triangle and a nonconvex quadrangle. If the vertex V_{v-4} is to be chosen inside one of the new triangles, then we place it anywhere in this triangle. If the vertex V_{v-4} must belong to the nonconvex quadrangle, then we place it in the region which is hatched in Figure 9. This is the region from which all the vertices of the cycle are visible. At each step, from a given visibility domain, we construct a new visibility domain (this is a nonempty open set; see Figure 10) and place the vertex V_{i-1} in this domain (i.e., in the visibility domain of the vertex V_i obtained at the previous step). □

1.2. The Euler Formula for Planar Graphs. For a convex polyhedron (in three-dimensional space), the following *Euler formula* holds: If v is the number of vertices of the polyhedron, e is the number of its edges, and f is the number of faces, then $v - e + f = 2$. The graph formed by the edges of a convex polyhedron in 3-space is planar; indeed, removing a point from the

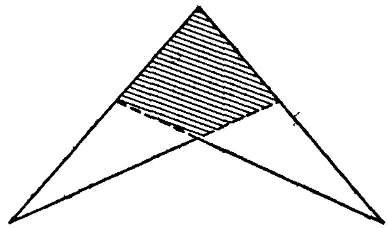

Figure 9. The domain of visibility for a nonconvex quadrangle

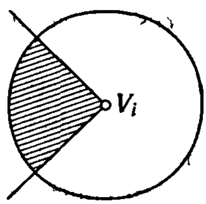

Figure 10. The domain of visibility for one of the new faces

surface of a convex polyhedron, we obtain a topological space homeomorphic to the plane.

For planar graphs, the Euler formula remains valid in the general case. We refer to the connected domains into which the plane is divided by an embedded planar graph as *faces*.

Theorem 1.6 (Euler formula). *Let G be a planar graph having s connected components. Suppose that v is the number of vertices of G and e is the number of edges. Then, for any embedding of G in the plane, the number f of faces is the same and equals $f = 1 + s - v + e$.*

Proof. If a graph contains no cycles, then it does not separate the plane into domains. The connected components of such a graph are called *trees*. By induction on the number of edges, it is easy to prove that the number of vertices of any tree is the number of edges plus one. Indeed, the removal of any edge splits the tree into two trees with fewer edges. Therefore, the Euler formula does hold for any graph consisting of one or several trees.

If a graph contains a cycle, then we consider the domain bounded by this cycle and not contained in another domain bounded by a cycle. For such a domain, the removal of one boundary edge reduces the number of faces by one and does not change the number of vertices. □

Corollary. *A connected planar graph (without loops and multiple edges) contains a vertex of degree at most* 5.

Proof. Any face has at least three edges; therefore, $3f \le 2e$. Substituting this inequality into the relation $3f = 6 - 3v + 3e$, we obtain $e \le 3v - 6$. Suppose that each vertex is incident to at least six edges. Then $6v \le 2e$, and hence $6v \le 2e \le 2(3v - 6) = 6v - 12$, which is impossible. □

Using this corollary, we can easily prove the following celebrated map coloring theorem.

Theorem 1.7 (five-color theorem). *Any planar graph (without loops and multiple edges) is 5-colorable, i.e., its vertices can be colored with five colors in such a way that any two vertices joined by an edge have different colors.*

Proof. Let G be a planar graph on n vertices. We use induction on n. For $n \le 5$, the assertion of the theorem is obvious. Suppose that it is proved for all planar graphs on at most $n - 1$ vertices. If the graph G has a vertex v of degree strictly less than 5, we consider the graph G' obtained from G by removing the vertex v and the edges incident to it. By the induction hypothesis, the graph G' is 5-colorable. In the graph G, the vertex v is joined by edges to fewer than five vertices; therefore, it can be colored differently from the neighboring vertices.

Now, suppose that the graph G has no vertices of degree strictly less than 5. Then it has a vertex v of degree precisely 5. The vertices neighboring v in G are not pairwise joined by edges, because if they were, then the graph G would contain the nonplanar graph K_5. Let v_1 and v_2 be vertices of G joined by edges to the vertex v and not joined to each other. First, consider the graph G' obtained from G by removing the vertex v and the edges incident to it. Then, consider the graph G'' obtained from G' by adding an edge joining the vertices v_1 and v_2. This additional edge can be composed of the edges v_1v and vv_2; therefore, the graph G'' is planar. Finally, in G'', we contract the additional edge to a point. As a result, we obtain a planar graph G''' on $n - 2$ vertices. By the induction hypothesis, this graph is 5-colorable. Its 5-coloring induces a coloring of the vertices of G' under which the vertices v_1 and v_2 are of the same color. This means that the vertices neighboring v in G are colored with at most four different colors. Therefore, the vertex v can be colored differently from the neighboring vertices. □

Remark 1.1. Actually, any planar graph is 4-colorable (this is the *four-color theorem*), but the proof is very complicated. The first published proof of the four-color theorem [**5**, **7**] was obtained with the help of a computer and occupied 150 pages, but a complete exposition of this proof [**6**] took 740 pages. Then, simpler proofs appeared. For example, the proof given in [**110**] takes a little more than 40 pages, but still, this proof is very involved. It was also obtained with the help of a computer.

Problem 8. (a) Let G be a planar graph such that each face of G contains an even number of edges. Prove that the vertices of this graph can be colored with two colors in such a way that any two vertices joined by an edge have different colors.

(b) Let γ be a smooth closed curve such that all its self-intersections are transversal. Prove that the domains into which γ divides the plane can be colored with two colors in such a way that any two domains adjacent along some arc have different colors.

The Euler formula has various consequences, including the following formula, which is used most frequently.

Theorem 1.8. *Let G be a planar graph. For each i, let v_i be the number of its vertices incident to i edges, and for each j, let f_j be the number of faces bounded by j edges (counting multiplicities[2]). Then $\sum_i (4-i)v_i + \sum_j (4-j)f_j = 4(1+s) \geq 8$, where s is the number of connected components of the graph G.*

Proof. Clearly, $\sum_i iv_i = 2e = \sum_j jf_j$ (each edge has precisely two end vertices and belongs to precisely two faces). Moreover, $\sum_i v_i = v$ and $\sum_j f_j = f$. Therefore, the Euler formula implies

$$\begin{aligned}\sum_i (4-i)v_i + \sum_j (4-j)f_j &= 4v - 2e + 4f - 2e \\ &= 4(v-e+f) = 4(1+s),\end{aligned}$$

where s is the number of connected components in G. □

Corollary. *If all faces of G are tetragonal, then $3v_1 + 2v_2 + v_3 \geq 8$.*

The following inequality is also useful.

Theorem 1.9. *If each face of G is bounded by a cycle with at least n edges, then $e \leq \frac{n(v-2)}{n-2}$.*

Proof. This follows from the inequalities $nv - ne + nf \geq 2n$ and $2e \geq nf$. □

Problem 9. Give a different proof, using Theorem 1.9, of the nonplanarity of the graphs K_5 and $K_{3,3}$.

[2]A boundary edge may have multiplicity 2. For example, cutting an annulus bounded by two circles along a radius, we obtain a rectangular, rather than triangular, face, although it is bounded by three edges (the radius should be counted twice).

1.3. Embeddings of Graphs in Three-Dimensional Space. Not every graph can be embedded in the plane. But any finite graph can be embedded in 3-space. Moreover, any graph can be embedded in 3-space in such a way that all of its edges are straight line segments. For example, if the vertices of a graph belong to the curve (t, t^2, t^3), then the segments joining them cannot intersect. Indeed, the points of this curve with parameters t_1, t_2, t_3, and t_4 are the vertices of a tetrahedron with volume

$$\pm\frac{1}{6}\begin{vmatrix} 1 & t_1 & t_1^2 & t_1^3 \\ 1 & t_2 & t_2^2 & t_2^3 \\ 1 & t_3 & t_3^2 & t_3^3 \\ 1 & t_4 & t_4^2 & t_4^3 \end{vmatrix} \neq 0;$$

in particular, the opposite edges of this tetrahedron are disjoint.

Now, consider the embeddings in $\mathbb{R}^3$ of the graph K_6 with six vertices pairwise joined by edges. Choose three vertices in K_6. Let C_1 be the cycle generated by these three vertices, and let C_2 be the cycle generated by the three remaining vertices. Consider a projection of the graph K_6 embedded in $\mathbb{R}^3$. We define $\omega(C_1, C_2)$ to be 0 or 1 depending on whether the number of the crosses where C_1 passes over C_2 is even or odd. In other words, $\omega(C_1, C_2) = \operatorname{lk}(C_1, C_2) \pmod 2$, where lk is the linking coefficient. In particular, $\omega(C_1, C_2) = \omega(C_2, C_1)$ (this property of the linking coefficient was proved in [102]). Therefore, we can consider the number $\lambda(K_6) = \sum \omega(C_i, C_j)$, where the summation is over the $\frac{1}{2}\binom{6}{3} = 10$ unordered pairs of three-element disjoint cycles.

Theorem 1.10 ([115, 29]). *For any embedding of the graph K_6 in 3-space, $\lambda(K_6) \equiv 1 \pmod 2$. In particular, under any such embedding, there is a pair of linked cycles.*

Proof. The graph K_6 can be embedded in $\mathbb{R}^3$ so that precisely two cycles are linked and all of the remaining cycles are unlinked (Figure 11).

Any embedding of K_6 in $\mathbb{R}^3$ can be transformed into such an embedding by using the transformations of edges shown in Figure 12.

Let us see what happens to $\lambda(K_6)$ under changing the type of crossing of a pair of edges e_i and e_j. The number $\omega(C_p, C_q)$ changes only if $e_i \subset C_p$ and $e_j \subset C_q$ (or $e_j \subset C_p$ and $e_i \subset C_q$). The edges e_i and e_j are contained in disjoint cycles C_p and C_q if and only if the edges e_i and e_j are nonadjacent. There are precisely two such pairs of cycles for given edges e_i and e_j: to the edge e_i we can add one of the two vertices not incident to e_i and e_j. Thus, the number $\sum \operatorname{lk}(C_i, C_j)$ remains the same when the self-crossing of an edge or the crossing of two adjacent edges is changed, and it increases or decreases by 2 when the crossing of nonadjacent edges is changed Therefore, the number

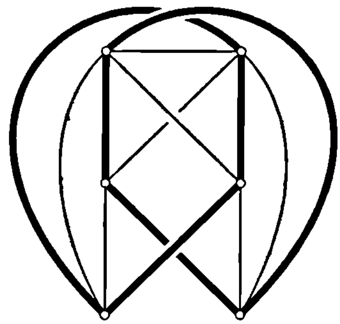

Figure 11. The graph K_6 with two linked cycles

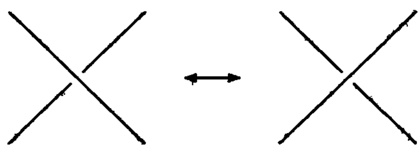

Figure 12. A change of a crossing type

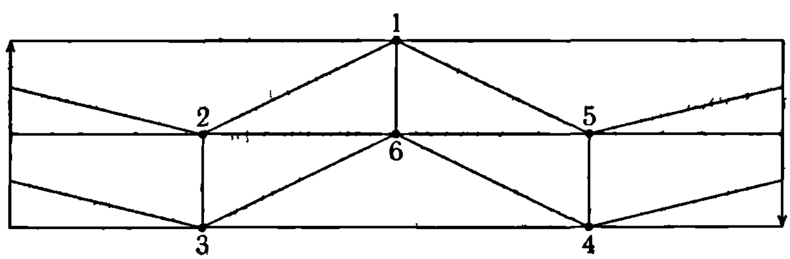

Figure 13. Embedding of the graph K_6 in the Möbius band

$\lambda(K_6) = \sum \mathrm{lk}(C_i, C_j) \pmod 2$ does not change under any transformations of the embedding of K_6. □

Corollary 1. *The boundary of the Möbius band embedded in $\mathbb{R}^3$ is always linked with its median.*

Proof (see [78]). Let us embed the graph K_6 in the Möbius band as shown in Figure 13.

The cycles 134 and 256 correspond to the boundary of the Möbius band and its median, respectively. It is easy to see that in all the other pairs of self-avoiding cycles, one of the cycles bounds a triangular domain contained in the Möbius band. Such cycles cannot be linked because if they were, then the Möbius band would have self-intersections. For any nonlinked cycles C_i

and C_j, we have $\omega(C_i, C_j) = 0$. Therefore, the cycles 134 and 256 are linked. □

Corollary 2. *The projective plane $\mathbb{R}P^2$ cannot be embedded in $\mathbb{R}^3$.*

Proof. Removing the disk D^2 from the projective plane embedded in $\mathbb{R}^3$, we obtain the Möbius band. Its median C is linked with $S^1 = \partial D^2$; therefore, C intersects D^2, which is impossible. □

Using Corollary 1 (that the boundary of the Möbius band is linked with its median under any embedding in $\mathbb{R}^3$), we can prove the following theorem.

Theorem 1.11. *Any smooth closed plane curve γ contains four points that are vertices of a rectangle.*

Proof. To each pair of points A, B on the curve γ we assign a point $f(A, B)$ as follows. From the midpoint of the line segment AB, we draw a perpendicular to the plane of the curve γ and take a point on it whose distance from the midpoint equals the length of AB. This point is $f(A, B)$. (All points $f(A, B)$ are chosen in the same half-space.) The case $A = B$ is not excluded; in this case, $f(A, A) = A$.

We have obtained a map from a topological space to $\mathbb{R}^3$. It is easy to see that this topological space is the torus $S^1 \times S^1$ whose points are identified by the rule $(x, y) \sim (y, x)$ (we do not distinguish between pairs A, B and B, A). This is the Möbius band.

Suppose that the map f is a bijection onto its image. Then it gives an embedding of the Möbius band into 3-space such that the boundary of the Möbius band is contained in some plane and the remaining part lies on one side of this plane. Hence the boundary of the Möbius band is not linked to the median. We have obtained a contradiction.

Thus, there exist two pairs of points A_1, B_1 and A_2, B_2 for which $f(A_1, B_1) = f(A_2, B_2)$. This means that the segments A_1B_1 and A_2B_2 have coinciding midpoints and equal lengths. Therefore, $A_1A_2B_1B_2$ is a rectangle. □

1.4. *k*-Connected Graphs. Two paths in a graph that join vertices x and y along edges are said to be *internally disjoint* if they have no common vertices except x and y.

A graph is *k-connected*[3] if it has at least $k + 1$ vertices and any two different vertices can be joined by at least k internally disjoint paths.

[3] In homotopy theory, this term has a quite different meaning.

Theorem 1.12 (Menger Whitney). *A graph G with at least $k+1$ vertices is k-connected if and only if every graph obtained from G by removing $k-1$ vertices (and the edges incident to them) is connected.*

Proof (see [**92**]). We prove a more general assertion. In this proof, we consider only paths containing at least two edges. We show that if $p(G,x,y)$ is the maximum number of internally disjoint paths from a vertex x to a vertex y and $q(G,x,y)$ is the minimum number of points different from x and y and such that any path from x to y passes through one of them, then $p(G,x,y) = q(G,x,y)$.

The inequality $p(G,x,y) \geq q(G,x,y)$ is fairly obvious. Indeed, let $\gamma_1,\dots,\gamma_p$ be internally disjoint paths from x to y, and let $x_1,\dots,x_q$ be points (different from x and y) such that any path from x to y passes through one from them. Since the paths $\gamma_1,\dots,\gamma_p$ are internally disjoint, it follows that each of them passes through at most one of the points $x_1,\dots,x_q$. On the other hand, each path from x to y passes through one of the points $x_1,\dots,x_q$; therefore, $p \geq q$.

Suppose that G is a graph with minimum number of edges for which the equality $p(G,x,y) = q(G,x,y)$ does not hold. Then $p = p(G,x,y) < q(G,x,y) = q$. The graph G has edges different from xy. Let λ be one of them. Suppose that $G-\lambda$ is the graph obtained from G by removing the edge λ, and $\widehat{G} = G/\lambda$ is the graph obtained from G by contracting λ to a point. Each of the graphs $G-\lambda$ and $\widehat{G}$ has strictly fewer edges than G; therefore, by assumption, $p(G-\lambda,x,y) = q(G-\lambda,x,y)$ and $p(\widehat{G},x,y) = q(\widehat{G},x,y)$, whence $q(G-\lambda,x,y) = p(G-\lambda,x,y) \leq p(G,x,y) = p < q$. Similarly, $q(\widehat{G},x,y) < q$. Thus, there are sets of vertices I and $\hat{J}$ in the graphs $G-\lambda$ and $\widehat{G}$, respectively, which separate x and y and have fewer than q elements. The set $\hat{J}$ corresponds to a set J of vertices separating x and y in G. Moreover, we have $|J| \leq |\hat{J}|+1$ and $|J| \geq q$. Therefore, $|J| = |\hat{J}|+1$. This means that both end vertices of the edge λ belong to the set J.

Let H_x be the set of vertices $z \in I \cup J$ joined to x by paths in G avoiding the other vertices from the set $I \cup J$; H_y is defined similarly. Any path from x to y in the graph G passes through at least one of the vertices from J; in particular, it passes through some vertices from $I \cup J$. The first of these vertices belongs to H_x, and the last belongs to H_y. Therefore, the sets H_x and H_y separate the vertices x and y in the graph G, whence $|H_x| \geq q$ and $|H_y| \geq q$.

Let $z \in H_x \cap H_y$. Then there are paths from x to z and from z to y in G that avoid the vertices from $I \cup J$ different from z. These two paths form one path γ from x to y. The path γ passes through precisely one vertex, z, from the set $I \cup J$. Hence γ avoids the edge λ because both end vertices of λ belong to J. Thus, the path γ belongs to the graph $G - \lambda$; hence it passes through one of the vertices from the set I. Only z can be this vertex. Moreover, the path γ passes through one of the vertices from J; again, this can be only the same vertex z. Thus, $z \in I \cap J$, i.e., $H_x \cap H_y \subset I \cap J$. Therefore,

$$|H_x| + |H_y| = |H_x \cap H_y| + |H_x \cup H_y| \leq |I \cap J| + |I \cup J| = |I| + |J|,$$

which is impossible because $|H_x| \geq q$, $|H_y| \geq q$, $|I| < q$, and $|J| = q$. □

Corollary. *Let G_1 and G_2 be k-connected subgraphs of the same graph. If $|G_1 \cap G_2| \geq k$, then the graph $G_1 \cup G_2$ is k-connected.*

Proof. According to the Menger–Whitney theorem, the graphs G_1 and G_2 remain connected after the removal of any $k-1$ vertices from the graph $G_1 \cup G_2$. The graphs G_1 and G_2 have a common vertex different from the removed vertices; therefore, $G_1 \cup G_2$ remains connected as well. □

An important example of n-connected graphs is the graphs formed by the edges of convex polyhedra in n-space. We say that a graph is *n-realizable* if it can be realized as the edge set of a (nondegenerate) convex polyhedron in $\mathbb{R}^n$.

Theorem 1.13 (Balinski [10]). *Any n-realizable graph is n-connected.*

Proof. Let P^n be a polyhedron in $\mathbb{R}^n$ whose edges form the graph under consideration. We must prove that removing any vertices $A_1, \dots, A_{n-1}$ and the edges incident to them, we always obtain a connected graph. Let V be the affine space spanned by the points $A_1, \dots, A_{n-1}$. There are two possible cases.

Case 1. The space V contains no interior points of the polyhedron P^n.

In this case, $V \cap P^n = F_1^k$ is a face of P^n. Let H_1 be the supporting hyperplane of P^n that contains the face F_1^k, and let H_2 be the supporting hyperplane parallel to H_1 (but not coinciding with H_1). We set $F_2^l = P^n \cap H_2$. If A is a vertex of the polyhedron P^n different from $A_1, \dots, A_{n-1}$, then we can descend from A to the hyperplane H_2 along edges of the polyhedron without ascending to hyperplane H_1 (in particular, avoiding the vertices $A_1, \dots, A_{n-1}$ and the edges incident to them). From another vertex B, we can arrive at some vertex of the polyhedron F_2^l in the same manner. It remains to note that the vertices of F_2^l form a connected graph.

Case 2. The space V contains interior points of the polyhedron P^n.

The dimension of V does not exceed $n-2$. Hence there exists a hyperplane H containing V and at least one vertex A_n of P^n that does not belong to V. Let H_1 and H_2 be the supporting hyperplanes of P^n parallel to H. The same argument as that used in Case 1 shows that we can go from any vertex A different from $A_1, \ldots, A_{n-1}$ to the vertex A_n without meeting the vertices $A_1, \ldots, A_{n-1}$ and the edges incident to them. To do this, we have to descend or ascend to the hyperplane H_1 or H_2. Clearly, if we can go from any vertex to A_n, then we can go from any vertex to any other vertex (through A_n). □

1.5. Steinitz' Theorem. The edges of a convex polyhedron (in 3-space) form a 3-connected graph (by Theorem 1.13 on p. 22). This graph is obviously planar because the surface of a convex polyhedron from which one point is removed is homeomorphic to the plane. It turns out that the 3-connectedness and planarity are not only necessary but also sufficient for a graph to be realizable as the edge set of a convex polyhedron.

Theorem 1.14 (Steinitz [122]). *A graph*[4] *can be realized as the edge set of a convex polyhedron in 3-space if and only if it is 3-connected and planar.*

Proof (see [14]). Recall that a graph is 3-connected if and only if it contains at least four vertices and the graph obtained from it by removing any two vertices and the edges incident to them is connected (see Theorem 1.12 on p. 21). A 3-connected graph cannot have vertices incident to fewer than three edges; therefore, a 3-connected graph on n vertices must have at least $n \cdot 3/2$ edges. Thus, the minimum number of edges is that of the 3-connected graph K_4 formed by the edges of the tetrahedron.

We prove Steinitz' theorem by induction on the number of edges of a 3-connected planar graph. We start with the graph K_4, which has six edges. The induction step is performed as follows.

(i) First, to a 3-connected planar graph G with more than six edges we associate a 3-connected planar graph G' with fewer edges.

(ii) Then, from the convex polyhedron P' whose edges form the graph G' we construct a convex polyhedron P whose edges form the graph G.

Let G be a graph with an edge e. We define the operation of *deletion* of the edge e as follows. First, we remove the edge e from G; then, we remove all vertices of degree 2 from the resulting graph (if they appear), i.e., replace two edges having a common vertex incident to no other edges by one edge (see Figure 14).

The graphs under consideration must have no loops and multiple edges; therefore, not every edge can be deleted because the deletion may result in the appearance of a loop or a multiple edge.

[4] Containing no loops and multiple edges.

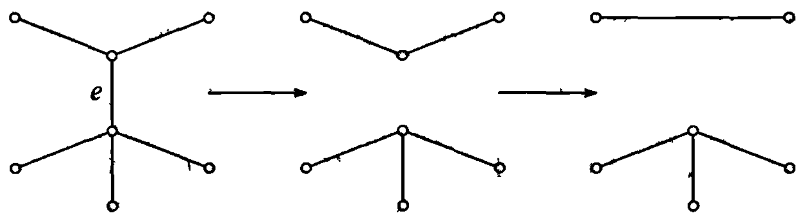

Figure 14. The deletion of an edge e

Step 1. Any 3-connected planar graph G with more than six edges has an edge e such that the graph G' obtained by deleting it is 3-connected and planar.

The planarity of the graph obtained by deleting an edge is obvious. Instead of proving the 3-connectedness, we prove a more general assertion.

Let $\Pi = \{\pi_0, \dots, \pi_n\}$ be a set of self-avoiding paths in a graph G. To the graph G and the set of paths Π we assign a one-dimensional complex G_Π, which may have loops and multiple edges. The vertex set of the complex G_Π consists of the end vertices of the paths $\pi_i \in \Pi$ in G and the vertices of G contained in at least two paths π_i. The edges of G_Π are the arcs of π_i cut out by the vertices of G_Π.

Lemma. *Let G be a 3-connected graph. Then there exists a set of paths $\{\pi_0, \dots, \pi_n\}$ such that the set $\Pi(k) = \{\pi_0, \dots, \pi_k\}$, where $1 \le k \le n$, has the following properties:*

(1) *the complex $G_{\Pi(k)}$ is a 3-connected graph;*

(2) $G_{\Pi(1)} = K_4$;

(3) $G_{\Pi(n)} = G$;

(4) *for $k = 1, \dots, n-1$, the path π_{k+1} can intersect $G_{\Pi(k)}$ only in its endpoints.*

Proof. We construct the set of paths $\{\pi_i\}$ by induction. First, we prove that any 3-connected graph G contains a subgraph homeomorphic to K_4. Let x and y be two different vertices of G. By assumption, there exist internally disjoint paths σ_1, σ_2, and σ_3 from x to y. Only one of these three paths can be an edge. Suppose for definiteness that the paths σ_2 and σ_3 pass through vertices z_2 and z_3 different from x and y. Removing x and y, we obtain a connected graph; therefore, the points z_2 and z_3 can be joined by a path σ avoiding x and y. The path σ may share segments with σ_2 and σ_3, but it also has a segment disjoint from $\sigma_2 \cup \sigma_3$ and joining vertices $v \in \sigma_2$ and $w \in \sigma_3$. The vertices x, y, v, and w and the arcs which they cut out of the paths σ, σ_1, σ_2, and σ_3 form a subgraph homeomorphic to K_4.

Among all subgraphs of G homeomorphic to K_4 we choose a subgraph T containing a maximum number of vertices of the graph G. Suppose that it has vertices x, y, v, and w. Taking the path $xyvw$ for π_0 and $vxwy$ for π_1, we obtain $G_{\Pi(1)} = K_4$.

Suppose that paths $\pi_0, \dots, \pi_k$ are already constructed and $G_{\Pi(k)} \neq G$. Then one of the following two conditions holds:

(a) the graph G has a vertex z which belongs to an edge of $G_{\Pi(k)}$ but is not a vertex of $G_{\Pi(k)}$;

(b) condition (a) does not hold but the graph $G_{\Pi(k)}$ has a vertex z which is an end vertex for an edge of G that is not an edge of $G_{\Pi(k)}$.

Indeed, since G is connected, it follows that if some vertex of G does not belong to $G_{\Pi(k)}$, then there exists an edge of G such that one of its endpoints belongs to $G_{\Pi(k)}$ and the other does not.

In case (a), consider the edge e of $G_{\Pi(k)}$ which contains the vertex z. Let z_1 and z_2 be the end vertices of e, and let z' be a vertex of $G_{\Pi(k)}$ different from z_1 and z_2. Since the graph G is 3-connected, it contains a path from z to z' which avoids z_1 and z_2. Therefore, G contains a path from an interior point of e to a point (not necessarily a vertex) of the graph $G_{\Pi(k)}$ such that this path has no common points with $G_{\Pi(k)}$ except its endpoints. For π_{k+1} we take such a path containing a maximum number of vertices of G.

In case (b), any edge of $G_{\Pi(k)}$ is also an edge of G. The graph G contains a path σ from the vertex z to a point of $G_{\Pi(k)}$ which has no other common points with $G_{\Pi(k)}$. The paths $\pi_0, \dots, \pi_k$ were chosen so as to contain a maximum number of vertices of G. Therefore, the path σ goes from z to a vertex of G not neighboring z. For π_{k+1} we take such a path σ containing a maximum number of vertices of the graph G.

In case (b), we add an edge to the graph; this cannot violate its 3-connectedness.

In case (a), we either choose an additional vertex u on one edge and join it by an edge to an existing vertex or take additional vertices u and v on two edges and join them by an edge uv. Clearly, the new graph remains connected after the removal of any two vertices different from u and v. The removal of the vertex u is equivalent to the deletion of the edge containing u from the old graph. Any graph obtained from a 3-connected graph by deleting one edge is at least 2-connected. Therefore, the new graph is 3-connected.

The remaining requirements to the path π_{k+1} are obviously satisfied. □

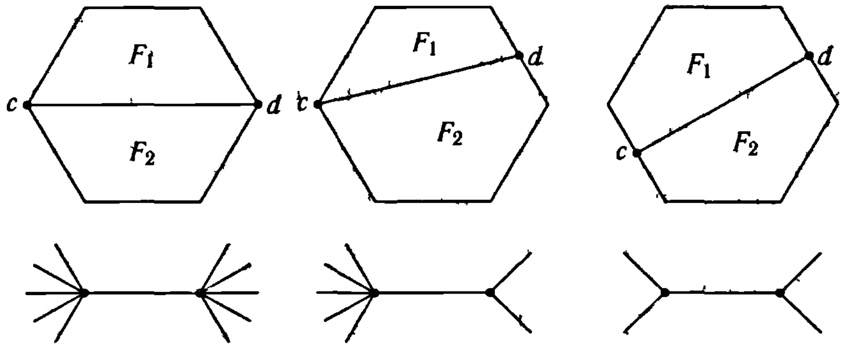

Figure 15. The three cases of edge deletion

After the lemma is proved, the proof of the assertion of Step 1 becomes very simple. Let $\{\pi_0, \dots, \pi_n\}$ be the set of paths constructed for a 3-connected graph G containing more than six edges. This graph is different from K_4; therefore, $n > 1$. The graph G has no vertices of degree 2; hence the path π_n consists of one edge of G. Deleting this edge, we obtain a 3-connected graph $G_{\Pi(n-1)}$, as required. $\square$

Now, we must make the second step, namely, to learn how to construct the convex polyhedron P corresponding to the graph G from the convex polyhedron P' corresponding to the graph G'. In the planar graph G, the edge to be deleted can be of one of the three types shown in Figure 15. These three types of edges deleted from the graph correspond to the three types of edges added to the polyhedron; they also are shown in the figure.

We might try to construct the required transformation of the polyhedron by slightly moving the faces F_1 and F_2 so as to make them noncoplanar (they are coplanar in the initial polyhedron P', while in the polyhedron P, they must be contained in different planes). This involves certain difficulties; namely, if the plane of a face passes through a vertex of an n-hedral angle with $n \geq 4$, then the arbitrary move will destroy the structure of the edge graph of the polyhedron. For example, none of the faces F_1 and F_2 of the polyhedron shown in Figure 16 can be arbitrarily moved, because moving them destroys the structure of the edges going from the vertices A and B. Thus, to achieve the goal, we must slightly move also the vertices A and B. In turn, a small displacement of a vertex may destruct the edge graph if this vertex belongs to a face with more than three sides.

We can try to overcome this difficulty by ordering the vertices and faces in such a way that the sequence of vertices and faces begins with F_1, F_2, c, d and no term of the sequence is incident[5] to more than three preceding

[5]A vertex A is incident to a face F (or a face F is incident to A) if $A \in F$.

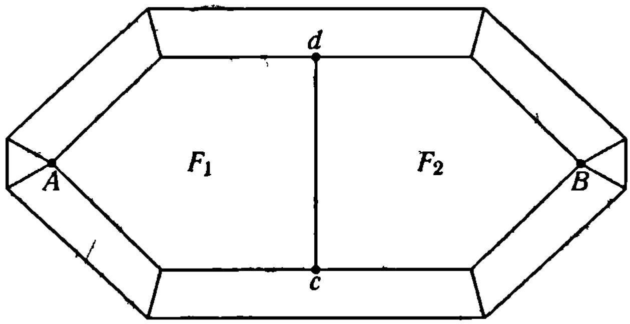

Figure 16. The faces F_1 and F_2 cannot be moved

terms. Indeed, if the vertices and faces can be ordered as described, then we can move the faces F_1 and F_2 and, then, successively move each term of the sequence so that it remain incident to those preceding terms to which it must be incident. If a vertex is incident to three preceding faces, then its position is determined uniquely. If a vertex is incident to $p < 3$ preceding faces, then there are $3 - p$ degrees of freedom in the choice of its position.

Step 2. The set of all vertices and faces of a 3-connected planar graph G can be ordered in such a way that any term of the sequence of vertices and faces is incident to at most three preceding terms. Moreover, for the four initial terms, two faces adjacent to a given edge and two end vertices of this edge can be taken.

First, for the planar graph G, we construct the planar graph $\widetilde{G}$ whose vertex set consists of the vertices of G and the additional vertices corresponding to the faces of G. Two vertices of the graph $\widetilde{G}$ are joined by an edge if they correspond to a vertex and a face incident to each other (see Figure 17).

We must order the vertices of $\widetilde{G}$ in such a way that in the sequence of vertices, each vertex is joined by edges to at most three preceding vertices. Moreover, for the first four vertices we must take the given vertices k_1, k_2, k_3, and k_4, which generate a cycle in the graph $\widetilde{G}$.

All of the faces of $\widetilde{G}$ are tetragonal; therefore, we can apply the corollary of Theorem 1.8 (see p. 17), according to which the graph $\widetilde{G}$ has at least eight vertices of degree 3 (obviously, it has no vertices of degrees 1 and 2). In particular, $\widetilde{G}$ has a vertex of degree 3 different from k_1, k_2, k_3, and k_4. We choose this vertex to be the last term of the sequence and denote it by k_n (here n is the number of vertices in the graph $\widetilde{G}$). Let $K(n)$ be the graph obtained from $\widetilde{G}$ by removing the vertex k_n and the edges incident to it.

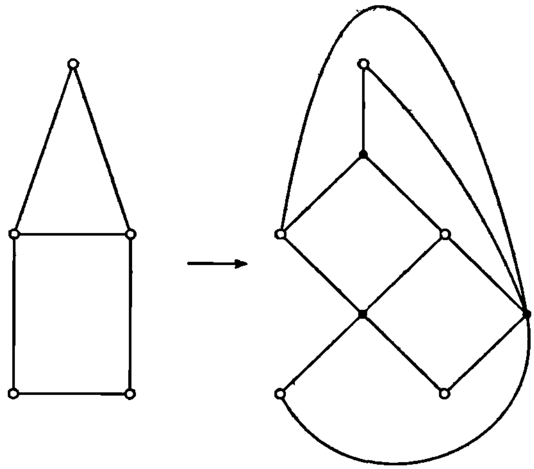

Figure 17. The graph $\widetilde{G}$

Suppose that vertices $k_n, k_{n-1}, \ldots, k_m$ are already chosen and graphs $K(n)$, $K(n-1)$, ..., $K(m)$ are constructed. If $m > 5$, then we must choose a vertex k_{m-1} and construct a graph $K(m-1)$. By assumption, the vertices k_1, k_2, k_3, and k_4 generate a cycle. In particular, the degree of each of these vertices is at least 2. If the graph $K(m)$ contains an isolated vertex or a vertex of degree 1, then we can take such a vertex for k_{m-1}. Suppose that all vertices of $K(m)$ have degrees at least 2. There are two possible cases.

Case 1. The subgraph of $K(m)$ generated by the vertices k_1, k_2, k_3, and k_4 is isolated.

We remove the vertices k_1, k_2, k_3, and k_4 from $K(m)$. To the obtained graph we again apply the corollary of Theorem 1.8, according to which this graph has at least one vertex of degree at most 3. We take this vertex for k_{m-1}.

Case 2. At least one of the vertices k_1, k_2, k_3, and k_4 is joined by an edge to a vertex k_i with $i \geq 5$ in the graph $K(m)$.

In this case, one of the vertices k_1, k_2, k_3, and k_4 has degree at least 3; therefore, the contribution of these vertices to $2v_2 + v_3$ does not exceed 7. This means, in particular, that the graph $K(m)$ has a vertex of degree at most 3 different from the vertices k_1, k_2, k_3, and k_4. We take this vertex for k_{m-1}.

In all cases, the graph $K(m-1)$ is obtained from $K(m)$ by removing the vertex k_{m-1}. $\square$

2. Homotopy Properties of Graphs

2.1. The Fundamental Group of a Graph. In graphs (one-dimensional complexes), many phenomena of homotopic topology manifest themselves; we study them in this section.

Two maps $f_0, f_1: X \to Y$ are said to be *homotopic* if there exists a continuous map $F: X \times [0,1] \to Y$ such that $F(x,0) = f_0(x)$ and $F(x,1) = f_1(x)$. In other words, maps f_0 and f_1 can be joined by a family of continuous maps $f_t: X \to Y$, where $0 \le t \le 1$, continuously depending on t. This family of continuous maps is called a *homotopy* between f_0 and f_1. If the maps f_0 and f_1 are homotopic, we use the notation $f_0 \simeq f_1$.

It is easy to verify that homotopy of maps is an equivalence relation. The implication $f \simeq g$, $g \simeq h \implies f \simeq h$ is proved by using the gluing theorem for continuous maps (Theorem 0.1 on p. 2).

Problem 10. Prove that the maps[6] f, g: $\mathrm{GL}(n,\mathbb{R}) \times \mathrm{GL}(n,\mathbb{R}) \to \mathrm{GL}(2n,\mathbb{R})$ defined by

$$f(A,B) = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \quad \text{and} \quad g(A,B) = \begin{pmatrix} AB & 0 \\ 0 & 1 \end{pmatrix}$$

are homotopic.

A map homotopic to a constant map is said to be *null-homotopic*.

Topological spaces X and Y are said to be *homotopy equivalent* if there exist continuous maps $f: X \to Y$ and $g: Y \to X$ such that the maps $f \circ g$ and $g \circ f$ are homotopic to the identity maps of the spaces Y and X, respectively. The homotopy equivalence between spaces X and Y is denoted by $X \sim Y$.

A topological space is *contractible* if it is homotopy equivalent to a point.

Exercise 9. Prove that the space $\mathbb{R}^n$ is contractible.

A topological space X is said to be *path-connected* if any two points x_0 and x_1 can be joined by a path in this space, i.e., there exists a continuous map f from the interval $I = [0,1]$ to X such that $f(0) = x_0$ and $f(1) = x_1$.

Problem 11. Prove that any path-connected space is connected.

On the set of points of a space X, we can introduce an equivalence relation by declaring two points to be equivalent if they can be joined by a path. The equivalence class of a point $x \in X$ is the maximal path-connected subset containing x. The equivalence classes are called *path-connected components*.

---

[6]On the set of $m \times n$ matrices, the topology is introduced as follows: each matrix is identified with a point of $\mathbb{R}^{mn}$ (or $\mathbb{C}^{mn}$, if the matrix elements are complex) and the induced topology is taken.

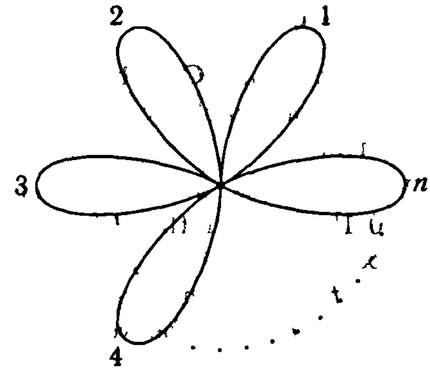

Figure 18. A wedge of circles

Problem 12. Prove that if each point in a connected space X has a path-connected neighborhood, then X is path-connected.

Problem 13. Prove that any connected open subset U of $\mathbb{R}^n$ is path-connected.

Problem 14. Consider the sets

$$X_1 = \{(x,y) : x = 0,\ -1 \le y \le 1\} \quad \text{and} \quad X_2 = \left\{(x,y) : x > 0,\ y = \sin\frac{1}{x}\right\}$$

in $\mathbb{R}^2$. Prove that $X = X_1 \cup X_2$ is connected but not path-connected.

Problem 15. Prove that the following topological spaces of matrices are path-connected: (a) the space of real matrices of order n with positive determinants; (b) the space $\mathrm{SO}(n)$ of orthogonal matrices of order n with determinant 1; (c) the space $\mathrm{U}(n)$ of unitary matrices of order n; (d) the space $\mathrm{SU}(n)$ of unitary matrices of order n with determinant 1.

Let X and Y be disjoint topological spaces in which two points $x_0 \in X$ and $y_0 \in Y$ are marked (in what follows, we refer to marked points of spaces as *base points*). The topological space $X \vee Y = X \cup Y/\{x_0, y_0\}$ is called the *wedge product*, or *wedge*, of the spaces X and Y. In other words, the space $X \vee Y$ is obtained by identifying the base points x_0 and y_0. The wedge $X \vee Y$ can also be defined as the subset of $X \times Y$ consisting of all points (x, y) such that $x = x_0$ or $y = y_0$. Similarly, the wedge of spaces $X_1, \dots, X_n$ with base points $x_1, \dots, x_n$ is defined as $X_1 \vee \dots \vee X_n = X_1 \cup \dots \cup X_n/\{x_1, \dots, x_n\}$. The wedge of n circles is shown in Figure 18.

Theorem 1.15. *Any finite connected one-dimensional complex is homotopy equivalent to a wedge of circles.*

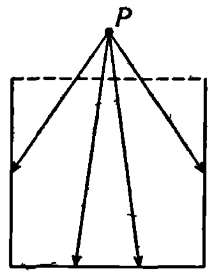

Figure 19. Extension of the map

Proof. Suppose that the end vertices of an edge A of a one-dimensional complex X do not coincide. Then A is a line segment rather that a circle; hence there exists a homotopy $f_t\colon A \to A$ between the identity map $f_0 = \mathrm{id}_A$ and the constant map $f_1\colon A \to * \in A$. Let us prove that the spaces X and X/A are homotopy equivalent in this case. The homotopy $f_t\colon A \to A$ can be extended to a homotopy $F_t\colon X \to X$ with $F_0 = \mathrm{id}_X$. In other words, the map of the set $(A \times I) \cup (X \times \{0\}) \subset X \times I$ can be extended to a map of the entire set $X \times I$. This extension is constructed as follows. Suppose that both end vertices of an edge B belong to the edge A. Then the map is defined on three of the four sides of the square $B \times I$; in Figure 19, these sides are shown by the solid lines, and the fourth side is shown by the dashed line. We send all points of each ray from P to the same point (namely, the image of the intersection point of the ray with one of the three distinguished sides). If one of the endpoints of B does not belong to A (or none of them does), then we define the map on the corresponding lateral side (or on both lateral sides) arbitrarily. Then we construct similar extensions for the edges adjacent to A and B, and so on.

Let $p\colon X \to X/A$ be the natural projection. The map F_1 has the property $F_1(A) = * \in A$. Hence there exists a (unique) map $q\colon X/A \to A$ for which $F_1 = q \circ p$. To prove the homotopy equivalence of the spaces X and X/A, it is sufficient to show that $q \circ p \sim \mathrm{id}_X$ and $p \circ q \sim \mathrm{id}_{X/A}$. By construction, the homotopy F_t joins the maps $F_1 = q \circ p$ and $F_0 = \mathrm{id}_X$. Since $F_t(A) \subset A$ for all t, it follows that $p \circ F_t = q_t \circ p$, where q_t is some homotopy between $q_0 = \mathrm{id}_{X/A}$ and $q_1 = p \circ q$.

Now, we take the one-dimensional complex X/A for X and repeat the procedure. After several repetitions, we obtain a one-dimensional complex which has no edges with noncoinciding end vertices. Such a complex is a wedge of circles. □

It is easy to see that any connected one-dimensional complex containing n_0 vertices and n_1 edges is homotopy equivalent to the wedge of $n_1 - n_0 + 1$

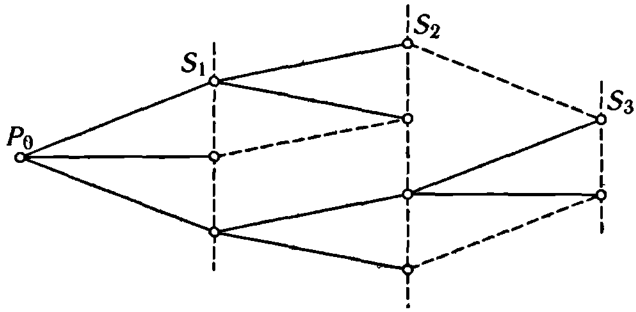

Figure 20. A maximal tree

circles. To show this, we construct a *maximal tree* for this complex, i.e., a contractible subcomplex containing all vertices of the given complex. For this purpose, we take any vertex P_0 and consider the sets S_n, where $n = 1, 2, \dots,$ consisting of the vertices for which the shortest paths to P_0 pass through precisely n edges. We join each vertex from the set S_{n+1} to a vertex from S_n which is joined with the given vertex by an edge (see Figure 20). As a result, we obtain a maximal tree. It contains $n_0 - 1$ edges; successively contracting them, we obtain a one-dimensional complex with one vertex and $n_1 - n_0 + 1$ edges, i.e., the wedge of $n_1 - n_0 + 1$ circles.

An important characteristic of a path-connected topological space X with a base point x_0 is its *fundamental group* $\pi_1(X, x_0)$. The elements of the fundamental group are the classes of homotopic *loops* in X based at x_0, i.e., maps $f: I \to X$ of the interval $I = [0,1]$ with $f(0) = f(1) = x_0$. The group structure on $\pi_1(X, x_0)$ is introduced as follows. We set

$$f_1 f_2(t) = \begin{cases} f_1(2t) & \text{if } 0 \le t \le 1/2, \\ f_2(2t-1) & \text{if } 1/2 \le t \le 1. \end{cases}$$

In other words, we spend half the time passing the loop f_1 at doubled speed, and in the remaining time, we pass the loop f_2 at doubled speed.

The identity element of the fundamental group is the class of the constant map $f: I \to x_0$. The class inverse to the class of a loop $f(t)$ is the class of the loop $g(t) = f(1-t)$. Indeed, the map

$$F_s(t) = \begin{cases} x_0 & \text{if } 0 \le t \le s/2, \\ f(2t-s) & \text{if } s/2 \le t \le 1/2, \\ f(2-2t-s) & \text{if } 1/2 \le t \le 1-s/2, \\ x_0 & \text{if } 1-s/2 \le t \le 1 \end{cases}$$

(see Figure 21) is a homotopy between $F_0 = fg$ and $F_1: I \to x_0$.

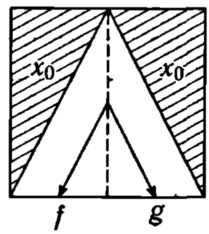

Figure 21. An inverse element in the fundamental group

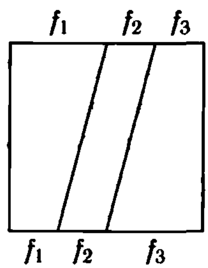

Figure 22. Associativity of multiplication

Using Figure 22, it is easy to construct a homotopy between the maps $f_1(f_2f_3)$ and $(f_1f_2)f_3$.

Let α be a path in X from x_1 to x_2, and let f be a loop based at x_1. Then $\alpha^{-1}f\alpha$ is a loop based at x_2. It is easy to verify that the map $f \mapsto \alpha^{-1}f\alpha$ induces an isomorphism between the groups $\pi_1(X, x_1)$ and $\pi_1(X, x_2)$. Paths α and β induce the same isomorphisms if and only if the class of the loop $\alpha\beta^{-1}$ belongs to the center of the group $\pi_1(X, x_1)$. Indeed, the loops $\alpha^{-1}f\alpha$ and $\beta^{-1}f\beta$ are homotopic if an only if the loops $f(\alpha\beta^{-1})$ and $(\alpha\beta^{-1})f$ are homotopic.

A path-connected space X is said to be *simply connected* if $\pi_1(X, x_0) = 0$ for some point $x_0 \in X$; in this case, $\pi_1(X, x_1) = 0$ for any point $x_1 \in X$.

A continuous map $f: X \to Y$ naturally induces a homomorphism f_* : $\pi_1(X, x_0) \to \pi_1(Y, y_0)$, where $y_0 = f(x_0)$. Under this homomorphism, the class of any loop $\omega(t)$ corresponds to the class of the loop $f(\omega(t))$. Clearly, $(fg)_* = f_*g_*$.

Theorem 1.16. *Let f_t be a homotopy between maps $f_0, f_1: X \to Y$. Then the homomorphism $(f_1)_*: \pi_1(X, x_0) \to \pi_1(Y, f_1(x_0))$ coincides with the composition of the homomorphism $(f_0)_*: \pi_1(X, x_0) \to \pi_1(Y, f_0(x_0))$ and the isomorphism $\pi_1(Y, f_0(x_0)) \to \pi_1(Y, f_1(x_0))$ determined by the path $\alpha(t) = f_t(x_0)$, which joins the points $f_0(x_0)$ and $f_1(x_0)$.*

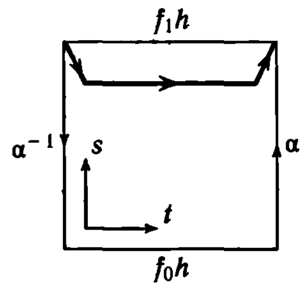

Figure 23. The homotopy

Proof. Let h be a loop in X based at x_0. We must prove that the loops $f_1(h(t))$ and $\alpha^{-1}f_0(h(t))\alpha$ are homotopic. Consider the map $F: I \times I \to Y$ defined by $F(s,t) = f_s(h(t))$. Figure 23 shows one of the paths forming a homotopy between the loops $f_1 h$ and $\alpha^{-1}(f_0 h)\alpha$. □

Theorem 1.17. *The fundamental groups of homotopy equivalent path-connected topological spaces are isomorphic.*

Proof. Let X and Y be homotopy equivalent path-connected topological spaces. Then there exist maps $f: X \to Y$ and $g: Y \to X$ such that $fg \sim \mathrm{id}_Y$ and $gf \sim \mathrm{id}_X$. According to Theorem 1.16, the homomorphisms $g_* f_*: \pi_1(X, x_0) \to \pi_1(X, gf(x_0))$ and $f_* g_*: \pi_1(Y, y_0) \to \pi_1(Y, fg(y_0))$ are compositions of the identity map and an isomorphism; thus, they are isomorphisms. Consider the homomorphisms

$$\pi_1(X, x_0) \xrightarrow{f_*^{(1)}} \pi_1(Y, f(x_0)) \xrightarrow{g_*} \pi_1(X, gf(x_0)) \xrightarrow{f_*^{(2)}} \pi_1(Y, fgf(x_0)).$$

(Here $f_*^{(1)}$ and $f_*^{(2)}$ are the homomorphisms of fundamental groups with different base points induced by the same map f.) The homomorphism $g_* f_*^{(1)}$ is an isomorphism; therefore, g_* is an epimorphism. The homomorphism $f_*^{(2)} g_*$ is an isomorphism; therefore, g_* is a monomorphism. Thus, g_* is an isomorphism. □

It follows from Theorems 1.15 and 1.17 that the fundamental group of any connected one-dimensional complex is isomorphic to the fundamental group of a wedge of circles. To be more precise, the fundamental group of a connected one-dimensional complex with n_0 vertices and n_1 edges is isomorphic to the fundamental group of the wedge of $n_1 - n_0 + 1$ circles.

2.2. Coverings of One-Dimensional Complexes. Let $\widetilde{X}$ and X be path-connected topological spaces (say, connected one-dimensional complexes). A map $p: \widetilde{X} \to X$ is called a *covering* if $p(\widetilde{X}) = X$ and each

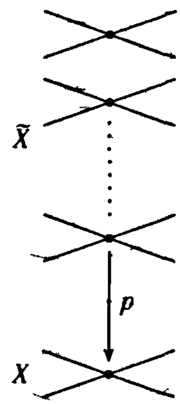

Figure 24. A covering of a one-dimensional complex

point $x \in X$ has a neighborhood U such that the preimage $p^{-1}(U)$ of this neighborhood is homeomorphic to $U \times D$, where D is a discrete set, and the restriction of p to $p^{-1}(U)$ coincides with the natural projection $U \times D \to U$ (see Figure 24). The space $\widetilde{X}$ is called the *covering space*, and X is the *base* of the covering. It can be shown by using the path-connectedness of X that the sets D have the same cardinality for all points $x \in X$ (this is essentially proved below in constructing a lifting of a path). If the discrete set D contains precisely n points, then we say that the covering is *n-fold*; sometimes (especially in algebraic geometry and complex analysis), the number n is referred to as the *degree* of the covering. The preimage of a point $x_0 \in X$ is called the *fiber* over the point x_0. Each fiber of an n-fold covering contains precisely n points.

Problem 16. (a) Let K_n be the complete graph on $n > 3$ vertices, and let $p\colon K_n \to G$ be a covering. Prove that this covering is odd-fold.

(b) Prove that there exists a covering $p\colon K_n \to G$ of any odd multiplicity.

In this section, we consider only coverings of one-dimensional complexes.

The real line $\mathbb{R}$ can be treated as a one-dimensional complex with vertices at the integer points. The map $\exp\colon \mathbb{R} \to S^1$ that takes each point $t \in \mathbb{R}$ to the point $\exp(2\pi i t) \in S^1$ is a covering (see Figure 25).

A *lifting* of a path $\gamma(t) \subset X$ is a path $\widetilde{\gamma}(t) \subset \widetilde{X}$ such that $p(\widetilde{\gamma}(t)) = \gamma(t)$ for all t. If x_0 is the starting point of the path $\gamma(t)$ and $\tilde{x}_1 \in p^{-1}(x_0)$, then there exists a unique lifting of $\gamma(t)$ which starts at $\tilde{x}_1$. The example of the map exp shows that the lifting of a closed path is not necessarily closed (see Figure 26). Any covering $p\colon \widetilde{X} \to X$ induces a homomorphism $p_*\colon \pi_1(\widetilde{X}, \tilde{x}_0) \to \pi_1(X, x_0)$, where $x_0 = p(\tilde{x}_0)$. The class of a loop $\gamma(t) \subset X$ based at a point x_0 belongs to the subgroup $p_*\pi_1(\widetilde{X}, \tilde{x}_0) \subset \pi_1(X, x_0)$ if and

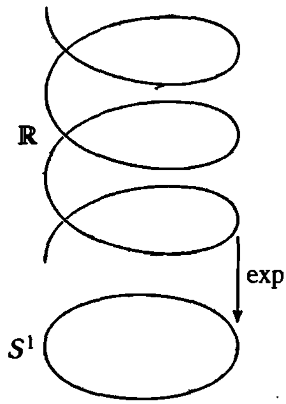

Figure 25. The exponential covering of the circle

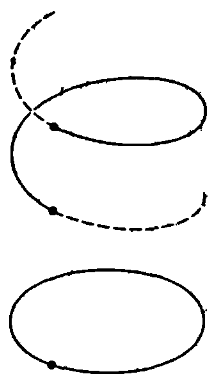

Figure 26. A nonclosed lifting of a closed path

only if the lifting of this loop starting at $\tilde{x}_0$ is closed. For a different point $\tilde{x}_1$ from the preimage of x_0, the groups $G_0 = p_*\pi_1(\widetilde{X}, \tilde{x}_0)$ and $G_1 = p_*\pi_1(\widetilde{X}, \tilde{x}_1)$ do not necessarily coincide. Indeed, we have $G_1 = \alpha^{-1}G_0\alpha$, where α is the projection of a path joining the points $\tilde{x}_0$ and $\tilde{x}_1$ in $\widetilde{X}$. The groups G_0 and G_1 coincide if and only if the liftings of any loop γ which start at $\tilde{x}_0$ and at $\tilde{x}_1$ are closed or nonclosed simultaneously. It is also clear that any lifting of a loop γ based at x_0 joins some points from the preimage of x_0. Thus, for a loop γ at x_0, its liftings starting at different points of the preimage of x_0 are closed or nonclosed simultaneously only if $\alpha^{-1}G_0\alpha = G_0$ for all $\alpha \in \pi_1(X, x_0)$, i.e., if $p_*\pi_1(\widetilde{X}, \tilde{x}_0)$ is a normal subgroup in $\pi_1(X, x_0)$. In this case, the covering p is said to be *regular*. An example of an irregular covering is given in Figure 27. Figure 28 shows the same covering differently.

Now we study the homomorphism $p_*\colon \pi_1(\widetilde{X}, \tilde{x}_0) \to \pi_1(X, x_0)$ in more detail. First, we show that p_* is a monomorphism. To this end, we must

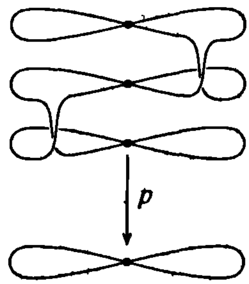

Figure 27. An irregular covering

Figure 28. The same irregular covering

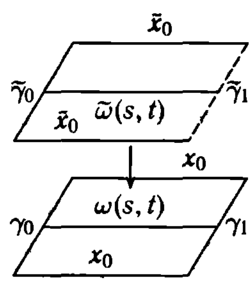

Figure 29. The lifting of a homotopy

verify that if the projections γ_0 and γ_1 of loops $\tilde{\gamma}_0$ and $\tilde{\gamma}_1$ based at $\tilde{x}_0$ are homotopic, then the loops $\tilde{\gamma}_0$ and $\tilde{\gamma}_1$ are homotopic as well. Let $\gamma_s(t)$ be a homotopy between γ_0 and γ_1. For fixed $t = t_0$, $\omega(s, t_0) = \gamma_s(t_0)$ is a path joining the points $\gamma_0(t_0)$ and $\gamma_1(t_0)$. Consider its lifting $\tilde{\omega}(s, t_0)$ starting at $\tilde{\gamma}_0(t_0)$ (see Figure 29).

Clearly, $p\tilde{\omega}(s,t) = \omega(s,t)$ and $\tilde{\omega}(s,0) = \tilde{\omega}(s,1) = \tilde{x}_0$. Therefore, if $\tilde{\omega}(s,t)$ continuously depends on t, then $\tilde{\gamma}_s(t) = \tilde{\omega}(s,t)$ is a homotopy between $\tilde{\gamma}_0$ and $\tilde{\gamma}_1$. Let s_0 be the least upper bound of all s such that $\tilde{\omega}(s,t)$ continuously depends on t. (Note that $\tilde{\omega}(0,t) = \tilde{\gamma}_0(t)$ continuously depends on t.) Consider the neighborhood U of $\omega(s_0,t)$ from the definition of a covering and choose $\varepsilon > 0$ such that $\omega(s,\tau) \in U$ for $|s - s_0| < \varepsilon$ and $|t - \tau| < \varepsilon$. Using the compactness of the interval, we can choose the same ε for all t. Take $t_0 \in [0,1]$. All points $\tilde{\omega}(s_0 - \varepsilon, t)$, where $t \in [t_0 - \varepsilon, t_0 + \varepsilon]$, belong to the same copy of the neighborhood U in $p^{-1}(U) \approx U \times D$. Therefore, for any $s \in [s_0, s_0 + \varepsilon]$, all points $\tilde{\omega}(s,t)$, where $t \in [t_0 - \varepsilon, t_0 + \varepsilon]$, belong to the same copy of U. Hence $\tilde{\omega}(s,t)$ continuously depends on t for $s \in [s_0, s_0+\varepsilon]$. Thus, $s_0 = 1$, which implies the required assertion.

Consider the subgroup $H \rightleftharpoons p_*\pi_1(\widetilde{X}, \tilde{x}_0) \subset \pi_1(X, x_0) = G$ and the right cosets Hg_i, where $g_i \in G$. Two cosets Hg_1 and Hg_2 coincide if $g_1 g_2^{-1} \in H$, and they are disjoint if $g_1 g_2^{-1} \notin H$. There is a natural one-to-one correspondence between the right cosets Hg_i and the points of $p^{-1}(x_0)$. To establish it, we use the base point $\tilde{x}_0$ in $p^{-1}(x_0)$. To each loop γ in X based at x_0 we assign the endpoint of the lifting of this loop starting at $\tilde{x}_0$. As a result, we obtain a map $G \to p^{-1}(x_0)$. We show that this map is a one-to-one correspondence between the right cosets and the set $p^{-1}(x_0)$. Let $\widetilde{\gamma}_1$ and $\widetilde{\gamma}_2$ be the liftings of γ_1 and γ_2 starting at $\tilde{x}_0$. The end of the path $\widetilde{\gamma}_1$ coincides with the end of $\widetilde{\gamma}_2$ if and only if $\widetilde{\gamma}_1\widetilde{\gamma}_2^{-1}$ is a closed path at x_0, i.e., $\widetilde{\gamma}_1\widetilde{\gamma}_2^{-1} \in H$. It remains to note that the image of the map $G \to p^{-1}(x_0)$ under consideration is the entire set $p^{-1}(x_0)$. Indeed, each point $\tilde{x}_1 \in p^{-1}(x_0)$ is the image of the element of $\pi_1(X, x_0)$ corresponding to the projection of a path in $\widetilde{X}$ starting at $\tilde{x}_0$ and ending at $\tilde{x}_1$; the projection of this path is a loop at x_0 in X. Thus, we have proved the following assertion.

Theorem 1.18. *If $p\colon \widetilde{X} \to X$ is a covering and $p(\tilde{x}_0) = x_0$, then there exists a one-to-one correspondence between the cosets $\pi_1(X, x_0)/p_*\pi_1(\widetilde{X}, \tilde{x}_0)$ and the fiber $p^{-1}(x_0)$.*

The coset space does not generally have a natural structure of a group. Indeed, if the product of classes Hg and Hg^{-1} is uniquely determined, then the equality $HgHg^{-1} = H$ (i.e., $gHg^{-1} = H$) must hold for all $g \in G$. This means that H must be a normal subgroup in G, i.e., p must be a regular covering. (Clearly, if H is a normal subgroup, then $Hg_1Hg_2 = Hg_1g_2$ because $g_1H = Hg_1$.)

Thus, if a covering p is regular, then the set G/H, which is in one-to-one correspondence with the set $p^{-1}(x_0)$, has the natural structure of a group. In this case, we can define a group structure on $p^{-1}(x_0)$ by fixing a point $\tilde{x}_0 \in p^{-1}(x_0)$. This group admits a more geometric description than the quotient group $\pi_1(X, x_0)/p_*\pi_1(\widetilde{X}, \tilde{x}_0)$. The point is that for regular coverings, we can insert the intermediate group $\mathrm{Aut}(p)$ in the correspondence $G/H \leftrightarrow p^{-1}(x_0)$:

$$G/H \leftrightarrow \mathrm{Aut}(p) \leftrightarrow p^{-1}(x_0).$$

Here $\mathrm{Aut}(p)$ is the *automorphism group* of the covering p; it is defined as follows.

A homeomorphism $f\colon \widetilde{X} \to \widetilde{X}$ is called an *automorphism* of a covering $p\colon \widetilde{X} \to X$ if $p(f(\tilde{x})) = p(\tilde{x})$ for all $\tilde{x} \in \widetilde{X}$. If $\tilde{y} = f(\tilde{x})$, then $p(\tilde{y}) = p(f(\tilde{x})) = p(\tilde{x})$; therefore, any automorphism of p permutes the points of each fiber.

Theorem 1.19. *Any automorphism of a covering is completely determined by the image of one point under this automorphism.*

Proof. We show that for a covering $p\colon \widetilde{X} \to X$, there exists at most one automorphism $f\colon \widetilde{X} \to \widetilde{X}$ mapping the point $\tilde{x}_0 \in \widetilde{X}$ to a given point $\tilde{x}_1 \in \widetilde{X}$. Take any point $\tilde{y}_0 \in \widetilde{X}$. Consider a path $\widetilde{\gamma}_0$ joining $\tilde{x}_0$ to $\tilde{y}_0$, Let $\gamma = p\widetilde{\gamma}_0$ be the projection of $\widetilde{\gamma}_0$, and let $\widetilde{\gamma}_1$ be the lifting of γ starting at $\tilde{x}_1$. Then the automorphism f takes the path $\widetilde{\gamma}_0$ to the path $\widetilde{\gamma}_1$; hence $f(\tilde{y}_0) = \tilde{y}_1$. Thus, the automorphism f is determined uniquely. Clearly, an automorphism f mapping $\tilde{x}_0$ to $\tilde{x}_1$ exists if and only if the point $\tilde{y}_1$ is uniquely determined by $\tilde{y}_0$; hence the lifting starting at $\tilde{x}_1$ of the projection of any closed path based at $\tilde{x}_0$ is closed. □

Exercise 10. Prove that each automorphism of the covering shown in Figure 27 is the identity automorphism.

Theorem 1.20. (a) *A covering $p\colon \widetilde{X} \to X$ is regular if and only if the group* $\operatorname{Aut}(p)$ *acts transitively on the fiber $p^{-1}(x_0)$, i.e., every element of this fiber is mapped to every other element by some automorphism.*

(b) *For a regular covering $p\colon \widetilde{X} \to X$, the group* $\operatorname{Aut}(p)$ *is isomorphic to $\pi_1(X, x_0)/p_*\pi_1(\widetilde{X}, \tilde{x}_0)$.*

Proof. (a) Suppose that p is a regular covering and $\tilde{x}_1, \tilde{x}_2 \in p^{-1}(x_0)$. We construct an automorphism $g \in \operatorname{Aut}(p)$ mapping $\tilde{x}_1$ to $\tilde{x}_2$. Take any point $\tilde{y}_1 \in \widetilde{X}$. Suppose that $\widetilde{\gamma}_1$ is an arbitrary path from $\tilde{x}_1$ to $\tilde{y}_1$, $\gamma = p\widetilde{\gamma}_1$ is the projection of $\widetilde{\gamma}_1$, and $\widetilde{\gamma}_2$ is the lifting of γ starting at $\tilde{x}_2$. We set $g(\tilde{y}_1) = \tilde{y}_2$, where $\tilde{y}_2$ is the end of the path $\widetilde{\gamma}_2$. The map g is well defined, i.e., $\tilde{y}_2$ does not depend on the choice of $\widetilde{\gamma}_1$. Indeed, since the covering p is regular, it follows that if a path $\widetilde{\gamma}_1\widetilde{\gamma}_1'$ is closed, then any lifting of $p(\widetilde{\gamma}_1\widetilde{\gamma}_1')$ is closed as well.

Now suppose that the group $\operatorname{Aut}(p)$ acts transitively on the fiber $p^{-1}(x_0)$. Let ω be a closed path based at $\tilde{x}_1 \in p^{-1}(x_0)$, and let g be an automorphism taking $\tilde{x}_1$ to $\tilde{x}_2$. Then $g\omega$ is the lifting of $p\omega$ starting at $\tilde{x}_2$. Clearly, the path $g\omega$ is closed.

(b) Let α be a loop in X based at x_0, and let $[\alpha] \in \pi_1(X, x_0)$ be the class of homotopic loops containing α. To the class $[\alpha]$ we assign the automorphism g_α of p defined as follows. Suppose that $\tilde{x}_0 \in p^{-1}(x_0)$ is the fixed point in the fiber and $\tilde{y}_0 \in \widetilde{X}$ is arbitrary. We join $\tilde{x}_0$ to $\tilde{y}_0$ by a path $\widetilde{\gamma}$. Consider the path $\gamma = p\widetilde{\gamma}$. We set $g_\alpha(\tilde{y}_0) = \tilde{y}_1$, where $\tilde{y}_1$ is the end of the lifting of $\gamma\alpha$ starting at $\tilde{x}_0$.

The kernel of the homomorphism $\pi_1(X, x_0) \to \operatorname{Aut}(p)$ is the subgroup $p_*\pi_1(\widetilde{X}, \tilde{x}_0)$. This homomorphism is an epimorphism. Indeed, take a point $\tilde{x}_i \in p^{-1}(x_0)$ and consider the projection α_i of a path from $\tilde{x}_0$ to $\tilde{x}_i$. The loop α_i corresponds to an automorphism of p which maps $\tilde{x}_0$ to $\tilde{x}_i$. But such an automorphism is unique. □

Corollary 1. *If $p\colon \widetilde{X} \to X$ is a covering and $\pi_1(\widetilde{X}) = 0$, then* $\mathrm{Aut}(p) \cong \pi_1(X)$.

Corollary 2. *If $p\colon \widetilde{X} \to X$ is a regular covering and $A = \mathrm{Aut}(p)$, then $X = \widetilde{X}/A$ and the covering has the form $p\colon \widetilde{X} \to \widetilde{X}/A$.*

Problem 17. Prove that a map $f\colon S^1 \to S^1$ is null-homotopic if and only if f can be represented as $f = f_1 f_2$, where $f_1\colon \mathbb{R} \to S^1$ and $f_2\colon S^1 \to \mathbb{R}$.

2.3. Coverings and Fundamental Groups. Coverings can be used to calculate the fundamental group of any one-dimensional complex. We start with calculating the fundamental group of the circle S^1.

Theorem 1.21. $\pi_1(S^1) = \mathbb{Z}$.

Proof. Consider the exponential covering $p\colon \mathbb{R} \to S^1$ which takes every point $t \in \mathbb{R}$ to the point $\exp(it) \in S^1$. The covering space $\mathbb{R}$ is contractible; therefore, $\pi_1(\mathbb{R}) = 0$. By Corollary 1 of Theorem 1.20, the group $\pi_1(S^1)$ is isomorphic to the automorphism group of the covering p.

Any automorphism $g \in \mathrm{Aut}(p)$ is uniquely determined by its action on the element $0 \in \mathbb{R}$. Clearly, $g(0) = 2\pi n_g$, where $n_g \in \mathbb{Z}$. Moreover, $g(t) = t + 2\pi n_g$, and hence $hg(t) = t + 2\pi(n_h + n_g)$. Thus, $\mathrm{Aut}(p) \cong \mathbb{Z}$. Each integer n corresponds to the automorphism $t \mapsto t + 2\pi n$, and this automorphism corresponds to the loop traversing the circle S^1 n times. □

We have already proved that the fundamental group of a connected one-dimensional complex is isomorphic to the fundamental group of some wedge of circles (see p. 34). It remains to calculate the fundamental group of a wedge of circles. Recall that the *free group* of rank n is the group F_n with generators $a_1, \ldots, a_n$ and no relations; i.e., in the group F_n, any irreducible word of the form $a_{i_1}^{\varepsilon_1} \cdots a_{i_k}^{\varepsilon_k}$, where $\varepsilon_l = \pm 1$, is different from the identity element ("irreducible" means that the word contains no fragments of the form $a_i^{\varepsilon} a_i^{-\varepsilon}$).

Theorem 1.22. *The fundamental group of the wedge of n circles is isomorphic to the free group on n generators.*

First proof. Let $\alpha_1, \ldots, \alpha_n$ be the elements of the group $G = \pi_1(\bigvee_{i=1}^n S_i^1)$ that correspond to the single traverses of the circles $S_1^1, \ldots, S_n^1$. Clearly, $\alpha_1, \ldots, \alpha_n$ generate the group G. We must only verify that there are no relations between them. It suffices to prove that the lifting of any irreducible loop $\alpha_{i_1}^{\varepsilon_1} \cdots \alpha_{i_k}^{\varepsilon_k}$ under some covering is a nonclosed path. There exists a covering $T_n \to \bigvee_{i=1}^n S_i^1$ of the wedge of circles with contractible covering

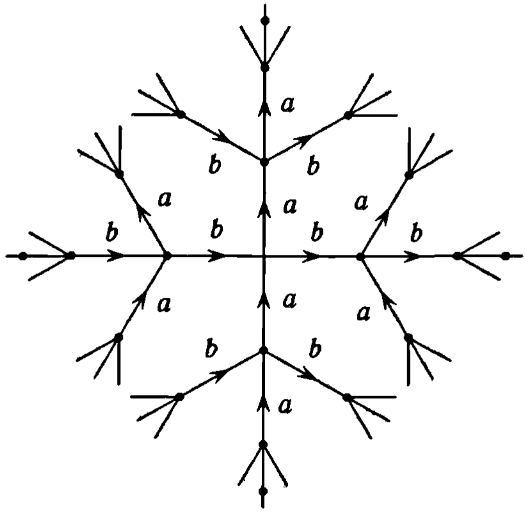

Figure 30. The universal covering of the wedge of two circles

space T_n; the structure of the covering space T_n for $n = 2$ is seen from Figure 30. For $n = 2$, this covering is as follows. Let a and b be the circles from the wedge with orientations. Each edge of the graph T_2 is labeled by a or b, and it is endowed with a direction. Every vertex of the graph T_2 is incident to two edges labeled by a and two edges labeled by b; one edge of each type is incoming and the other is outgoing. The edges labeled by a are mapped onto the circle a in accordance with their directions. On the edges labeled by b the map is similar.

The graph T_n contains no loops; therefore, the lifting of an irreducible loop $\alpha_{i_1}^{\varepsilon_1} \cdots \alpha_{i_k}^{\varepsilon_k}$ cannot be a closed path. $\square$

Second proof. Suppose that $\alpha_{i_1}^{\varepsilon_1} \cdots \alpha_{i_k}^{\varepsilon_k}$ is an irreducible loop. We construct a $(k + 1)$-fold covering for which one of the liftings of this loop is nonclosed.

Take $k + 1$ copies of the wedge of n circles arranged above one another. Each copy of the wedge is identically mapped to the wedge. We remove two equal small arcs from the i_1th circles in the two lowest copies of the wedge and join the endpoints of the remaining arcs criss-cross (such a construction was used to obtain the covering in Figure 27). If $i_1 \neq i_2$, then we apply the same procedure to the i_2th circles in the second and third (from below) copies of the wedge. If $i_1 = i_2$, then we remove one more arc from the second copy of the i_1th circle in such a way that the lifting of the loop $\alpha_{i_1}^{2\varepsilon_1}$ ascends from the first copy to the third. Namely, if $\varepsilon_1 = 1$, then, at the second step, we remove the arc which is passed after the arc removed at the first

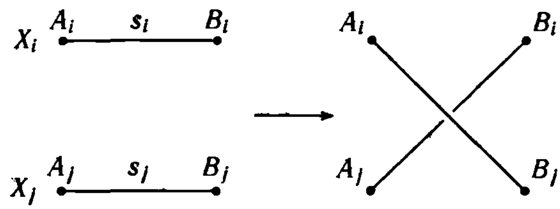

Figure 31. Reconstruction of a graph

step in traversing the i_1th circle in the positive direction, and if $\varepsilon_1 = -1$, then we remove the arc passed before the removed arc. In the case where the expression for the loop contains several identical consecutive letters, the procedure is similar.

As a result, we obtain a $(k+1)$-fold covering such that the lifting of the irreducible loop $\alpha_{i_1}^{\varepsilon_1} \cdots \alpha_{i_k}^{\varepsilon_k}$ ascends from the very bottom copy of the wedge to the very top one. $\square$

If $p\colon \widetilde{X} \to X$ is a covering, then the map $p_*\colon \pi_1(\widetilde{X}, \tilde{x}_0) \to \pi_1(X, x_0)$ is a monomorphism (see p. 36). This means that the fundamental group of the covering space $\widetilde{X}$ is isomorphic to a subgroup of the fundamental group of the base X. We show that each subgroup of the fundamental group of the base corresponds to some covering.

Theorem 1.23. *Let X be a one-dimensional complex and let $G = \pi_1(X, x_0)$. Then, for any subgroup $H \subset G$, there exists a covering $p\colon \widetilde{X} \to X$ such that $p_*\pi_1(\widetilde{X}, \tilde{x}_0) = H$.*

Proof. We say that two loops γ_1 and γ_2 based at x_0 are equivalent if the homotopy class of the loop $\gamma_1\gamma_2^{-1}$ belongs to H. Let U be the set of all loops whose homotopy classes belong to H, and let $U_1 = U, \ldots, U_i, \ldots$ be the equivalence classes of loops. For each equivalence class, consider a copy X_i of the complex X. Choose a maximal tree T in X; we denote its copy in X_i by T_i. We leave the edges of the trees T_i unchanged and reconstruct the remaining edges of the complexes X_i as follows. Let s be a directed edge of X not contained in the maximal tree T; it corresponds to an element $\hat{s} \in \pi_1(X, x_0)$. If $U_i\hat{s} = U_j$, then we replace the edge s_i with end vertices A_i and B_i and the edge s_j with end vertices A_j and B_j by the edges A_iB_j and A_jB_i, respectively (see Figure 31). After performing all such replacements in the complexes X_i, we obtain a complex $\widetilde{X}$ for which a natural covering $p\colon \widetilde{X} \to X$ is defined. We show that $\widetilde{X}$ is connected and $p_*\pi_1(\widetilde{X}, \tilde{x}_0) = H$.

The complex $\widetilde{X}$ is connected because, for any two classes U_i and U_j, there exists a loop γ_{ij} such that $U_i\gamma_{ij} = U_j$. We prove that $p_*\pi_1(\widetilde{X}, \tilde{x}_0) = H$. For definiteness, we assume that the point $\tilde{x}_0$ belongs to the complex X_1. A loop $e_1 \dots e_n$, where $e_1, \dots, e_n$ are edges of X, corresponds to a class of homotopic loops from the subgroup $p_*\pi_1(\widetilde{X}, \tilde{x}_0)$ if and only if its lifting starting at $\tilde{x}_0$ is closed. On the other hand, the end of the lifting (starting at $\tilde{x}_0$) of the loop $e_1 \dots e_n$ belongs to the complex corresponding to the class $Ue_1 \dots e_n$. This lifting is closed if and only if $Ue_1 \dots e_n = U$, i.e., the homotopy class of the loop $e_1 \dots e_n$ belongs to H. □

The subgroups of the fundamental group $G = \pi_1(X, x_0)$ are partially ordered by inclusion: some of these subgroups are contained in some other subgroups. The covering spaces of X are partially ordered too: some of them cover other covering spaces. These two partial orderings are related to each other.

Theorem 1.24. *Let X be a one-dimensional complex and $G = \pi_1(X, x_0)$. Suppose that $p_i\colon \widetilde{X}_i \to X$ $(i = 1, 2)$ are the coverings corresponding to subgroups $H_i \subset G$ (here $H_i = (p_i)_*\pi_1(\widetilde{X}_i, \tilde{x}_i)$ and $p_i(\tilde{x}_i) = x_0$). Then a covering $p\colon \widetilde{X}_1 \to \widetilde{X}_2$ with $p(\tilde{x}_1) = \tilde{x}_2$ and $p_2 p = p_1$ exists if and only if $H_1 \subset H_2$.*

Proof. If $p_1 = p_2 p$, then the image of the map $(p_1)_*$ is contained in the image of $(p_2)_*$, i.e., $H_1 \subset H_2$. Now, suppose that $H_1 \subset H_2$. Take an arbitrary point $\tilde{y}_1 \in \widetilde{X}_1$ and a path $\tilde{\gamma}_1$ from $\tilde{x}_1$ to $\tilde{y}_1$ and consider the projection $\gamma = p_1\tilde{\gamma}_1$ of $\tilde{\gamma}_1$. We set $p(\tilde{y}_1) = \tilde{y}_2$, where $\tilde{y}_2$ is the end of the lifting of γ starting at $\tilde{x}_2$. The map p is well defined if and only if the path $\tilde{\gamma}_2$ is closed whenever $\tilde{\gamma}_1$ is closed. This means that if the class of the loop γ belongs to H_1, then it also belongs to H_2. This condition does hold; therefore, the map p is well defined. □

Corollary. *If $H_1 = H_2$, then the one-dimensional complexes $\widetilde{X}_1$ and $\widetilde{X}_2$ are homeomorphic.*

Proof. There is a bijection between the preimage of any point under the map p and the coset space H_2/H_1. If $H_1 = H_2$, then the map p is one-to-one. □

If $H_1 = 0$, then the space $\widetilde{X}_1$ covers any space covering X. For this reason, a covering space with trivial fundamental group, as well as the covering $p\colon \widetilde{X}_1 \to X$, is said to be *universal*. For any one-dimensional complex, the universal covering space exists and is determined uniquely up to homeomorphism; the universal covering space of a one-dimensional complex is always a tree.

Let $R = \{r_1, \dots, r_m\}$ be a set of elements of the free group F_n on generators $a_1, \dots, a_n$, and let N be the minimal normal subgroup containing

R, i.e., the intersection of all normal subgroups containing R. The group $G = F_n/N$ is referred to as the group *defined by the generators* $a_1, \dots, a_n$ *and relations* $r_1, \dots, r_m$.

Theorem 1.25. *Let G be a group defined by n generators and m relations. Then there exists a regular covering of the wedge of n circles such that its automorphism group is isomorphic to G.*

Proof. The fundamental group of the wedge of n circles is isomorphic to the free group F_n. According to Theorem 1.23, there exists a covering under which the image of the fundamental group of the covering space in the fundamental group of the base coincides with the subgroup $N \subset F_n$. The subgroup N is normal; therefore, the covering is regular, and hence its automorphism group is isomorphic to $F_n/N = G$. $\square$

Problem 18. Construct regular coverings of the wedge of two circles with the following automorphism groups: (a) $\mathbb{Z}$; (b) $\mathbb{Z}\oplus\mathbb{Z}$; (c) $\mathbb{Z}_n$; and (d) $\mathbb{Z}_2\oplus\mathbb{Z}_3$.

Coverings of one-dimensional complexes can be used to prove various properties of free groups. Examples of such properties are given in the following two problems.

Problem 19. (a) Prove that any subgroup of a free group G is free.

(b) Prove that if H is a subgroup of a free group G and the index $[G : H] = k$ is finite, then $\operatorname{rk} H = (\operatorname{rk} G - 1)k + 1$.

Problem 20. Prove that the free group of rank 2 contains the free group of any rank n (including $n = \infty$) as a subgroup.

It is convenient to construct the universal covering of a graph G (which may contain multiple edges and loops) by using the matrix $R(G)$ defined as follows. We start with partitioning the vertices of G into sets $V_1, \dots, V_n$ so that the number of edges joining each vertex $v \in V_i$ to vertices from V_j be the same for all $v \in V_i$ (although, it may depend on $j = 1, \dots, n$); we assume that a loop with vertex $v \in V_i$ corresponds to two edges joining v to vertices from V_i. Such a partition can be constructed as follows. At the first step, we sort vertices by their degrees. Then, we refine the partition $V_1', \dots, V_k'$ thus obtained by sorting the vertices from each V_i' according to the numbers of edges joining these vertices to vertices from V_j'. The second step is repeated until the process stabilizes.

By definition, the matrix $R(G)$ has size $n \times n$; its element r_{ij} is equal to the number of edges joining each vertex $v \in V_i$ to vertices from the set V_j.

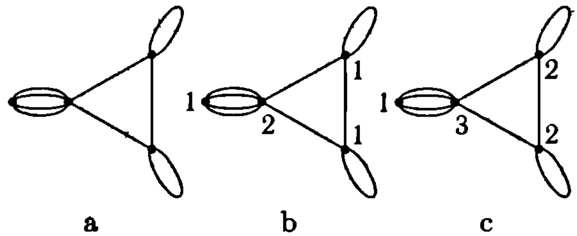

Figure 32. Calculation of the matrix $R(G)$

Example. For the graph shown in Figure 32, we obtain two sets of vertices at the first step and three sets of vertices at the second step. For this graph,

$$R(G) = \begin{pmatrix} 0 & 0 & 4 \\ 0 & 3 & 1 \\ 4 & 2 & 0 \end{pmatrix}.$$

Theorem 1.26. (a) *If a graph $\widetilde{G}$ covers a graph G, then $R(\widetilde{G}) = R(G)$; to be more precise, the matrices coincide up to a numbering of the sets into which the vertices are divided.*

(b) *The universal covering space of any graph G is uniquely determined by the matrix $R(G)$.*

Proof. (a) Sets $\widetilde{V}_1, \ldots, \widetilde{V}_n$ form the required partition of the vertex set of $\widetilde{G}$ if and only if the sets $V_1 = p(\widetilde{V}_1), \ldots, V_n = p(\widetilde{V}_n)$ form the required partition of the vertex set of G (here $p\colon \widetilde{G} \to G$ is a covering).

(b) It is easy to verify that any connected graph without cycles is uniquely determined by the matrix $R(G)$. □

Theorem 1.26 is used to prove the following assertion.

Theorem 1.27 (see [74]). *If finite connected graphs G and G' have the same universal covering, then they have a common finite covering; i.e., there exists a finite graph H which covers both G and G'.*

Proof. According to Theorem 1.26, $R(G) = R(G') = R = (r_{ij})$. Let $V_1, \ldots, V_\alpha$ and $V_1', \ldots, V_\alpha'$ be the corresponding partitions of the vertex sets of G and G'. For convenience, we turn G and G' into directed graphs by replacing each edge by a pair of edges with opposite directions and each loop by a pair of directed loops. Let $n_i = |V_i|$ and m_{ij} be the numbers of vertices of types i and $i \to j$, respectively, in the graph G. We define s as the least common multiple of the numbers m_{ij} over all i and j. We set $a_i = s/n_i$ and $b_{ij} = s/m_{ij}$ (if $m_{ij} = 0$, then the number b_{ij} is not defined). It is seen from the definition that $m_{ij} = n_i r_{ij}$, and a_i and b_{ij} are integers. It is also clear that $m_{ij} = m_{ji}$, and therefore $b_{ij} = b_{ji}$.

The most important property of the numbers a_i and b_{ij} is that they are completely determined by the matrix R; i.e., they coincide for the graphs G and G'. Let us prove this. First, we show that the number $f_i = n_i/n_1$ depends only on the matrix R. Indeed, if $r_{i1} \neq 0$, then $f_i = \frac{m_{i1}/r_{i1}}{m_{1i}/r_{1i}} = \frac{r_{1i}}{r_{i1}}$ because $m_{i1} = m_{1i}$. Of course, it may happen that $r_{i1} = 0$. In this case, we take a sequence of numbers $1 = j_1, j_2, \ldots, j_{h_i} = i$ such that $r_{j_l j_{l+1}} \neq 0$ for $l = 1, 2, \ldots, h_i - 1$ (it exists because the graph G is connected) and set $f_i = \prod_{l=0}^{h_i-1} \left(r_{j_l j_{l+1}}/r_{j_{l+1} j_l}\right)$. Now, we can determine the numbers a_i and b_{ij} from the relations

$$a_1 = n_1^{-1}\mathrm{LCM}(m_{ij}) = n_1^{-1}\mathrm{LCM}(n_i r_{ij}) = n_1^{-1}\mathrm{LCM}(f_i n_1 r_{ij}) = \mathrm{LCM}(f_i r_{ij}),$$
$$a_i = s/n_i = a_1 n_1/n_i = a_1/f_i, \quad \text{and} \quad b_{ij} = a_i/r_{ij}.$$

We enumerate the edges of type $i \to j$ going from a vertex $v \in V_i$ by the integers from 0 to $r_{ij} - 1$; we denote the number of an edge e by $g(v, e)$. The numbers $g'(v', e')$ for the graph G' are defined similarly.

Consider the directed graph H defined as follows. The vertices of H have the form (i, v, v', p), where $1 \leq i \leq \alpha$, $v \in V_i$, $v' \in V_i'$, and $0 \leq p < a_i$. The edges of H have the form (i, j, e, e', q), where $1 \leq i, j \leq \alpha$; e and e' are edges of type $i \to j$ in the graphs G and G', respectively; and $0 \leq q < b_{ij}$. An edge (k, j, e, e', q) has head (starting vertex) (i, v, v', p) if and only if $i = k$, v is the head of e, v' is the head of e', $q = [p/r_{ij}]$, and $g(v, e) - g'(v', e') \equiv p \pmod{r_{ij}}$; an edge (j, k, e, e', q) has tail (i, v, v', p) if and only if $i = k$, v is the tail of e, v' is the tail of e', $q = [p/r_{ij}]$, and $g(v, -e) - g'(v', -e') \equiv p \pmod{r_{ij}}$, where $-e$ and $-e'$ are the edges e and e' with opposite directions.

It follows from $a_i = r_{ij} b_{ij}$ that the head of each edge (i, j, e, e', q) is determined uniquely. Indeed, suppose that $x \equiv g(v, e) - g'(v', e') \pmod{r_{ij}}$ and $0 \leq x < r_{ij}$. We set $p = q r_{ij} + x$. Clearly, the conditions $0 \leq p < a_i = r_{ij} b_{ij}$, $q = [p/r_{ij}]$, $0 \leq q < b_{ij}$, and $p \equiv g(v, e) - g'(v', e') \pmod{r_{ij}}$ uniquely determine the number p. The head of the edge (i, j, e, e', q) is (i, v, v', p). It is also clear that a vertex (i, v, v', p) is the tail of an edge (j, k, e, e', q) if and only if it is the head of the edge $-(j, k, e, e', q) = (k, j, -e, -e', q)$. Therefore, the tail of each edge is uniquely determined as well. It follows from $b_{jk} = b_{kj}$ that the change of direction of edges is well defined. Thus, the graph H is well defined, and its edges are divided into pairs of edges with opposite directions.

We define a covering $p\colon H \to G$ by sending each edge (j, k, e, e', q) of the graph H to the edge e of G; clearly, each vertex (i, v, v', p) is then mapped to the vertex v. We must only verify that this map establishes a one-to-one correspondence between the edges going from a vertex (i, v, v', p) and the edges going from the vertex v. Consider any edge e of type $i \to j$ going from v. The corresponding edge e' in the graph G' has type $i \mapsto j$ and head v', and $g'(v', -e') \equiv g(v, -e) + p \pmod{r_{ij}}$. The edge e is the image of precisely one edge going from (i, v, v', p), namely, of (i, j, e, e', q), where $q = [p/r_{ij}]$.

The projection of the edge $-(i, j, e, e', q) = (j, i, -e, -e', q)$ is $-e$; therefore, we can construct a covering of the initial (undirected) graph G from the covering $p\colon H \to G$ of the directed graphs by replacing each pair of edges with opposite directions by one undirected edge.

A covering $p_1 : H \to G_1$ is constructed similarly.

The graph H is not necessarily connected, but each of its connected components has the required property. □

3. Graph Invariants

The graphs which we consider in this section may have loops and multiple edges.

Let e be an edge of a graph G. The graphs obtained from G by deleting the edge e and by contracting e to a point are denoted by $G - e$ and G/e, respectively. Note that if e is a loop, then $G - e = G/e$. It is easy to verify that the operations of edge contraction and deletion commute; i.e., if e_1 and e_2 are two edges of G, then $(G/e_1)/e_2 = (G/e_2)/e_1$, $(G - e_1) - e_2 = (G - e_2) - e_1$, and $(G/e_1) - e_2 = (G - e_2)/e_1$.

We say that graphs G_1 and G_2 are *isomorphic* if there exists a homeomorphism $h\colon G_1 \to G_2$ which acts on the vertices as a one-to-one map. A *graph invariant* is a map from the set of all graphs to some other set which takes isomorphic graphs to the same element. A *polynomial invariant* is an invariant taking values in the polynomial ring; in other words, it assigns a polynomial to each graph, and the polynomials assigned to isomorphic graphs coincide.

The most important polynomial invariants of graphs satisfy the relation

$$F(G) = aF(G/e) + bF(G - e), \tag{1}$$

where a and b are some fixed polynomials (or constants). Relation (1) may hold for all edges e (including loops), or it may hold only for edges e with different end vertices.

Given an arbitrary graph, we can always obtain the graph $\overline{K}_n$ consisting of n isolated vertices (the complement of the complete graph K_n) by contracting and deleting its edges. Therefore, if relation (1) holds for any edge e, then the values of the polynomial F for the graphs $\overline{K}_n$ completely determine this polynomial. If relation (1) holds only for edges that are not loops, then we must specify the values of F for graphs consisting of several isolated vertices to which several loops may be attached.

Relation (1) might be self-contradictory in the sense that eliminating and contracting edges in different orders, we might obtain different polynomials. Thus, we must verify that different successions of operations in the calculation of $F(G)$ always give the same result.

Theorem 1.28. *The polynomial $F(G)$ is well defined.*

Proof. Let e_1 and e_2 be edges of a graph G. Then

$$\begin{aligned} aF(G/e_1) + bF(G - e_1) &= a^2F((G/e_1)/e_2) + abF((G/e_1) - e_2) \\ &\quad + abF((G - e_1)/e_2) + b^2F(G - e_1 - e_2) \\ &= a^2F((G/e_2)/e_1) + abF((G/e_2) - e_1) \\ &\quad + abF((G - e_2)/e_1) + b^2F(G - e_1 - e_2) \\ &= aF(G/e_2) + bF(G - e_2). \end{aligned}$$

Thus, the result of calculation does not depend on the order in which the edges are deleted and contracted. □

Clearly, if graphs G_1 and G_2 are isomorphic, then, deleting respective edges of these graphs simultaneously, we obtain the same result; therefore, $F(G_1) = F(G_2)$, i.e., F is a polynomial invariant of graphs. In some cases, this polynomial can be used to recognize nonisomorphic graphs.

Taking different polynomials a and b and assigning different values of F to the graphs $\overline{K}_n$ (or to the graphs consisting of isolated vertices with loops), we obtain different polynomials F. Some of them have interesting geometric interpretations.

3.1. The Chromatic Polynomial. The *chromatic polynomial* $P(G,t)$ is defined by the relation

$$P(G,t) = -P(G/e,t) + P(G - e,t),$$

which must hold for every edge e. The value of $P(G,t)$ for the graph G consisting of n isolated vertices is set to be equal to t^n.

Theorem 1.29. *For any positive integer t, $P(G,t)$ is equal to the number of different colorings of the vertices of G with t colors under which the end vertices of any edge have different colors.*

Proof. Let $t \in \mathbb{N}$, and let $\widetilde{P}(G,t)$ be the number of t-colorings of G. Note that if the graph G has at least one loop, then $\widetilde{P}(G,t) = 0$ (the end vertices of the loop coincide, and they cannot have different colors). If G consists of n isolated vertices, then $\widetilde{P}(G,t) = t^n = P(G,t)$. Therefore, it suffices to show that $\widetilde{P}(G,t) = -\widetilde{P}(G/e,t) + \widetilde{P}(G-e,t)$ for any edge e of G.

First, suppose that e is not a loop. Let v_1 and v_2 be the end vertices of e. The number of t-colorings of the graph $G-e$ under which v_1 and v_2 have the same color is $\widetilde{P}(G/e,t)$, and the number of t-colorings under which v_1 and v_2 have different colors is $\widetilde{P}(G,t)$. Therefore, $\widetilde{P}(G-e,t) = \widetilde{P}(G,t)+\widetilde{P}(G/e,t)$, as required.

Now, suppose that e is a loop. In this case, we have $\widetilde{P}(G,t) = 0$ and $\widetilde{P}(G/e,t) = \widetilde{P}(G-e,t)$, because the graphs G/e and $G-e$ coincide. □

Corollary. *The number of t-colorings of a graph G depends on t polynomially.*

Exercise 11. Prove that if K_n is the complete graph on n vertices, then $P(K_n,t) = t(t-1)\cdots(t-n+1)$.

Theorem 1.30 (Whitney [143]). *Let G be a loopless graph on n vertices. Then*

$$P(G,t) = t^n - a_1t^{n-1} + a_2t^{n-2} - a_3t^{n-3} + \cdots, \quad \textit{where } a_i \geq 0.$$

Proof. If the graph G has one vertex and no edges, then $P(G,t) = t$. Let $\{e_1,\dots,e_k\}$ be the edge set of G. Then

$$\begin{aligned} P(G) &= P(G-e_1) - P(G/e_1) \\ &= P(G-e_1-e_2) - P((G-e_1)/e_2) - P(G/e_1); \end{aligned}$$

each of the graphs G/e_1 and $(G-e_1)/e_2$ has $n-1$ vertices. Clearly, the graph $G-e_1-e_2-\cdots-e_k = \overline{K}_n$ consists of n isolated vertices. Hence

$$P(G) = P(\overline{K}_n) - g_1 - \cdots - g_k = t^n - g_1 - \cdots - g_k,$$

where $g_1,\dots,g_k$ are the chromatic polynomials of graphs on $n-1$ vertices. □

Theorem 1.31 (Whitney [143]). *The chromatic polynomial of a graph can be calculated by the combinatorial formula*

$$P(G,t) = \sum_{H\subset G} (-1)^{e(H)}t^{c(H)},$$

where the summation is over all subgraphs $H \subset G$ whose vertex sets coincide with that of the graph G; here $e(H)$ is the number of edges and $c(H)$ is the number of connected components of H.

Proof. Consider the polynomial

$$\widetilde{P}(G,t) = \sum_{H\subset G} (-1)^{e(H)}t^{c(H)}.$$

The graph $G = \overline{K}_n$ has precisely one subgraph H with the same vertices as G, namely, $H = G = \overline{K}_n$. We have $e(H) = 0$ and $c(H) = n$. For this graph,

$$\widetilde{P}(G,t) = \sum_{H \subset G} (-1)^{e(H)} t^{c(H)} = t^n = P(G,t).$$

It remains to show that $\widetilde{P}(G,t) = -\widetilde{P}(G/e,t) + \widetilde{P}(G-e,t)$. For this purpose, we represent the polynomial $\widetilde{P}(G,t)$ in the form $\widetilde{P}(G,t) = \sum_{e \in H} + \sum_{e \notin H}$. It is easy to see that $\sum_{e \in H} = -\widetilde{P}(G/e,t)$ and $\sum_{e \notin H} = \widetilde{P}(G-e,t)$; the minus sign in the first equality appears because the graph H/e has one edge fewer than H. □

3.2. A Polynomial in Three Variables. In [93], a polynomial invariant $f(G; t, x, y)$ satisfying the relation $f(G) = xf(G/e) + yf(G-e)$ (for all edges e, including loops) and taking the value t^n for $\overline{K}_n$ was introduced.

Exercise 12. (a) Prove that if G is a connected tree with n edges, then $f(G) = t(x+ty)^n$.

(b) Prove that if G is a cycle of length n, then $f(G) = t(x+ty)^n + (t-1)x^n$.

The coefficients of the polynomial f have the following combinatorial interpretation.

Theorem 1.32. *Let G be a graph with v vertices and e edges. Then*

$$f(G) = \sum_{i=0}^{e} \sum_{j=1}^{v} b_{ij} t^j x^{e-i} y^i;$$

for every i and j, b_{ij} is the number of subsets Y of the edge set of G such that each Y has i elements and the graph obtained from G by deleting all edges belonging to Y has j connected components.

Proof. It is seen from the definition of the polynomial $f(G)$ that this polynomial can be calculated as follows. First, we partition the edges of G into two sets X and Y. Then, we contract all edges from X and delete all edges from Y. This sequence of operations corresponds to the monomial $t^j x^{e-i} y^i$. To calculate the polynomial $f(G)$, we must consider all sets Y and take the sum of the obtained monomials. □

Corollary 1. *The coefficients of $f(G)$ are always nonnegative.*

Corollary 2. *If graphs K and H have no common vertices, then $f(K \cup H) = f(K)f(H)$.*

Corollary 3. *If the intersection of graphs K and H consists of precisely one vertex, then $f(K \cup H) = t^{-1} f(K) f(H)$.*

Proof. Let Y_1 and Y_2 be subsets of the edge sets of K and H, respectively. Suppose that, by deleting the edges that belong to Y_1 and Y_2 from K and H, we obtain graphs with j_1 and j_2 connected components. Then, by deleting the edges that belong to $Y_1 \cup Y_2$ from $K \cup H$, we obtain a graph with j_1+j_2-1 connected components (two connected components merge because the graphs K and H have a common vertex). □

Corollary 3 makes it possible to construct examples of nonisomorphic graphs with the same polynomial $f(G)$. Namely, taking graphs K and H and successively identifying one vertex of K with different vertices of H, we obtain graphs that are not necessarily isomorphic but have the same polynomial $f(G)$. The following theorem describes yet another transformation of graphs which produces graphs with the same polynomial $f(G)$.

Theorem 1.33. *Let u_1 and u_2 be vertices of a graph K, and let v_1 and v_2 be vertices of a graph H. Suppose that G_1 is the graph obtained from K and H by identifying u_1 with v_1 and u_2 with v_2, and G_2 is the graph obtained from K and H by identifying u_1 with v_2 and u_2 with v_1. Then $f(G_1) = f(G_2)$.*

Proof. Take some subsets X_1 and X_2 in the edge sets of K and H. These subsets are in one-to-one correspondence with subsets of the edge sets of each of the graphs G_i $(i = 1, 2)$. The edge joining v_1 and v_2 in the graph G_1 belongs to the set $X = X_1 \cup X_2$ if and only if it belongs to X in the graph G_2. Therefore, in G_1 and G_2, the numbers of connected components of the graph formed by the edges from X are equal. □

3.3. The Bott–Whitney Polynomial. The *Bott Whitney polynomial* $R(G,t)$ is defined by the relation

$$R(G,t) = R(G/e,t) - R(G-e,t), \tag{2}$$

which must hold only for the edges e that are not loops. The value of R for the graph consisting of one vertex and n loops is equal to $(t-1)^n$; the value of R for a union of such graphs is equal to the product of values for the factors.

Exercise 13. Prove that $R(G,1) = 0$.

The Bott Whitney polynomial has the following interpretation.

Theorem 1.34. *Suppose that G is a graph, H is a set of edges of G, and $\overline{H}$ is the complement of H in G (i.e., the graph containing the edges of G that are not contained in H; the vertices of $\overline{H}$ coincide with those of G). Then the graph $\overline{H}$ is homotopy equivalent to a disjoint union of wedges of circles. Let $b_1(\overline{H})$ be the number of all circles in these wedges. Then*

$$R(G,t) = \sum_{H \subset G} (-1)^{e(H)} t^{b_1(\overline{H})}, \tag{3}$$

where $e(H)$ is the number of elements in H and the summation is over all sets of edges, including the empty set.

Proof. Suppose that the graph G consists of several isolated vertices with loops. If the total number of these loops is m, then

$$\sum_{i=0}^{m} \sum_{|H|=i} (-1)^i t^{m-i} = \sum_{i=0}^{m} \binom{m}{i} (-1)^i t^{m-i} = (t-1)^m = R(G,t).$$

It remains to prove (3). The polynomial $R(G)$ can be represented as $R(G) = \sum_{e\notin H} + \sum_{e\in H}$. It is easy to verify that $\sum_{e\notin H} = R(G/e)$ and $\sum_{e\in H} = -R(G-e)$. Indeed, let $e \notin H$; then $e \in \overline{H}$. By assumption, e is not a loop; hence the graphs $\overline{H}$ and $\overline{H}/e$ are homotopy equivalent, which means that $b_1(\overline{H}) = b_1(\overline{H}/e)$. Now, suppose that $e \in H$, i.e., $H = H_1 \cup \{e\}$. Then $e(H) = e(H_1) + 1$, and $(-1)^{e(H)} = -(-1)^{e(H_1)}$. The complement of H in G coincides with the complement of H_1 in $G - e$. □

Theorem 1.35. *Suppose that a graph G consists of two graphs G_1 and G_2, which are either disjoint or have one common vertex and no common edges. Then $R(G) = R(G_1)R(G_2)$.*

Proof. We apply (3). Let us represent the set H as $H = H_1 \cup H_2$, where each H_i consists of the edges of G_i. Then $e(H) = e(H_1) + e(H_2)$ and $b_1(\overline{H}) = b_1(\overline{H}_1) + b_1(\overline{H}_2)$, where $\overline{H}_1$ and $\overline{H}_2$ are the complements of H in the graphs G_1 and G_2, respectively. Therefore, $R(G) = R(G_1)R(G_2)$. □

Corollary 1. *If a graph G has a pendant edge (i.e., an edge one of whose end vertices is not incident to any other edge), then $R(G) = 0$.*

Proof. It is easy to show that if a graph G_1 consists of one edge, then $R(G_1) = 0$. The graph G with a pendant edge can be represented as the union of such a graph G_1 and some graph G_2 intersecting G_1 in one vertex. Therefore, $R(G) = R(G_1)R(G_2) = 0$. □

Corollary 2. *If a graph G is a cycle, then $R(G) = t - 1$.*

Proof. Let e be an edge of G. Then the graph $G-e$ has a pendant edge, and hence $R(G-e) = 0$. The graph G/e is a cycle with fewer edges. It remains to note that for a graph G consisting of one loop, we have $R(G) = t-1$. □

The Bott–Whitney polynomial, unlike the chromatic polynomial, is a *topological invariant*; i.e., the Bott–Whitney polynomials of homeomorphic graphs coincide.

Theorem 1.36. *The Bott Whitney polynomial is a topological graph invariant.*

Proof. Graphs G and G' are homeomorphic if and only if there exists a sequence of graphs with first term G and last term G' in which all pairs of neighboring graphs are related by the following transformation: on an edge e, an additional vertex v is taken, and e is replaced by two edges e_1 and e_2 with the common vertex v. Therefore, it is sufficient to show that if a graph G' is obtained from a graph G by such a transformation, then $R(G') = R(G)$. The edge e_1 is not a loop; hence $R(G') = R(G'/e_1) - R(G' - e_1)$. The graph $G' - e_1$ has the pendant edge e_2; therefore, $R(G' - e_1) = 0$. Clearly, the graph G'/e_1 is isomorphic to G. □

In [**144**], Whitney defined a set of graph invariants which coincides with the set of coefficients of the polynomial $R(G)$. In [**18**], Bott independently defined a polynomial invariant of finite CW-complexes which coincides with $R(G)$ in the one-dimensional case. The properties of the Bott Whitney polynomial were studied in detail in [**137**].

3.4. The Tutte Invariants. Let $g(G)$ be a function on the set of graphs taking values in a commutative associative ring with identity. The function g is called a *Tutte invariant*, or a *V-function*,[7] if the following conditions hold:

(i) $g(\varnothing) = 1$;

(ii) if an edge e is not a loop, then $g(G) = g(G/e) + g(G - e)$;

(iii) if a graph G is the disjoint union of graphs K and H, then $g(G) = g(K)g(H)$.

Each Tutte invariant is completely determined by its values on the graphs consisting of a single vertex and several loops.

One of the most important Tutte invariants is the *dichromatic polynomial*

$$Q(G,t,z) = \sum_{H \subset G} z^{b_1(H)} t^{c(H)}$$

(introduced by Tutte), where the summation is over all subgraphs $H \subset G$ whose vertex sets coincide with that of the graph G, $b_1(H)$ is the number of independent cycles in H ($b_1(H)$ is the total number of circles in the disjoint union of wedges of circles that is homotopy equivalent to the graph H), and $c(H)$ is the number of connected components of H.

Theorem 1.37. *The dichromatic polynomial is a Tutte invariant.*

Proof. Obviously, condition (i) holds. To prove (ii), we represent the dichromatic polynomial as $Q(G) = \sum_{e \notin H} + \sum_{e \in H}$. It is clear that $\sum_{e \notin H} = Q(G - e)$; the equality $\sum_{e \in H} = Q(G/e)$ follows from $b_1(H) = b_1(H/e)$ and

[7] This term was used by Tutte.

$c(H) = c(H/e)$. Condition (iii) holds because the functions b_1 and c are additive for disjoint graphs. □

An edge e with end vertices v_1 and v_2 in a graph G is called a *bridge* if any path from v_1 to v_2 in G passes through e. Tutte [**127**] introduced a polynomial $T(G,x,y)$ which satisfies condition (ii) only for edges e that are neither loops nor bridges. To be more precise, the *Tutte polynomial* $T(G,x,y)$ has the following properties.

(a) For a graph G containing precisely one edge, $T(G,x,y) = x$ if this edge is a bridge, and $T(G,x,y) = y$ if this edge is a loop.

(b) If e is an edge of G which is neither a loop nor a bridge, then $T(G,x,y) = T(G-e,\ x,y) + T(G/e,\ x,y)$.

(c) If e is a bridge, then $T(G,x,y) = xT(G/e,\ x,y)$, and if e is a loop, then $T(G,x,y) = yT(G-e,x,y)$.

Clearly, properties (a) (c) allow us to calculate the Tutte polynomial for any connected graph G. These properties are consistent because the Tutte polynomial can be expressed by the combinatorial formula

$$T(G,x,y) = \sum_{H\subset G} (x-1)^{c(H)-c(G)}(y-1)^{e(H)-v(G)+c(H)},$$

where the summation is over all subgraphs $H \subset G$ with the same vertices as G.

Topology in Euclidean Space

1. Topology of Subsets of Euclidean Space

1.1. The Distance from a Point to a Set. Let A be an arbitrary subset of $\mathbb{R}^n$, and let $x \in \mathbb{R}^n$. The quantity $d(x, A) = \inf_{a\in A} \|x - a\|$ is called the *distance* from the point x to the set A.

Theorem 2.1. (a) *The function $f(x) = d(x, A)$ is continuous for any set $A \subset \mathbb{R}^n$.*

(b) *If A is closed, then the function $f(x) = d(x, A)$ takes positive values for all $x \notin A$.*

Proof. (a) Let $x, y \in \mathbb{R}^n$. Then $d(x, A) = \inf_{a\in A} \|x - a\| \leq \|x - y\| + \inf_{a\in A} \|y - a\| = \|x - y\| + d(y, A)$, i.e., $d(x, A) - d(y, A) \leq \|x - y\|$. Similarly, $d(y, A) - d(x, A) \leq \|x - y\|$. Thus, $|f(x) - f(y)| \leq \|x - y\|$, which means that f is continuous.

(b) If A is closed, then $\mathbb{R}^n \setminus A$ is open. Hence, for any point $x_0 \in \mathbb{R}^n \setminus A$, there exists a $\delta > 0$ such that the ball of radius δ centered at x_0 is contained in $\mathbb{R}^n \setminus A$. Therefore, $d(x, A) \geq \delta > 0$. □

Remark 2.1. Theorem 2.1, as well as Theorem 2.2 below, remains valid for any metric space. The proof is the same.

For arbitrary sets $A, B \subset \mathbb{R}^n$, the quantity $d(A, B) \coloneqq \inf_{a\in A,\ b\in B} \|a - b\|$ is called the *distance* between A and B.

Theorem 2.2. *If $A \subset \mathbb{R}^n$ is closed and $C \subset \mathbb{R}^n$ is compact, then there exists a point $c_0 \in C$ for which $d(A, C) = d(A, c_0)$. If, in addition, the set A is compact, then there is also a point $a_0 \in A$ for which $d(A, C) = d(a_0, c_0)$.*

Proof. The function $f(x) = d(x, A)$ is continuous on the compact set C; therefore, it attains its minimum at some point $c_0 \in C$. If the set A is compact, then the continuous function $g(x) = d(c_0, x)$ on A attains its minimum at some point $a_0 \in A$. □

Problem 21. Is it true that $d(A, C) \leq d(A, B) + d(B, C)$?

A distance between sets which satisfies the triangle inequality is defined as follows. Take any sets $A, B \subset \mathbb{R}^n$. Consider the set T of all positive numbers t with the following properties:

- for any $a \in A$, there exists a $b \in B$ such that $\|a - b\| \leq t$;
- for any $b \in B$, there exists an $a \in A$ such that $\|a - b\| \leq t$.

The number $d_H(A, B) = \inf_{t \in T} t$ is called the *Hausdorff distance* between the sets A and B.

Problem 22. Prove that $d_H(A, C) \leq d_H(A, B) + d_H(B, C)$.

1.2. Extension of Continuous Maps. The problem of extending a continuous map $f \colon A \to Y$, where $A \subset X$, to a continuous map from the entire space X to Y often arises in topology. In the simplest case where $X = \mathbb{R}^n$ and $Y = \mathbb{R}$, a solution to the extension problem is provided by the following assertion, known as *Urysohn's lemma.*

Theorem 2.3 (Urysohn's lemma [**129**]). *If A and B are disjoint closed subsets of $\mathbb{R}^n$, then there exists a continuous map $f \colon \mathbb{R}^n \to [-1, 1]$ such that $f(A) = \{-1\}$ and $f(B) = \{1\}$.*

Proof. By assumption, any point $x \in \mathbb{R}^n$ is either outside A or outside B. The sets A and B are closed; therefore, we have either $d(x, A) > 0$ (in the former case) or $d(x, B) > 0$ (in the latter case). In any case, $d(x, A) + d(x, B) > 0$; hence the function

$$f(x) = \frac{d(x, A) - d(x, B)}{d(x, A) + d(x, B)}$$

is well defined for all $x \in \mathbb{R}^n$. The continuity of the functions $d(x, A)$ and $d(x, B)$ implies the continuity of $f(x)$. Clearly, $f(A) = \{-1\}$ and $f(B) = \{1\}$. Moreover, for any point x, we have

$$-1 \leq \frac{-d(x, B)}{d(x, A) + d(x, B)} \leq \frac{d(x, A) - d(x, B)}{d(x, A) + d(x, B)} \leq \frac{d(x, A)}{d(x, A) + d(x, B)} \leq 1. \qquad \square$$

Corollary. *If A and B are disjoint closed subsets of $\mathbb{R}^n$, then there exist disjoint open sets $U \supset A$ and $V \supset B$ with disjoint closures.*

Proof. Let $f: \mathbb{R}^n \to [-1,1]$ be a continuous function such that $f(A) = \{-1\}$ and $f(B) = \{1\}$. For U and V we can take the preimages of the sets $[-1, -\frac{1}{2})$ and $(\frac{1}{2}, 1]$. □

Using Urysohn's lemma, it is possible to prove the existence of an extension for any continuous function on a closed subset of Euclidean space.

Theorem 2.4 (Tietze). *Suppose that $X \subset \mathbb{R}^n$ is a closed set and $f: X \to [-1,1]$ is a continuous function. Then there exists a continuous function $F: \mathbb{R}^n \to [-1,1]$ such that its restriction to X coincides with f.*

Proof. Let $r_k = \frac{1}{2}\left(\frac{2}{3}\right)^k$ for $k = 1, 2, \ldots$. Then $3r_1 = 1$ and $r_k \to 0$ as $k \to \infty$. First, we construct a sequence of continuous functions $f_1, f_2, \ldots$ on X such that $-3r_k \le f_k \le 3r_k$ and a sequence of continuous functions $g_1, g_2, \ldots$ on $\mathbb{R}^n$. We define $f_1 = f$. Suppose that the functions $f_1, \ldots, f_k$ are already constructed. Consider the disjoint closed sets

$$A_k = \{x \in X : f_k(x) \le -r_k\} \quad \text{and} \quad B_k = \{x \in X : f_k(x) \ge r_k\}.$$

Applying Urysohn's lemma to these sets, we obtain a continuous map $g_k: \mathbb{R}^n \to [-r_k, r_k]$ such that $g_k(A_k) = \{-r_k\}$ and $g_k(B_k) = \{r_k\}$. On the set A_k, the functions f_k and g_k take values between $-3r_k$ and $-r_k$; on the set B_k, they take values between r_k and $3r_k$; at all the other points of X, these functions take values between $-r_k$ and r_k. We set $f_{k+1} = f_k - g_k \restriction X$. The function f_{k+1} is continuous on X, and $|f_{k+1}(x)| \le 2r_k = 3r_{k+1}$ for all $x \in X$.

Now, consider the sequence of functions $g_1, g_2, \ldots$ on $\mathbb{R}^n$. By construction, $|g_k(y)| \le r_k$ for all $y \in \mathbb{R}^n$. The series $\sum_{k=1}^{\infty} r_k = \frac{1}{2}\sum_{k=1}^{\infty}\left(\frac{2}{3}\right)^k$ converges; therefore, the series $\sum_{k=1}^{\infty} g_k(x)$ converges uniformly on $\mathbb{R}^n$ to some continuous function $F(x) = \sum_{k=1}^{\infty} g_k(x)$. We have

$$\begin{aligned}(g_1 + \cdots + g_k) \restriction X &= (f_1 - f_2) + (f_2 - f_3) + \cdots + (f_k - f_{k+1}) \\ &= f_1 - f_{k+1} = f - f_{k+1}.\end{aligned}$$

But $\lim_{k\to\infty} f_{k+1}(y) = 0$ for any point $y \in \mathbb{R}^n$; hence $F(x) = f(x)$ for all $x \in X$. Moreover,

$$\begin{aligned}|F(x)| &\le \sum_{k=1}^{\infty} |g_k(x)| \le \sum_{k=1}^{\infty} r_k = \sum_{k=1}^{\infty} \frac{1}{2}\left(\frac{2}{3}\right)^k \\ &= \frac{1}{3}\sum_{k=0}^{\infty}\left(\frac{2}{3}\right)^k = \frac{1}{3}\left(1 - \frac{2}{3}\right)^{-1} = 1.\end{aligned}$$

□

Corollary. *Suppose that $X \subset \mathbb{R}^n$ is a closed set and $f: X \to \mathbb{R}$ is a continuous function. Then there exists a continuous function $F: \mathbb{R}^n \to \mathbb{R}$ whose restriction to X coincides with f.*

Proof. Consider the homeomorphism $g: \mathbb{R} \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ defined by $g(x) = \arctan(x)$. The function $g(f(x))$ has a continuous extension G over $\mathbb{R}^n$ such that $|G(x)| \leq \pi/2$ for all $x \in \mathbb{R}^n$. Consider the closed set $A = \{y \in \mathbb{R}^n : |G(x)| = \pi/2\}$. Clearly, $A \cap X = \varnothing$; therefore, by Urysohn's lemma, there exists a continuous function $\varphi: \mathbb{R}^n \to [0, 1]$ such that $\varphi(A) = \{0\}$ and $\varphi(X) = \{1\}$. We set $F(y) = \tan(\varphi(y)G(y))$. If $x \in X$, then $F(x) = \tan(\arctan f(x)) = f(x)$. Moreover, $\varphi(y)G(y) < \pi/2$ for all $y \in \mathbb{R}^n$; therefore, the function F is well defined. □

The Tietze extension theorem and its corollary remain valid when f is a map to $\mathbb{R}^m$; for such maps, they are proved applying the Tietze theorem coordinatewise.

The Tietze extension theorem is often used to construct extensions of continuous maps. The following theorem is an interesting example of such an application.

Theorem 2.5. *If $A \subset \mathbb{R}^m \times \{0\}$ and $B \subset \{0\} \times \mathbb{R}^n$ are homeomorphic closed subsets of $\mathbb{R}^{m+n} = \mathbb{R}^m \times \mathbb{R}^n$, then the sets $\mathbb{R}^{m+n} \setminus A$ and $\mathbb{R}^{m+n} \setminus B$ are homeomorphic.*

Proof. Let $f_a: A \to B$ and $f_b: B \to A$ be mutually inverse homeomorphisms. According to the Tietze theorem, they can be extended to maps $F_a: \mathbb{R}^m \to \mathbb{R}^n$ and $F_b: \mathbb{R}^n \to \mathbb{R}^m$. Consider the maps $\mathcal{F}_a, \mathcal{F}_b: \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^n$ defined by

$$\mathcal{F}_a(x, y) = (x, y - F_a(x)) \quad \text{and} \quad \mathcal{F}_b(x, y) = (x - F_b(y), y).$$

These maps are invertible (for example, $\mathcal{F}_a^{-1}(x, y) = (x, y + F_a(x))$). Clearly, $\mathcal{F}_a$ and $\mathcal{F}_b$ take the set

$$\begin{aligned} X &= \{(x, y) \in \mathbb{R}^{m+n} : x \in A,\ y = f_a(x)\} \\ &= \{(x, y) \in \mathbb{R}^{m+n} : y \in B,\ x = f_b(y)\} \end{aligned}$$

to A and B, respectively. Therefore, $\mathbb{R}^{m+n} \setminus A \approx \mathbb{R}^{m+n} \setminus X \approx \mathbb{R}^{m+n} \setminus B$. □

Remark 2.2. The sets $(\mathbb{R}^m \times \{0\}) \setminus A$ and $(\{0\} \times \mathbb{R}^n) \setminus B$ may not be homeomorphic. For example, $\mathbb{R}^3 \setminus S^1$, where S^1 is the circle standardly embedded in $\mathbb{R}^3$, is not homomorphic to $\mathbb{R}^3 \setminus K$, where K is a trefoil (see Figure 13 in Chapter 6).

1.3. The Lebesgue Covering Theorems. Let $\mathcal{U}$ be an open cover of a topological space $A \subset \mathbb{R}^n$. The *Lebesgue number* of the cover $\mathcal{U}$ is defined as the least upper bound for all numbers $\delta \geq 0$ such that any subset $B \subset A$ of diameter[1] less than δ is contained in some element of $\mathcal{U}$ (i.e., in one of the open sets which constitute the cover $\mathcal{U}$).

Theorem 2.6 (Lebesgue). *If A is a compact subset of $\mathbb{R}^n$, then the Lebesgue number of any open cover $\mathcal{U}$ of A is strictly positive.*

Proof. The cover $\mathcal{U}$ has a finite subcover $\{U_1, \dots, U_k\}$. We set $f_i(x) = d(x, A \setminus U_i)$ and $f = \max(f_1, \dots, f_k)$. The function f is continuous. Moreover, if $a \in A$, then $f(a) > 0$. Indeed, $a \in U_i$ for some i; hence $f_i(a) > 0$, because the set $A \setminus U_i$ is closed. Therefore, the image of A under the continuous map $f \colon A \to \mathbb{R}$ is compact and does not contain 0. Thus, $d(0, f(A)) > 0$, i.e., there exists a number $\delta > 0$ such that $f(a) > \delta$ for any $a \in A$. This means that $f_i(a) > \delta$ for some i, i.e., the intersection of A with the ball of radius δ centered at a is contained in U_i. Therefore, any set $B \subset A$ of diameter less than δ belongs to some U_i. $\square$

Problem 23. Using the Lebesgue theorem, prove that any continuous function f on a compact set $A \subset \mathbb{R}^n$ is uniformly continuous on this set.

Lebesgue suggested the following definition of the topological dimension of a compact set $X \subset \mathbb{R}^n$. Let $\mathcal{U}$ be a finite cover of X by closed sets. The *order* of the cover $\mathcal{U}$ is defined as the minimum integer m for which at least one point $x \in X$ belongs to m elements of the cover $\mathcal{U}$ and no point $x \in X$ belongs to more than m elements of $\mathcal{U}$. We say that the *topological dimension* of a compact set $X \subset \mathbb{R}^n$ is equal to k if k is the least nonnegative integer with the following property: for any $\varepsilon > 0$, there exists a finite cover of X by closed sets of diameter less than ε which has order $k + 1$.

Theorem 2.7 (Lebesgue). *The topological dimension of any n-dimensional simplex Δ^n equals n.*

Proof (Sperner [121]). First, we prove that if $\mathcal{U}$ is a finite cover of the simplex Δ^n by closed sets of sufficiently small diameter, then the order of $\mathcal{U}$ is at least $n+1$. Let $\Delta_0^{n-1}, \dots, \Delta_n^{n-1}$ be the $(n-1)$-faces of Δ^n, and let a_i be the vertex of Δ^n opposite to the face Δ_i^{n-1}. In the topological space Δ^n, the subsets $\Delta^n \setminus \Delta_i^{n-1}$ are open. Clearly, they form a cover of Δ^n. Let $\varepsilon > 0$ be the Lebesgue number of this open cover. We show that if $\mathcal{U}$ is a finite cover of Δ^n by closed sets of diameter less than ε, then the order of $\mathcal{U}$ is at least $n+1$. Suppose that $\mathcal{U} = \{U_0, \dots, U_m\}$. Since the diameter of each U_j is less than ε, it follows that U_j is entirely contained in some $\Delta^n \setminus \Delta_i^{n-1}$,

[1]The *diameter* of a set is the least upper bound of pairwise distances between its points.

i.e., does not intersect the face Δ_i^{n-1}. Every vertex a_i belongs to some U_j, and this U_j cannot contain other vertices of the simplex Δ^n.

To each set U_i we assign one of the faces $\Delta_{\varphi(i)}^{n-1}$ disjoint from this set. We obtain a correspondence $\varphi\colon \{0,\dots,m\} \to \{0,\dots,n\}$. For $k = 0,\dots,n$, consider the union A_k of those U_i for which $\varphi(i) = k$. Clearly, $\bigcup_{k=0}^n A_k = \bigcup_{i=0}^m U_i = \Delta^n$, $a_k \in A_k$, and $A_k \cap \Delta_k^{n-1} = \varnothing$. Using Sperner's lemma (see p. 81), we can deduce from these relations (and the closedness of all A_k) that the sets A_k have a common point x. Indeed, let us label each point of the simplex Δ^n by the minimum number k for which A_k contains this point. According to Sperner's lemma, at least one of the simplices from the pth barycentric subdivision[2] of Δ^n is completely labeled. Choose an arbitrary point x_p in this simplex. The sequence $\{x_p\}$ has a convergent subsequence $\{x_{p_q}\}$. The point $x = \lim_{q\to\infty} x_{p_q}$ belongs to all of the A_k. Indeed, each set A_k contains one of the vertices of the simplex in which the point x_{p_q} was chosen, and the edge lengths of such simplices tend to zero as $q \to \infty$.

It remains to construct a cover of order $n+1$ of Δ^n by closed sets of arbitrarily small diameter. Consider the $(m+1)$st barycentric subdivision of Δ^n. For each vertex of the mth barycentric subdivision, consider the set of all closed n-simplices from the $(m+1)$st barycentric subdivision that contain this vertex. Such sets form the required cover. To show this, it suffices to consider the first barycentric subdivision. The barycenter belongs to $n+1$ sets, and the remaining points belong to fewer sets. □

The definition of topological dimension involves a metric quantity, the diameters of sets in a cover. Nevertheless, the topological dimension is indeed a topological invariant, i.e., it is preserved by homeomorphisms.

Theorem 2.8. *If X and Y are homeomorphic compact subsets of Euclidean space, then their topological dimensions coincide.*

Proof. Suppose that X and Y have topological dimensions k_X and k_Y, respectively. By assumption, there exists a homeomorphism $h\colon X \to Y$. For given $\varepsilon > 0$, consider the cover of Y by open balls of diameter ε and the cover of X by the (open) preimages of these balls under the map h. Let δ be the Lebesgue number of the cover of the compact set X. By the definition of topological dimension, there exists a cover of X by closed sets $U_1,\dots,U_m$ of diameter less than δ which has order k_X+1. The family $\{h(U_1),\dots,h(U_m)\}$ is a cover of Y by closed sets of diameter less than ε which has order k_X+1. Thus, $k_Y \le k_X$. Similarly, $k_X \le k_Y$. □

Now, we can prove the celebrated *Brouwer theorem on the invariance of dimension* [24].

[2]The definition of the barycentric subdivision of a simplex is given on p. 81.

Theorem 2.9 (Brouwer). *If $m \neq n$, then no open subset $U \subset \mathbb{R}^m$ can be homeomorphic to an open subset $V \subset \mathbb{R}^n$.*

Proof. Let $h\colon U \to V$ be a homeomorphism. The set U contains an m-simplex Δ^m. The topological dimension of $h(\Delta^m) \subset \mathbb{R}^n$ equals m. The compact set $h(\Delta^m)$ is contained in some simplex Δ^n. Any order n cover of Δ^n by closed sets of small diameter induces an order n cover of $h(\Delta^m)$ by closed sets of small diameter. Therefore, $m \leq n$. Similarly, $m \geq n$. □

1.4. The Cantor Set. Each number $x \in [0,1]$ can be written in ternary notation, i.e., as $x = a_1 3^{-1} + a_2 3^{-2} + \cdots$, where $a_i = 0$, 1, or 2. The *Cantor set* is the set $C \subset [0,1]$ of all numbers that have a ternary expansion without digit 1. For example, the number $1 \cdot 3^{-1} = 2 \cdot 3^{-2} + 2 \cdot 3^{-3} + 2 \cdot 3^{-4} + \cdots$ belongs to C.

Let C_k be the set of numbers $x \in [0,1]$ that have a ternary expansion with kth digit 0 or 2; for example, $C_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$. Each set C_k is closed, and $C = \bigcap_{k=1}^{\infty} C_k$; therefore, C is closed as well.

Theorem 2.10. *Any closed subset $A \subset C$ is a retract of C; i.e., there exists a continuous map $r\colon C \to A$ whose restriction to A is the identity map.*

Proof. Any closed set $A \subset [0,1]$ is compact; therefore, for each point $c \in C$, there exists an $a \in A$ with $d(c, A) = d(c, a)$. There cannot be more than two such points a. First, suppose that $c \in C$, $d(c, A) = d(c, a_1) = d(c, a_2)$, and $a_1 < a_2$. In this case, $a_1 < c < a_2$. The complement of C is everywhere dense; hence we can choose $y \notin C$ so that $a_1 < y < c < a_2$. For every $x \in C \cap [a_1, y)$, we set $r(x) = a_1$, and for every $x \in C \cap (y, a_2]$, we set $r(x) = a_2$. This gives us a map r at all points $c \in C$ for which $d(c, A) = d(c, a_1) = d(c, a_2)$. This map is well defined, because the interval (a_1, a_2) contains no points of A, and therefore the closed intervals $[a_1, a_2]$ and $[a'_1, a'_2]$ constructed for different points c' and c do not intersect.

Suppose that $c \in C$ is a point at which the map r is not yet defined. There exists precisely one point $a \in A$ for which $d(c, A) = d(c, a)$. We set $r(c) = a$.

For the points $a \in A$, the map r can be defined by either of the two methods; in both cases, $r(a) = a$. □

Theorem 2.10 implies the following very unexpected assertion.

Theorem 2.11 (Alexandroff [4]). *Any nonempty compact set $X \subset \mathbb{R}^n$ is the image of the Cantor set C under a continuous map.*

Proof. Let $U_1, U_2, \ldots$ be a countable base of open sets in X. For $c \in C$, consider its ternary expansion $0.c_1 c_2 c_3 \ldots$ containing no 1's (it is unique).

To the point c we assign the set $P(c) = \bigcap_{i=1}^{\infty} \varphi_i(c)$, where

$$\varphi_i(c) = \begin{cases} \overline{U}_i & \text{if } c_i = 0, \\ X \setminus U_i & \text{if } c_i = 2. \end{cases}$$

It is easy to verify that the set $P(c)$ contains at most one point. Indeed, suppose that $a, b \in X$ and $a \neq b$. Then there exists an i such that $a \in U_i$ and $b \notin \overline{U}_i$. If $\varphi_i(c) = \overline{U}_i$, then $b \notin \varphi_i(c)$, and if $\varphi_i(c) = X \setminus U_i$, then $a \notin \varphi_i(c)$. Therefore, $P(c)$ cannot contain both points a and b.

If $P(c)$ consists of one point, then we set $g(c) = P(c)$. We have defined a map g on $A = \{c \in C : \bigcap_{i=1}^{\infty} \varphi_i(c) \neq \varnothing\}$.

It is easy to show that the map $g\colon A \to X$ is surjective. Indeed, let $x \in X$. For the point $c = 0.c_1c_2\cdots \in C$ defined by

$$c_i = \begin{cases} 0 & \text{if } x \in U_i, \\ 2 & \text{if } x \notin U_i, \end{cases}$$

we have $g(c) = x$.

Now, let us show that the map g is continuous. Suppose that $c = 0.c_1c_2\cdots \in A$ $(c_i \neq 1)$ and $\varepsilon > 0$. Choose a set U_k such that $g(c) \in U_k$ and the diameter of $\overline{U}_k$ is less than ε. Take any point $a = 0.a_1a_2\cdots \in A$ $(a_i \neq 1)$ for which $|c - a| < 3^{-2k}$. The inequality $|c - a| < 3^{-2k}$ implies $c_k = a_k$. Therefore, $g(a) \in \varphi_k(a) = \varphi_k(c) = \overline{U}_k$. Thus, $\|g(a) - g(c)\| < \varepsilon$, which means that the map g is continuous at the point c.

Finally, we show that the set A is closed in C, i.e., $C \setminus A$ is open in C. Take $c \in C \setminus A$. We have $\bigcap_{i=1}^{\infty} \varphi_i(c) = \varnothing$, i.e., $\bigcup_{i=1}^{\infty} (X \setminus \varphi_i(c)) = X$. The sets $X \setminus \varphi_i(c)$ form an open cover of X. This cover has a finite subcover; therefore, $\bigcup_{i=1}^{m} (X \setminus \varphi_i(c)) = X$ for some $m \geq 1$. We have $\bigcap_{i=1}^{m} \varphi_i(c) = \varnothing$. Let $a \in C$ be any point satisfying the condition $|c - a| < 3^{-2m}$. Then $a_i = c_i$ for $i = 1, \ldots, m$. Hence $\bigcap_{i=1}^{m} \varphi_i(a) = \varnothing$, i.e., $a \in C \setminus A$. This means that the set $C \setminus A$ is open.

We have constructed a continuous map $g\colon A \to X$, where $A \subset C$ is closed. According to Theorem 2.10, there exists a continuous retraction $r\colon C \to A$. The composition $C \xrightarrow{r} A \xrightarrow{g} X$ is the required map. $\square$

Corollary (Peano). *There exists a surjective map of the interval I onto the k-dimensional cube I^k.*

Proof. First, we construct a continuous map $f\colon C \to I^k$. The Cantor set C is closed in I; therefore, by the Tietze extension theorem, the map f can be extended to a continuous map $F\colon I \to I^k$. $\square$

2. Curves in the Plane

2.1. The Jordan Curve Theorem. The term *Jordan curve* is used for the image C of the circle S^1 under a continuous injective (i.e., one-to-one) map $f\colon S^1 \to \mathbb{R}^2$. In his course of analysis [**62**], Jordan made an attempt to prove that the set $\mathbb{R}^2 \setminus C$ is disconnected and has precisely two path-connected components (this is the *Jordan curve theorem*). His proof was not quite rigorous. The first complete proof of the Jordan theorem was suggested by Veblen [**132**].

We have already proved the Jordan curve theorem in the case where the curve C is a finite-sided polygonal curve (see p. 6). The general Jordan curve theorem can be derived from the piecewise linear version by approximating the curve C by polygons. Such a proof was given in [**128**]. We give another proof, borrowed from [**125**]; it is based on the nonplanarity of the graph $K_{3,3}$ (see Theorem 1.2 on p. 7; recall that the proof of this theorem uses only the piecewise linear version of the Jordan theorem).

Our proof of the Jordan theorem uses Lemmas 1 and 2 proved below; the former asserts that an arbitrary path in the open set $\mathbb{R}^2 \setminus C$ can be replaced by a self-avoiding polygonal curve, and the latter says that if a graph can be drawn in the plane so that its edges be pairwise disjoint paths, then such paths can be replaced by disjoint polygonal curves.

Lemma 1. *Let $U \subset \mathbb{R}^2$ be an open set, and let $\gamma(t)$ be a continuous path between $p = \gamma(0)$ and $q = \gamma(1)$ in U. Then the points p and q can be joined by a finite-sided self-avoiding polygonal curve contained entirely in U.*

Proof. It is easy to replace an arbitrary polygonal curve joining p and q by a self-avoiding polygonal curve; thus, we will not try to make the curve self-avoiding. Let t_0 be the least upper bound for the t such that p can be joined with $\gamma(t)$ by a polygonal curve contained in U. Since U is open, some disk of radius $\delta > 0$ centered at $\gamma(t_0)$ is contained entirely in U. Thus, the case $t_0 < 1$ (as well as the case in which $t_0 = 1$ and p cannot be joined with $\gamma(1)$ by a polygonal curve) cannot occur. □

Lemma 2. *Suppose that there exists a continuous map from a graph to a plane which takes all vertices to different points and the images of two edges can intersect only in the image of their common vertex (if it exists). Then there exists a map with the same properties such that, additionally, the images of edges of this graph are pairwise disjoint polygonal curves in the plane.*

Proof. Each edge (the image of a closed interval under a continuous map) is a compact subset of the plane; therefore, we can choose a positive number ε which is smaller than the distance from any vertex to any edge not incident

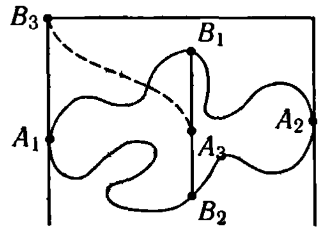

Figure 1. A Jordan curve and the graph $K_{3,3}$

to it and than half the distance between any two vertices. For each vertex, consider the circle of radius ε centered at this vertex. The intersection of this circle with every edge incident to the vertex is compact; hence, on each edge, we can choose two points at distance ε from its end vertices in such a way that the part of the edge between these points is outside the circles. We join the chosen points with the centers of the corresponding circles by straight line segments and successively replace the curves joining these points by disjoint polygonal lines, using the argument from the proof of Lemma 1. □

According to Lemmas 1 and 2, we can assume all paths and arcs encountered in the proof of the Jordan theorem to be finite-sided polygonal curves.

First, we prove that any Jordan curve separates the plane.

Theorem 2.12. *If C is a Jordan curve, then the set $\mathbb{R}^2 \setminus C$ is not path-connected.*

Proof. Let us draw two parallel supporting lines for the curve C and choose two points A_1 and A_2 belonging to C on these lines. On the two arcs of C determined by the points A_1 and A_2, we can choose points B_1 and B_2 so that the (open) interval B_1B_2 will not intersect C (see Figure 1); indeed, each of these two arcs is a compact set, and therefore the intersection of the arc with any straight line parallel to the supporting lines is compact. Take a point A_3 on the interval B_1B_2. Choose a point B_3 on the supporting line passing through A_1 but outside the orthogonal projection of the curve C on this line. If we could join the points A_3 and B_3 by a path avoiding C, then we would obtain an embedding of the graph $K_{3,3}$ in the plane, which does not exist. □

Now, let us prove an auxiliary assertion, which says that a nonclosed arc of the curve does not separate the plane.

Theorem 2.13. *Let A be a simple arc in the plane, i.e., the image of the interval I under an injective continuous map $f\colon I \to \mathbb{R}^2$. Then the set $\mathbb{R}^2 \setminus A$ is connected.*

Proof. Take $x, y \in \mathbb{R}^2 \setminus A$. The set A is compact; therefore, we can choose a positive number d such that the distances from x and y to A are larger than $3d$. The map f is uniformly continuous; hence A can be partitioned into arcs $A_1, \dots, A_k$ (each arc A_i joins points a_i and a_{i+1}) in such a way that the distance from each point a_i to any point of the arc A_i be at most d for $i = 1, \dots, k$. Denote by d' the minimum distance between the points of the arcs A_i and A_j, where $1 \le i \le j-2 \le k-2$. Clearly, $d' \le d$. We partition each arc A_i into arcs $A_{i1}, \dots, A_{ik_i}$ (each arc A_{ij} joins points a_{ij} and $a_{i,j+1}$) in such a way that the distance from a_{ij} to any point of the arc A_{ij} be less than $d'/4$. Let G_i be the graph formed by the edges of the squares centered at the points a_{ij} and such that the sides of all squares are parallel to two fixed straight lines and have length $d'/2$. The graphs G_i and G_j intersect if and only if $|i-j| \le 1$.

The graph $G = G_1 \cup \dots \cup G_k$ separates the plane into connected domains. Precisely one of them is unbounded; let us denote it by F. Each point of the arc A belongs to some of the bounded domains; therefore, A does not intersect F. Thus, it suffices to prove that $x, y \in F$.

Suppose that the point x belongs to some of the bounded domains determined by G. The graph G is 2-connected, and hence it contains a cycle C surrounding the point x. We choose the cycle C so that it belong to the graph $G_i \cup G_{i+1} \cup \dots \cup G_j$ and the difference $j - i$ be minimal. Let us show that $j - i \le 1$. Suppose that $j - i \ge 2$. We can assume that the cycle C has minimum number of edges not belonging to G_{j-1}. Each of the disjoint graphs G_{j-2} and G_j has at least one edge contained in C (and not contained in G_{j-1}). Moreover, the deletion of all edges contained in G_{j-1} violates the connectedness of C. This means that the cycle C has at least two disjoint segments contained in the graph G_{j-1}. These two segments can be joined by a path γ along edges of G_{j-1}. The path γ breaks the cycle C into two cycles. The point x is inside one of these cycles. But each of the cycles has strictly fewer edges not belonging to G_{j-1} than C. This contradicts the minimality assumption.

Thus, the point x belongs to the interior domain determined by the graph $G_i \cup G_{i+1}$. But this is impossible, because x is outside the disk of radius $3d$ centered at a_i, while the graph $G_i \cup G_{j+1}$ is contained in this disk. This contradiction shows that the point x belongs to the unbounded domain determined by G. The point y belongs to the same domain; therefore, x and y can be joined by a path contained in $\mathbb{R}^2 \setminus A$. □

We have already proved that any Jordan curve separates the plane. Now, we can prove the remaining part of the Jordan theorem.

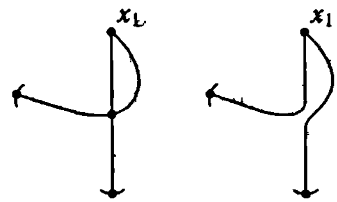

Figure 2. Modification of the arcs

Theorem 2.14. *Any Jordan curve C divides the plane into precisely two path-connected domains, and the boundary of each of these domains is the curve C.*

Proof. Let Ω be one of the path-connected domains into which the curve C separates the plane, and let c be any point on C. Removing an arbitrarily small arc δ containing c from C, we obtain the arc $A = C \setminus \delta$, which does not separate the plane. Therefore, any point $x \in \Omega$ can be joined to a point y contained in the other connected component by a polygonal arc γ avoiding A. The arc γ must intersect C; therefore, it intersects the arc δ. The arc γ has a segment such that it joins x to some point on the arc δ and all of its points (except the point on δ) belong to Ω. Thus, the boundary of Ω contains an everywhere dense subset of C; being closed, it contains the entire curve C.

It remains to prove that the set $\mathbb{R}^2 \setminus C$ cannot have more than two connected components. Suppose that points x_1, x_2, and x_3 belong to three different components Ω_1, Ω_2, and Ω_3 of the set $\mathbb{R}^2 \setminus C$. Let δ_1, δ_2, and δ_3 be pairwise disjoint arcs of the curve C. In the domain Ω_1, the point x_1 can be joined by an arc γ_{1j} to some point of the arc δ_j. Moreover, we can make the arcs γ_{11}, γ_{12}, and γ_{13} intersect only at the point x_1: it suffices to modify them in a neighborhood of their intersection point as shown in Figure 2.

For points x_2 and x_3, we define paths γ_{2i} and γ_{3i} in a similar way. Adding segments of the arcs δ_i to the paths γ_{ij}, where $i, j = 1, 2, 3$, we obtain an embedding of the graph $K_{3,3}$ in the plane, which does not exist. □

2.2. The Whitney–Graustein Theorem. Suppose that $S^1 = \{e^{2\pi i s} : s \in \mathbb{R}\}$ and $\gamma\colon S^1 \to \mathbb{R}^2$ is a smooth closed curve, i.e., $\gamma(s) = (x(s), y(s))$, where x and y are continuously differentiable functions of s and $v(s) = \frac{d\gamma(s)}{ds} \neq 0$ for all $s \in \mathbb{R}$. The *degree*[3] of the smooth curve γ is defined as the number of revolutions made by the vector $v(s)$ as s changes from 0 to 1. Each counterclockwise revolution is counted with the plus sign, and each

[3]The degree of a smooth closed curve is not the same thing as the degree of an algebraic curve.

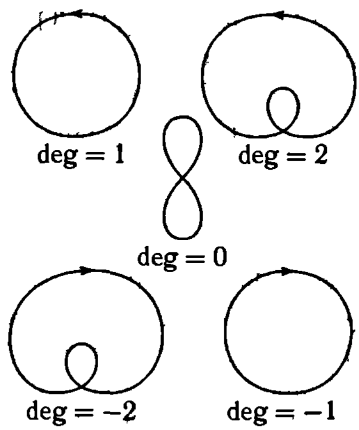

Figure 3. Examples of curves of small degrees

clockwise revolution is counted with the minus sign. Examples of curves of small degrees are shown in Figure 3.

We say that smooth closed curves γ_0 and γ_1 are *regularly homotopic* if there exists a family of smooth closed curves γ_t which continuously depends on $t \in [0,1]$ (the curves γ_0 and γ_1 corresponding to $t = 0$ and 1, respectively, are the given ones). By a continuous dependence on t we mean that the map $(s,t) \mapsto \gamma_t(s)$ from $[0,1] \times [0,1]$ to $\mathbb{R}^2$ is continuous.

Theorem 2.15 (Whitney–Graustein [**146**]). *Curves γ_0 and γ_1 are regularly homotopic if and only if they have equal degrees.*

Proof. Suppose that curves γ_0 and γ_1 are regularly homotopic and N_t is the degree of the smooth curve γ_t for each $t \in [0,1]$. Clearly, N_t continuously depends on t and takes only integer values. Therefore, N_t is constant, and $N_0 = N_1$.

Now, suppose that γ_0 and γ_1 are smooth closed curves of degree N. Using regular homotopies, we can replace the curves γ_0 and γ_1 by curves of length 1 for which $\gamma_0(0) = \gamma_1(0) = (0,0)$ and $\gamma_0'(0) = \gamma_1'(0) = (1,0)$. Thus, we can assume that $s \in [0,1]$ is the natural parameter, i.e., $\|\gamma_0'(s)\| = \|\gamma_1'(s)\| = 1$ for all s.

Let us write the velocity vectors for the curves γ_0 and γ_1 as $v_0(s) = e^{i\varphi_0(s)}$ and $v_1(s) = e^{i\varphi_1(s)}$, where $\varphi_0(0) = \varphi_1(0) = 0$ and $\varphi_0(1) = \varphi_1(1) = 2\pi N$. We set $\varphi_t(s) = (1-t)\varphi_0(s) + t\varphi_1(s)$. Consider the curve $\widetilde{\gamma}_t$ with velocity $v_t(s) = e^{i\varphi_t(s)}$; it is defined by $\widetilde{\gamma}_t(s) = \int_0^s e^{i\varphi_t(\tau)}\,d\tau$. For $t \neq 0,1$, the curve $\widetilde{\gamma}_t$ may not be closed, but it determines the closed curve $\gamma_t(s) = \widetilde{\gamma}_t(s) - s\widetilde{\gamma}_t(1) = \int_0^s e^{i\varphi_t(\tau)}\,d\tau - s\int_0^1 e^{i\varphi_t(\tau)}d\tau$. We must only verify that the curve γ_t is smooth, i.e., $\frac{d}{ds}\gamma_t(1) = \frac{d}{ds}\gamma_t(0)$ and $\frac{d}{ds}\gamma_t(s) \neq 0$. Clearly,

$\frac{d}{ds}\gamma_t(s) = e^{i\varphi_t(s)} - \int_0^1 e^{i\varphi_t(\tau)}d\tau = v_t(s) - \widetilde{\gamma}_t(1)$. The equality of the velocities at $s = 0$ and $s = 1$ follows from the equality $v_t(0) = v_t(1)$, which holds because $\varphi_t(0) = 0$ and $\varphi_t(1) = 2\pi N$. To prove that $v_t(s) \neq \widetilde{\gamma}_t(1)$, it suffices to note that $\|v_t(s)\| = 1$ and $\|\widetilde{\gamma}_t(1)\| < 1$, because $\|\widetilde{\gamma}_t(1)\| = \left|\int_0^1 e^{i\varphi_t(\tau)}d\tau\right| \leq \int_0^1 \left|e^{i\varphi_t(\tau)}\right| d\tau \leq 1$ and the function $e^{i\varphi_t(\tau)}$ is not constant. □

Yet another approach to the proof of the Whitney Graustein theorem is as follows. Slightly moving the curve if necessary, we can assume that it has only finitely many self-intersections. We call a segment ω of a curve γ a *simple loop* if it has the following properties: (i) ω starts and ends at a self-intersection point of γ and (ii) ω has no self-intersections (but it may intersect other parts of γ). It is easy to prove that any smooth curve which has a finite (nonzero) number of self-intersections contains a simple loop. For any simple loop ω of a curve γ, there exists a regular homotopy which changes only ω (and, possibly, the curve γ in a small neighborhood of the vertex of ω) and gives a new simple loop ω' disjoint from γ. The homotopy depends on whether the vertex angle of ω is larger or smaller than 180°. If the angle is smaller than 180°, then the homotopy is constructed directly. If it is larger than 180°, then we first obtain two loops and then delete one of them by changing the curve only in a small neighborhood of the vertex of ω. As a result, we obtain a circle with small loops, external and internal. We can interchange these small loops by dragging them through each other. Moreover, it is easy to construct a regular homotopy which kills pairs of internal and external small loops.

Theorem 2.16 (see [57,138]). *The degree of a smooth closed self-avoiding curve γ is equal to ± 1.*

Proof (Hopf [57]). After applying a regular homotopy, we can assume that the curve γ has length 1 and the map $\gamma\colon S^1 = \{e^{2\pi is}\} \to \mathbb{R}^2$ is such that $\left\|\frac{d\gamma}{ds}\right\| = 1$ for all $s \in [0,1]$. Let T be the triangle determined by the inequalities $0 \leq x \leq y \leq 1$ in the plane with coordinates x and y.

Consider the map $f\colon T \to S^1$ defined by

$$f(x,y) = \begin{cases} \frac{\gamma(y)-\gamma(x)}{\|\gamma(y)-\gamma(x)\|} & \text{if } 0 < y - x < 1, \\ \gamma'(x) & \text{if } x = y, \\ -\gamma'(0) & \text{if } x = 0 \text{ and } y = 1. \end{cases}$$

(Note that if $x = y$, then $\gamma'(x) = \gamma'(y)$, and if $x = 0$ and $y = 1$, then $-\gamma'(0) = -\gamma'(1)$.) For the covering $p\colon \mathbb{R}^1 \to S^1$ given by the formula $p(s) = e^{2\pi is}$, there exists a lifting of f, i.e., a map $F\colon T \to \mathbb{R}^1$ such that $pF = f$. Moreover, $2\pi \deg\gamma = F(1,1) - F(0,0) = [F(1,1) - F(0,1)] + [F(0,1) - F(0,0)]$.

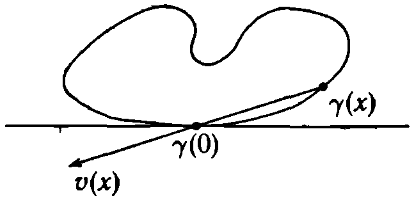

Figure 4. The choice of the point $\gamma(0)$

The difference $F(1,1) - F(0,1)$ corresponds to the angle through which the vector $\gamma(1) - \gamma(x) = \gamma(0) - \gamma(x) = v(x)$ rotates as x varies from 0 to 1 (along the upper side of the triangle T). If $\gamma(0)$ is the point of tangency of γ with some support line (see Figure 4), then this rotation angle is $\pm\pi$ (the sign coincides with that of $\deg \gamma$). The difference $F(0,1) - F(0,0)$ corresponds to the angle through which the vector $\gamma(y) - \gamma(0) = -v(y)$ rotates as y varies from 0 to 1. This rotation angle also equals $\pm\pi$, and its sign again coincides with that of $\deg \gamma$, because the vectors v and $-v$ rotate in the same direction. $\square$

Suppose that $\gamma \colon S^1 \to \mathbb{R}^2$ is a smooth closed curve with finitely many self-intersections and all of its self-intersections are double. Choose a non-self-intersection point x_0 on the curve γ. For the ith self-intersection x_i, we define the number W_i as follows. We go from the point x_0 along the curve γ in the direction of its orientation. When we pass through x_i for the first time, we draw a tangent vector v_1 in the direction of the motion; when we pass through this point for the second time, we draw a second tangent vector v_2. If the frame of reference (v_1, v_2) has negative orientation, then we set $W_i = 1$, and if this frame has positive orientation, then we set $W_i = -1$. The *Whitney number* is defined as $W(\gamma, x_0) = \sum W_i$, where the summation is over all self-intersection points of the curve γ.

Theorem 2.17 (Whitney [**146**]). *The degree* $\deg \gamma$ *and the Whitney number* $W(\gamma, x_0)$ *of any curve* γ *satisfy the relation* $\deg \gamma = W(\gamma, x_0) \pm 1$.

Proof. If the curve γ is self-avoiding, then we can apply Theorem 2.16. Suppose that γ is self-intersecting. Let us go from the point x_0 along the curve γ in the direction of its orientation until we pass twice through some self-intersection point (this must not be the first self-intersection of γ which we encounter). Then, we modify the curve γ as shown in Figure 5. We obtain a curve γ_1 containing x_0 and a self-avoiding curve γ_2.

We show that $\deg \gamma - W(\gamma, x_0) = \deg \gamma_1 - W(\gamma_1, x_0)$, i.e., $\deg \gamma - \deg \gamma_1 = W(\gamma, x_0) - W(\gamma_1, x_0) = W_i$, where $W_i = \pm 1$ is the number assigned to the cross destroyed under the reconstruction. Applying a regular homotopy, we

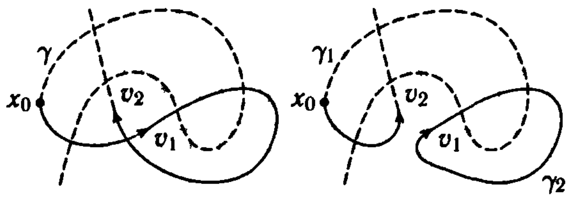

Figure 5. The reconstruction of the curve γ

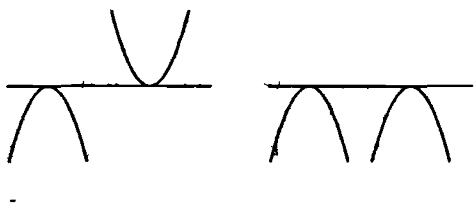

Figure 6. Internal and external double tangents

can transform γ in such a way that the angle between the vectors v_1 and v_2 become arbitrarily small. Clearly, $\deg \gamma = \deg \gamma_1 + \deg \gamma_0$. It is also easy to show that $\deg \gamma_0 = W_i$. $\square$

2.3. Double Points, Double Tangents, and Inflection Points. Let γ be a closed differentiable curve in the plane $\mathbb{R}^2$ composed of finitely many convex arcs not tangent to each other at interior points. Then γ has a finite number $D(\gamma)$ of self-intersection points and a finite number $F(\gamma)$ of inflection points. We assume that all self-intersections of γ are double points, i.e., the curve γ has no points through which more than two of its branches pass. We also assume that γ has no triple tangents, i.e., a straight line is tangent to γ at no more than two different points. There are two types of double tangents, internal and external (see Figure 6). We denote the numbers of external and internal double tangents by $I(\gamma)$ and $II(\gamma)$, respectively.

Theorem 2.18 (see [34]). $II(\gamma) - I(\gamma) = D(\gamma) + \frac{F(\gamma)}{2}$.

Proof. Let us orient the curve γ, i.e., specify a traverse direction. For each point $a \in \gamma$, consider the straight line l tangent to γ at a. The point a breaks l into the ray l_+ whose direction coincides with the traverse direction of γ and the ray l_- whose direction is opposite to the traverse direction of γ. Let us move the point a along γ in the positive direction until it traverses the entire curve once. For each position of a, we define N_+ as the number of intersection points (different from a) of the ray l_+ with γ. The number N_- is defined similarly. The numbers N_+ and N_- change only when a passes through a double point or an inflection point and when the line l becomes a

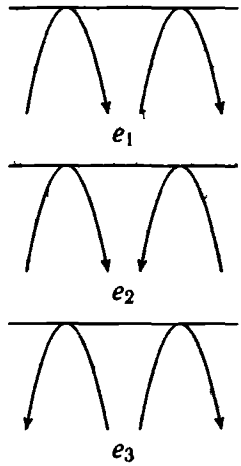

Figure 7. The three types of external double tangents

double tangent, i.e., a passes through one of the tangency points of a double tangent with the curve γ.

When the point a passes through a double point or an inflection point, N_+ decreases by 1 and N_- increases by 1.

For an oriented curve, the external tangents can be of three types, depending on the directions of the tangent lines at the points of tangency: they may have the same direction, be directed toward each other, or be directed away from each other (see Figure 7). We denote the number of external tangents of each of these three types by e_1, e_2, and e_3, respectively. Similarly, for the internal tangents, we use the symbols i_1, i_2, and i_3.

Consider the two points of tangency between a tangent line of type e_1 (i_1) and the curve γ. As a passes through one of these points, N_+ increases (respectively, decreases) by 2, and as a passes through the other point, N_- decreases (respectively, increases) by 2.

For tangents of type e_2 (i_2), as a passes through any of the two points of tangency, N_+ increases (decreases) by 2.

For tangents of type e_3 (i_3), as a passes through any of the two points of tangency, N_- decreases (increases) by 2.

After a complete traverse of the curve, N_+ increases by $2e_1 + 4e_2$ and decreases by $2i_1 + 4i_2 + 2D(\gamma) + F(\gamma)$ (each double point is passed twice, and each inflection point is passed once). Thus, N_+ does not change, and therefore $2e_1 + 4e_2 = 2i_1 + 4i_2 + 2D(\gamma) + F(\gamma)$. Applying the same argument to N_-, we see that $2e_1 + 4e_3 = 2i_1 + 4i_3 + 2D(\gamma) + F(\gamma)$. Summing these equalities, we obtain $4(e_1+e_2+e_3)-4(i_1+i_2+i_3)=4D(\gamma)+2F(\gamma)$, i.e., $4II(\gamma)-4I(\gamma)=4D(\gamma)+2F(\gamma)$. □

A similar assertion can be proved for generic closed polygons with finitely many vertices (a polygon is said to be generic if no three of its vertices are collinear and no three edges intersect in one point); see [**11**, **35**].

The equality

$$II(\gamma)-I(\gamma)=D(\gamma)+\frac{F(\gamma)}{2} \tag{1}$$

is a necessary condition for the existence of a curve γ having $II(\gamma)$ external double tangents, $I(\gamma)$ internal double tangents, $D(\gamma)$ double points, and $F(\gamma)$ inflection points. But this condition is not sufficient. For example, if $F(\gamma)=0$, then the curve γ must be convex; therefore, $I(\gamma)=II(\gamma)=0$, while equality (1) implies only $I(\gamma)=II(\gamma)$. If $F(\gamma)\neq 0$, then condition (1) is not only necessary but also sufficient for the existence of such a curve (see [**47**]). But if $F(\gamma)=0$, then the additional condition $I(\gamma)\leq D(\gamma)^2-D(\gamma)$ must hold, and the number $I(\gamma)$ must be even (see [**95**]).

3. The Brouwer Fixed Point Theorem and Sperner's Lemma

3.1. The Brouwer Fixed Point Theorem. We use the following notation:

$D^n=\{x\in\mathbb{R}^n\colon \|x\|\leq 1\}$ (the unit disk, or ball);

$S^n=\{x\in\mathbb{R}^n\colon \|x\|=1\}\subset D^n$ (the unit sphere).

Let $f\colon X\to X$ be a map. A point $x\in X$ is called a *fixed point* of f if $f(x)=x$.

Theorem 2.19 (Brouwer). *Any continuous map $f\colon D^n\to D^n$ has a fixed point.*

Remark 2.3. Brouwer proved this theorem in [**23**]. Earlier, assertions equivalent to the fixed point theorem were proved by Henri Poincaré [**98**] and the Latvian mathematician Bohl [**15**]. But the theorem is most commonly known as the *Brouwer fixed point theorem.*

Let $A\subset X$. A continuous map $r\colon X\to A$ is called a *retraction* if $r\restriction A=\mathrm{id}_A$, i.e., $r(a)=a$ for any point $a\in A$. If there exists a retraction $r\colon X\to A$, then A is a *retract* of X.

Problem 24. Prove that A is a retract of a space X if and only if any continuous map $f\colon A\to Y$ can be extended to X.

Problem 25. Prove that if A is a retract of X and any continuous self-map of X has a fixed point, then any continuous self-map of A has a fixed point as well.

Theorem 2.20. *There exists no retraction $r\colon D^n \to S^{n-1}$.*

It is easy to verify that the Brouwer theorem is equivalent to Theorem 2.20. Indeed, suppose that $f\colon D^n \to D^n$ is a fixed-point-free continuous map. For each point $x \in D^n$, consider the ray from $f(x)$ passing through x. Let $r(x)$ be the intersection point of this ray with the sphere S^{n-1}. Clearly, r is a retraction of D^n onto S^{n-1}.

Now, suppose that $r\colon D^n \to S^{n-1}$ is a retraction. If $i\colon S^{n-1} \to S^{n-1}$ is a fixed-point-free map (say, $i(x) = -x$), then the map $ir\colon D^n \to S^{n-1} \subset D^n$ has no fixed points.

Theorem 2.20 is equivalent also to the following assertion.

Theorem 2.21. *Let $v(x)$ be a continuous vector field on D^n such that $v(x) = x$ for all $x \in S^{n-1}$. Then $v(x) = 0$ for some $x \in D^n$.*

Indeed, if $r\colon D^n \to S^{n-1}$ is a retraction, then the formula $v(x) = r(x)$ defines a nowhere vanishing vector field on D^n. If $v(x)$ is a vector field on D^n such that $v(x) = x$ for all $x \in S^{n-1}$ and $v(x) \neq 0$ for $x \in D^n$, then $x \mapsto \frac{v(x)}{\|v(x)\|}$ is the required retraction.

There exist several different proofs of Theorems 2.19 2.21. In most cases, it is more convenient to prove the nonexistence of a retraction of the disk D^n to the sphere S^{n-1}. In this section, three such proofs are given; we consider only smooth maps.

We can pass to continuous maps by using their smooth approximations. Indeed, suppose we have a continuous retraction $r\colon D^n \to S^{n-1}$. Let us show that there exists a smooth retraction $\tilde{r}\colon D^n \to S^{n-1}$. If $\|x\| = 1$, then $r(x) = x$. Therefore, for any $\varepsilon_1 > 0$, we can find $\delta > 0$ such that $\|r(x) - x\| \leq \varepsilon_1$ for $1 - \delta \leq \|x\| \leq 1$. By the Weierstrass theorem, there exist a smooth map $f\colon \mathbb{R}^n \to \mathbb{R}^n$ such that $\|f(x) - (r(x) - x)\| \leq \varepsilon_1$ for $\|x\| \leq 1$ and a smooth function $\psi(t)$ such that $0 \leq \psi(t) \leq 1$ for $0 \leq t \leq 1$, $\psi(1) = 0$, and $1 - \varepsilon_2 \leq \psi(t)$ for $t^2 \leq 1 - \delta$. We set $g(x) = x + \varphi(x)f(x)$, where $\varphi(x) = \psi(\|x\|^2)$. If $\|x\| \leq 1 - \delta$, then

$$\begin{aligned}
\|g(x)\| &= \|x + \varphi(x)f(x)\| \\
&= \|r(x) + \varphi(x)[f(x) - r(x) + x] + (\varphi(x) - 1)(r(x) - x)\| \\
&\geq \|r(x)\| - \varphi(x)\|f(x) - r(x) + x\| - (1 - \varphi(x))\|r(x) - x\| \\
&\geq 1 - 1 \cdot \varepsilon_1 - \varepsilon_2 \cdot 2 = 1 - \varepsilon_1 - 2\varepsilon_2.
\end{aligned}$$

If $1 - \delta \le \|x\| \le 1$, then

$$\begin{aligned}\|g(x)\| &= \|x + \varphi(x)f(x)\| \\ &= \|x + \varphi(x)[f(x) - r(x) + x] + \varphi(x)(r(x) - x)\| \\ &\ge \|x\| - \varphi(x)\|f(x) - r(x) + x\| - \varphi(x)\|r(x) - x\| \\ &\ge 1 - \delta - 1 \cdot \varepsilon_1 - 1 \cdot \varepsilon_1 = 1 - \delta - 2\varepsilon_1.\end{aligned}$$

Note that $\delta \to 0$ as $\varepsilon_1 \to 0$. Hence we can assume that $\varepsilon_1, \varepsilon_2, \delta \le 1/4$. We then have $\|g(x)\| \ge 1/4 > 0$ for all $x \in D^n$. If $\|x\| = 1$, then $\varphi(x) = 0$ and $g(x) = x$. The required retraction $\tilde{r}\colon D^n \to S^{n-1}$ is defined by $\tilde{r}(x) = g(x)/\|g(x)\|$.

Below, we give three proofs for the nonexistence of a retraction of the disk to the sphere; they use some facts that have not been mentioned in this book so far (we shall prove them further on). A quite elementary proof of the Brouwer theorem, which is equivalent to the nonexistence of a retraction of the disk onto the sphere, is given on p. 81.

3.1.1. *The first proof of the nonexistence of a retraction of the disk onto the sphere* (Hirsch [53]). Suppose that $r\colon D^n \to S^{n-1}$ is a smooth retraction and $a \in S^{n-1}$ is a regular value of r. Then $r^{-1}(a)$ is a union of one-dimensional submanifolds, and its boundary is contained in S^{n-1}. The set $r^{-1}(a)$ is compact because it is a closed subset of a compact space. Any compact one-dimensional manifold is either a circle or a closed interval; therefore, the boundary of $r^{-1}(a)$ consists of an even number of points. But $r^{-1}(a)$ and S^{n-1} intersect in precisely one point, a. This contradiction completes the proof. □

Remark 2.4. It can be proved in a similar way that if M^n is a compact manifold with nonempty boundary W^{n-1}, then there exists no retraction $r\colon M^n \to W^{n-1}$.

3.1.2. *The second proof of the nonexistence of a retraction of the disk onto the sphere.* Suppose that $r\colon D^n \to S^{n-1}$ is a smooth retraction. Consider the differential $(n-1)$-form $\omega = x_1 dx_2 \wedge \cdots \wedge dx_n$. Its differential is the n-form $d\omega = dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n$. These two forms are defined on the entire space $\mathbb{R}^n$; therefore, they can be regarded both as forms on D^n and as forms on S^{n-1}. By Stokes' theorem,

$$\begin{aligned}\int_{S^{n-1}} \omega &= \int_{r(S^{n-1})} \omega = \int_{S^{n-1}} r^*\omega = \int_{D^n} dr^*\omega \\ &= \int_{D^n} r^* d\omega = \int_{r(D^n)} d\omega = \int_{S^{n-1}} d\omega = 0.\end{aligned}$$

The last equality holds because there are no nonzero n-forms on the $(n-1)$-manifold S^{n-1}, i.e., the n-form $d\omega$ vanishes on S^{n-1}.

On the other hand, the same Stokes' theorem implies

$$\int_{S^{n-1}} \omega = \int_{D^n} d\omega = \text{volume}(D^n) > 0. \qquad \square$$

3.1.3. *The third proof of the nonexistence of a retraction of the disk onto the sphere* (see [111]). Suppose that there exists a continuously differentiable retraction $f\colon D^n \to S^{n-1}$. For $x \in D^n$ and $0 \le t \le 1$, we set

$$\begin{aligned} g(x) &= f(x) - x, \\ f_t(x) &= x + tg(x) = (1-t)x + tf(x). \end{aligned}$$

The continuous differentiability of g implies the existence of a positive constant c such that $\|g(x) - g(y)\| \le c\|x - y\|$ for all $x, y \in D^n$. The map f_t is injective for $0 \le t < 1/c$. Indeed, if $f_t(x) = f_t(y)$, then $\|x - y\| = \|tg(x) - tg(y)\| \le tc\|x - y\|$. Therefore, for $0 \le t < 1/c$, we have $\|x - y\| = 0$.

The partial derivatives of g are uniformly bounded; hence the Jacobian matrix

$$\left(\frac{\partial f_t}{\partial x_1}, \dots, \frac{\partial f_t}{\partial x_n}\right) = I_n + \left(\frac{\partial g}{\partial x_1}, \dots, \frac{\partial g}{\partial x_n}\right) \tag{1}$$

is invertible for small t. By the inverse function theorem, the map f_t with $t \le t_0$ takes $\operatorname{int} D^n$ (the interior of D^n) to some open set G_t. Take $e \in D^n \setminus G_t$. Let b be the intersection point of the straight line segment joining e to an arbitrary point of G_t with the boundary of G_t. The set $f_t(D^n)$ is compact; therefore, $b = f_t(x)$ for some $x \in D^n$. Since $b \notin G_t = f_t(\operatorname{int} D^n)$, we have $x \notin \operatorname{int} D^n$, i.e., $x \in S^{n-1}$. Hence $b = x$ and $e = b = x \in S^{n-1}$. Thus, f_t maps $\operatorname{int} D^n$ onto $\operatorname{int} D^n$. Moreover, f_t bijectively maps S^{n-1} to S^{n-1}, and, as shown above, it injectively maps D^n to D^n. Therefore, f_t is a bijection from D^n to D^n (for $t \le t_0$).

Consider the integral

$$I(t) = \int_{D^n} \det\left(\frac{\partial f_t}{\partial x_1}\right) = \int \cdots \int_{D^n} \det\left(\frac{\partial f_t}{\partial x_1}, \dots, \frac{\partial f_t}{\partial x_n}\right) dx_1 \cdots dx_n.$$

For $0 \le t \le t_0$, this integral is equal to the volume of the unit ball D^n. Formula (1) shows that $I(t)$ is a polynomial in t. Therefore, $I(t)$ is a positive constant for $0 \le t \le 1$.

On the other hand, $f_1(x) = f(x) \in S^{n-1}$; hence $f_1(x) \cdot f_1(x) = 1$, which means that $\frac{\partial f_1}{\partial x_i} \cdot f_1 = 0$ for $i = 1, \dots, n$. The vectors $\frac{\partial f_1}{\partial x_i}$ belong to the same hyperplane; therefore, they are linearly dependent. Thus, $\det\left(\frac{\partial f_t}{\partial x_1}, \dots, \frac{\partial f_t}{\partial x_n}\right) = 0$, and $I(1) = 0$. This contradiction completes the proof. $\square$

Let $\mathcal{F}(f)$ be the set of fixed points of a continuous map $f\colon D^n \to D^n$. By the Brouwer theorem, this set is nonempty. Clearly, it is also closed. It turns out that any nonempty closed subset of D^n is the set of fixed points for some continuous map.

Theorem 2.22 (see [109]). *Let $F \subset D^n$ be a nonempty closed set. Then there exists a continuous map $f\colon D^n \to D^n$ for which $\mathcal{F}(f) = F$.*

Proof. Setting $d(x, F) = \inf_{y\in F} \|x - y\|$ for each point $x \in D^n$, we obtain a continuous function on D^n (the distance from x to the closed set F). Now, we define the map $f\colon D^n \to D^n$ as

$$f(x) = \begin{cases} x - d(x, F)\frac{x-x_0}{\|x-x_0\|} & \text{for } x \neq x_0, \\ x_0 & \text{for } x = x_0. \end{cases}$$

This map f is continuous, and $\mathcal{F}(f) = F$. □

The Brouwer theorem and the nonexistence of a retraction of the disk to the sphere have various consequences. Below, we give several examples.

Theorem 2.23. *Suppose that $f\colon D^n \to D^n$ and $f(S^{n-1}) \subset S^{n-1}$.*

(a) *If $f \restriction S^{n-1}$ is the identity map, then* $\operatorname{Im} f = D^n$;

(b) *If $f \restriction S^{n-1}$ has no fixed points, then* $\operatorname{Im} f = D^n$.

Proof. Suppose that $\operatorname{Im} f \neq D^n$. Then there exists a point $O \in D^n \setminus \operatorname{Im} f$. Let $r\colon D^n \setminus O \to S^{n-1}$ be a projection of $D^n \setminus O$ onto S^{n-1} from O. The map r is a retraction. Since $O \notin \operatorname{Im} f$, $rf\colon D^n \to S^{n-1}$ is well defined.

(a) The map rf is a retraction, which is impossible.

(b) By the Brouwer theorem, the map $rf\colon D^n \to S^{n-1} \subset D^n$ has a fixed point. But $\operatorname{Im}(rf) \subset S^{n-1}$, and rf has no fixed points on S^{n-1}. □

Theorem 2.24. *Suppose that a continuous path α joins two points on opposite sides of a rectangle, and a continuous path β joins two points on the two other opposite sides of the same rectangle. If both paths α and β lie inside the rectangle, then they are disjoint.*

Proof. Suppose that $\alpha(s) = (\alpha_1(s), \alpha_2(s))$ and $\beta(t) = (\beta_1(t), \beta_2(t))$, where $s, t \in [-1, 1]$. It suffices to prove the required assertion for the square I^2 given in the plane with coordinates x_1 and x_2 by the inequalities $|x_i| \leq 1$, where $i = 1, 2$. Thus, we can assume that $\alpha_1(\varepsilon) = \varepsilon$ and $\beta_2(\varepsilon) = \varepsilon$ for $\varepsilon = \pm 1$.

Suppose that $\alpha(s) \neq \beta(t)$ for all $s, t \in [-1, 1]$. Let

$$N(s, t) = \max_{i=1,2}\{|\alpha_i(s) - \beta_i(t)|\}.$$

Consider the map $F\colon I^2 \to I^2$ defined by

$$F(s,t) = \frac{1}{N(s,t)}\big(\beta_1(t) - \alpha_1(s), \alpha_2(s) - \beta_2(t)\big).$$

The square I^2 is homeomorphic to the disk D^2; therefore, according to the Brouwer theorem, F has a fixed point (s_0, t_0).

The image of F consists of points of the form $(\pm 1, t)$ or $(s, \pm 1)$; therefore, either $s_0 = \pm 1$ or $t_0 = \pm 1$. Clearly,

$$N(s,t)F(\pm 1,t) = \big(\beta_1(t) \mp 1,\ \alpha_2(\pm 1) - \beta_2(t)\big),$$
$$N(s,t)F(s,\pm 1) = \big(\beta_1(\pm 1) - \alpha_1(s),\ \alpha_2(s) \mp 1\big).$$

By assumption, the paths α and β lie inside the square I^2; hence $|\beta_1(t)| \leq 1$ and $|\alpha_2(s)| \leq 1$. Thus, ± 1 cannot have the same sign as $\beta_1(t) \mp 1$ or $\alpha_2(s) \mp 1$. This contradicts the obvious inequality $N(s,t) > 0$. □

3.2. The Jordan Curve Theorem as a Corollary of the Brouwer Theorem. In this section, we derive the Jordan curve theorem from the Brouwer fixed point theorem, following [**79**]. Recall that we have already given one proof of the Jordan theorem (see pp. 63–66).

To prove the Jordan curve theorem, we need Theorem 2.24, which we have proved by using the Brouwer theorem. We also need the Tietze extension theorem for continuous maps (Theorem 2.4 on p. 57).

First, note that if C is a Jordan curve, then the set $\mathbb{R}^2 \setminus C$ has precisely one unbounded connected component. This follows from the boundedness of C. It is also clear that any connected component of $\mathbb{R}^2 \setminus C$ is path-connected and open.

Step 1. The set $\mathbb{R}^2 \setminus C$ has a bounded connected component.

Proof. The set C is compact; hence there are two points $a, b \in C$ such that the distance between them is maximal. We can assume that $a = (-1, 0)$ and $b = (1, 0)$. Then the rectangle R given by the inequalities $|x| \leq 1$ and $|y| \leq 2$ contains the entire curve C, and the intersection of its boundary with C consists of precisely two points, a and b (see Figure 8).

The points a and b are the midpoints of two sides of the rectangle R. Let n and s be the midpoints of the two other sides of this rectangle. According to Theorem 2.24, the segment $[n, s]$ intersects the curve C. Let l be the intersection point nearest to n. The points a and b divide C into two arcs. We denote the arc containing l by C_n and the other arc by C_s. Let m be the point of the set $C_n \cap [n, s]$ farthest from n. Then the segment $[m, s]$ intersects C_s; otherwise, the path going first from n to l along a straight line, next from l to m along C_n, and, finally, from m to s along a straight

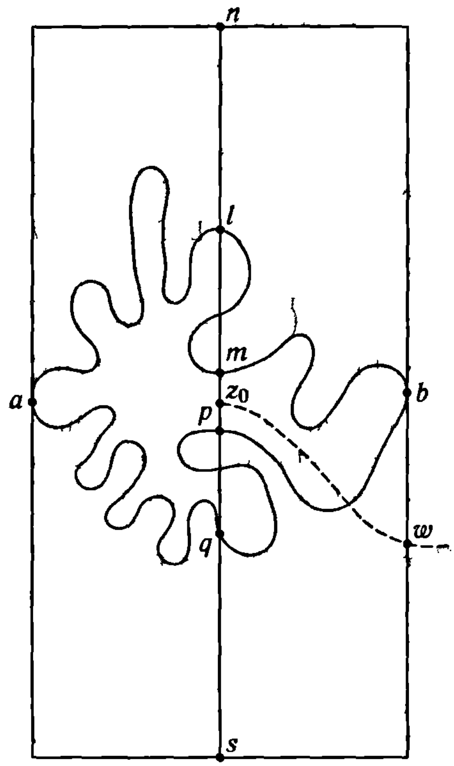

Figure 8. **A Jordan curve in a rectangle**

line does not intersect the arc C_s, which contradicts Theorem 2.24. Let p be the point of $C_s \cap [m, s]$ farthest from s, and let z_0 be the midpoint of $[m, p]$.

We show that the connected component of $\mathbb{R}^2 \setminus C$ containing z_0 is bounded. Suppose that z_0 can be joined by a path γ to a point outside the rectangle R. Let w be the first point at which γ meets the boundary of R. Slightly moving the path γ if necessary, we can assume that the point w is different from a and b. Suppose for definiteness that w is in the bottom part of R (i.e., it is closer to s than to n). Consider the path from n to s which goes from n to l along a straight line, from l to m along the arc C_n, from m to z_0 along a straight line, from z_0 to w along γ, and finally, from w to s along the boundary of R, missing a and b. This path does not intersect the arc C_s, which contradicts Theorem 2.24. □

Step 2. The boundary of each connected component of $\mathbb{R}^2 \setminus C$ coincides with C.

Proof. Suppose that U is a connected component of $\mathbb{R}^2 \setminus C$, $\overline{U}$ is the closure of U, and $\partial U = \overline{U} \setminus U$ is the boundary of U. None of the points of

∂U can belong to another connected component W, because W is open and $U \cap W = \varnothing$. Thus, $\partial U \subset C$. Suppose that $\partial U \neq C$. Then ∂U is contained in some arc A of C. We show that this is impossible.

According to step 1, $\mathbb{R}^2 \setminus C$ has a bounded connected component; take a point z_0 in this component. If U is bounded itself, then we choose z_0 in U. Let D^2 be a disk centered at z_0 and containing C, and let $S^1 = \partial D^2$ be its boundary circle. Then S^1 is contained entirely in the unbounded component of $\mathbb{R}^2 \setminus C$. The arc A is homeomorphic to the unit interval $[0, 1]$; therefore, by the Tietze theorem, the identity map $A \to A$ can be extended to a continuous map $f\colon D^2 \to A$. We define a map $g\colon D^2 \to D^2$ as follows. If the component U is bounded, then we set

$$g(z) = \begin{cases} f(z) & \text{for } z \in \overline{U}, \\ z & \text{for } z \in D^2 \setminus U. \end{cases}$$

If the component U is unbounded, then we set

$$g(z) = \begin{cases} z & \text{for } z \in \overline{U}, \\ f(z) & \text{for } z \in D^2 \setminus U. \end{cases}$$

The intersection of the closed sets $\overline{U}$ and $D^2 \setminus U$ is contained in A, and the restriction of f to A is the identity map; therefore, the map g is continuous and well defined. It is easy to verify that z_0 does not belong to the image of g and the restriction of g to S^1 is the identity map. Hence the composition of g and the projection of D^2 onto S^1 from the point z_0 gives a retraction of D^2 onto S^1, which cannot exist. □

Step 3. The set $\mathbb{R}^2 \setminus C$ has precisely one bounded connected component.

Proof. We use the notation of Step 1 (see Figure 8). Suppose that in addition to the connected component U containing z_0, $\mathbb{R}^2 \setminus C$ has one more bounded connected component W. Clearly, $W \subset R$. Let q be the point of $C_s \cap [n, s]$ nearest to s. Consider the path β which goes from n to l along a straight line, from l to m along the arc C_s, from m to p along a straight line, from p to q along the arc C_s, and finally, from q to s along a straight line. Clearly, β avoids W and the points a and b. Choose neighborhoods of a and b disjoint from β. According to Step 1, the closure of W contains a and b; therefore, the chosen neighborhoods contain points a_1 and b_1 belonging to W. Let us draw a straight line segment from a to a_1, extend it from a_1 to b_1 along a path entirely contained in W, and finally, join b_1 to b by a straight line segment. The path thus obtained does not intersect β, which contradicts Theorem 2.24. □

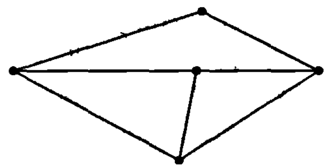

Figure 9. A partition which is not a triangulation

3.3. Sperner's Lemma. We have already given three proofs of the Brouwer fixed point theorem. This theorem has many other proofs. For example, it can be derived from Sperner's lemma, which has various applications. Sperner himself used this lemma to obtain a new proof of another well-known theorem of Brouwer, on the invariance of dimension. The Brouwer fixed point theorem was derived from Sperner's lemma in [68].

Let P^n be a convex polyhedron in $\mathbb{R}^n$. A partition of P^n into n-simplices is called a *triangulation* if the intersection of any two simplices is a common face for these simplices (the empty set is regarded as a face of dimension -1). For example, the partition shown in Figure 9 is not a triangulation.

Suppose that the vertices of a k-simplex are labeled by $0, 1, \ldots, k$. We say that the simplex has *complete set of labels*, or is *completely labeled*, if each of the numbers $0, 1, \ldots, k$ is the label of some vertex (in this case, it is the label of precisely one vertex).

We derive Sperner's lemma from the following theorem, which sometimes is also called Sperner's lemma.

Theorem 2.25. *Suppose that all vertices of a triangulation of a convex polyhedron P^n are labeled by the numbers $0, 1, \ldots, n$. Then the number of completely labeled n-simplices in the triangulation of P^n is odd if and only if the number of completely labeled $(n-1)$-simplices in the triangulation of the boundary of P^n is odd.*

Proof. Consider an n-simplex one of whose faces is a completely labeled $(n-1)$-simplex. If the vertex opposite to this face is labeled by n, then the n-simplex has precisely one completely labeled $(n-1)$-face, and if the opposite vertex is labeled by one of the numbers $0, 1, \ldots, n-1$, then it has precisely two such faces. Therefore, the number of completely labeled n-simplices is congruent modulo 2 to the number of pairs consisting of an n-simplex and its $(n-1)$-face with complete set of labels. Moreover, each completely labeled $(n-1)$-simplex contained in the boundary of P^n belongs to precisely one n-simplex, and each of the other $(n-1)$-simplices belongs to precisely two n-simplices. Hence the number of pairs under consideration is congruent modulo 2 to the number of completely labeled $(n-1)$-simplices in the triangulation of the boundary. ☐

Theorem 2.26 (Sperner's lemma [**121**]). *Suppose that the vertices of a triangulation of a completely labeled n-simplex are labeled in such a way that if a vertex of the triangulation belongs to some face of the initial simplex, then the label of this vertex coincides with that of one of the vertices of the face. Then the triangulation contains a completely labeled n-simplex. Moreover, the number of such simplices is odd.*

Proof. Any odd number is different from zero; therefore, it suffices to prove that the number of completely labeled n-simplices is odd. According to Theorem 2.25, we must show that the number of completely labeled $(n-1)$-simplices on the boundary is odd. The assumption about the labeling of the triangulation implies that any completely labeled $(n-1)$-simplex on the boundary belongs to a completely labeled $(n-1)$-face of the initial simplex. Therefore, the validity of Sperner's lemma for n-simplices follows from that for $(n-1)$-simplices. For $n = 0$, the lemma is trivially true. □

The proof of the Brouwer theorem we give here uses barycentric coordinates. Take a point X in a simplex $A_0 \dots A_n$. The *barycentric coordinates* of X with respect to this simplex are $(x_0, \dots, x_n)$, where x_i is the ratio of the volume of the simplex $XA_0 \dots A_{i-1}A_{i+1} \dots A_n$ to the volume of $A_0 \dots A_n$. Clearly, the numbers $x_0, \dots, x_n$ are nonnegative, and their sum is equal to 1. The x_i coordinate is proportional to the distance from X to the face $A_0 \dots A_{i-1}A_{i+1} \dots A_n$; therefore, the barycentric coordinates of the point X uniquely determine its position.

Proof of the Brouwer theorem. The unit disk D^n is homeomorphic to the n-simplex Δ^n; therefore, it suffices to prove that any continuous map $f\colon \Delta^n \to \Delta^n$ has a fixed point. We label the points of Δ^n as follows. Let $(x_0, \dots, x_n)$ be the barycentric coordinates of a point $X \in \Delta^n$, and let $(y_0, \dots, y_n)$ be the barycentric coordinates of its image under the map f. For the label of X we take the smallest number j for which $y_j \le x_j \ne 0$. Note that the labels of the vertices of any triangulation satisfy the conditions of Sperner's lemma. Indeed, if a point belongs to a face $A_{i_0} \dots A_{i_k}$, then the only nonzero barycentric coordinates of this point are those with numbers $i_0, \dots, i_k$.

Let M be the barycenter of the simplex Δ^n, and let M_i be the barycenter of its ith face. Consider the partition of Δ^n into simplices such that the vertices of each smaller simplex are M, M_i, and some $n-1$ vertices of the ith face ($i = 0, 1, \dots, n$). This partition is called the *barycentric subdivision* of the simplex Δ^n. If d is the maximum edge length of Δ^n, then the length of any edge of a simplex from the barycentric subdivision is at most $\frac{n}{n+1}d$.

Indeed, the vertices of the barycentric subdivision have the form

$$v_1, \quad \frac{v_1+v_2}{2}, \quad \frac{v_1+v_2+v_3}{3}, \quad \dots, \quad \frac{v_1+\dots+v_{n+1}}{n+1},$$

where $v_1, \dots, v_{n+1}$ are the vertices of Δ^n. Clearly,

$$\begin{aligned} &\frac{v_1+\dots+v_{p+q}}{p+q} - \frac{v_1+\dots+v_p}{p} \\ &\quad = \frac{q}{p+q}\left(\frac{v_{p+1}+\dots+v_{p+q}}{q} - \frac{v_1+\dots+v_p}{p}\right) \\ &\quad = \frac{q}{p+q}(a-b), \end{aligned}$$

where a and b are some points of Δ^n. It remains to note that

$$\frac{q}{p+q} \le \frac{p+q-1}{p+q} = 1 - \frac{1}{p+q} \le \frac{n}{n+1},$$

because $p+q \le n+1$.

According to Sperner's lemma, among the n-simplices of the first barycentric subdivision there is a simplex with complete set of labels. Take an arbitrary point X_1 in this simplex. Now, consider the *second barycentric subdivision*, which consists of the barycentric subdivisions of all simplices of the first barycentric subdivision. In the second barycentric subdivision, we choose a completely labeled n-simplex and take any point X_2 in it. In a similar way, we choose points X_3, X_4, and so on. The infinite sequence $\{X_i\}$ has a convergent subsequence $\{X_{i_k}\}$. Let us show that the point $X = \lim_{k\to\infty} X_{i_k}$ is fixed.

Let $(x_{0k}, \dots, x_{nk})$ and $(y_{0k}, \dots, y_{nk})$ be the barycentric coordinates of the point X_{i_k} and its image, and let $(x^l_{0k}, \dots, x^l_{nk})$ and $(y^l_{0k}, \dots, y^l_{nk})$ be the barycentric coordinates of the vertices of the simplex containing X_{i_k} and their images. The simplices under consideration have complete sets of labels; therefore, for any $j = 0, \dots, n$, we have $y^l_{jk} \le x^l_{jk}$ for some l. The edge length of the i_kth simplex tends to zero as $k \to \infty$; hence $\lim_{k\to\infty} x^l_{jk} = \lim_{k\to\infty} x_{jk} = x_j$, where $(x_0, \dots, x_n)$ are the barycentric coordinates of X. Thus, if $(y_0, \dots, y_n)$ are the barycentric coordinates of the image of X, then $y_j \le x_j$ for $j = 0, 1, \dots, n$. But $\sum x_j = 1 = \sum y_j$; hence $x_j = y_j$ for $j = 0, 1, \dots, n$. This means that the point X is fixed. □

Below, we give yet another proof of Sperner's lemma; this proof has also turned out to be useful in obtaining other combinatorial theorems.

The constructive proof of Sperner's lemma [28]. First, we prove the required assertion for the 1-simplex (interval). If two neighboring vertices of the triangulation are labeled by the same number, then one of them can be removed. This does not change the number of completely labeled intervals. After removing all such vertices, we obtain a partition of the interval with labels 01010...101. The number of completely labeled intervals equals $2k-1$, where k is the number of zeros (it coincides with the number of ones).

Now, suppose that $n \geq 1$ and the required assertion is proved for the n-simplices. Given a triangulation of an $(n+1)$-simplex, consider all completely labeled n-simplices in this triangulation. For each of the $(n+1)$-simplices of this triangulation, one of the following three cases is possible:

(0) one of the numbers $0, 1, \ldots, n$ is missing among the labels of vertices; in this case, the number of completely labeled n-faces is 0;

(1) the simplex is completely labeled; in this case, the number of completely labeled n-faces is 1;

(2) the vertices of the simplex are labeled by $0, 1, \ldots, n$, and one of these labels occurs twice; in this case, the number of completely labeled n-faces is 2.

In case 2, we join the barycenters of the completely labeled n-faces by a straight line segment. In case 1, we mark the barycenter of the completely labeled n-face. As a result, we obtain several pairwise disjoint polygonal curves (some of them may degenerate into base points). The endpoint of any such line is either a base point (it corresponds to a completely labeled $(n+1)$-simplex) or the barycenter of a completely labeled n-face contained in an n-face of the initial simplex (the latter face is necessarily completely labeled). Thus, the parity of the number of completely labeled $(n+1)$-simplices coincides with that of the number of completely labeled n-faces contained in a (completely labeled) face of the initial simplex. By the induction hypothesis, this number is odd. □

In the constructive proof, we can track the orientations of simplices and obtain the following refinement of Sperner's lemma.

Theorem 2.27 (see [25]). *Under the assumptions of Sperner's lemma, the number of completely labeled simplices whose orientations*[4] *coincide with that of the initial simplex is equal to the number of completely labeled simplices with opposite orientation plus* 1.

Proof. For the unit interval, the proof is virtually the same: for partitions with labels 01010...101, the number of positively oriented completely labeled intervals is equal to $k+1$, and that of negatively oriented completely labeled intervals is equal to k (here k is the number of zeros).

[4]We mean the orientations determined by the labeling.

Now, suppose that $n \geq 1$ and the required assertion is proved for the n-simplices. Consider a triangulation of an $(n+1)$-simplex with labeled vertices. Take an $(n+1)$-simplex in this triangulation. We mark each of its completely labeled n-faces Δ^n by $+$ or $-$ according to the following rule. The face Δ^n is endowed with two orientations, one of which is induced by the set of labels $0, 1, \ldots, n$ and the other is defined as the orientation of a face of an $(n+1)$-simplex induced by the orientation of this simplex (all $(n+1)$-simplices have the same orientations as the initial simplex). If these orientations of the face Δ^n coincide, then we mark it by $+$; otherwise, we mark it by $-$. Suppose that a face Δ^n belongs to two $(n+1)$-simplices Δ_1^{n+1} and Δ_2^{n+1}. These simplices induce opposite orientations on Δ^n, while the orientation determined by the marks $0, 1, \ldots, n$ is the same for both simplices Δ_1^{n+1} and Δ_2^{n+1}. Therefore, the sign of Δ^n as a face of Δ_1^{n+1} is opposite to that of Δ^n as a face of Δ_2^{n+1}. This immediately implies that the endpoints of one of the polygonal curves from the constructive proof of Sperner's lemma are the barycenters of

(1) either two completely labeled $(n+1)$-simplices with opposite orientations,

(2) or two completely labeled n-simplices with opposite orientations belonging to a completely labeled n-face of the initial simplex,

(3) or one completely labeled $(n+1)$-simplex and one completely labeled n-simplex with consistent orientations.

Thus, the difference between the numbers of positively and negatively oriented completely labeled $(n+1)$-simplices equals the difference between the numbers of positively and negatively oriented completely labeled n-simplices contained in a completely labeled n-face of the initial simplex. $\square$

3.4. The Kakutani Fixed Point Theorem. The Brouwer theorem can be generalized to maps of the simplex Δ^n which take points to subsets of this simplex rather than to points. These maps must have certain properties. First, we consider only those maps $x \mapsto \Phi(x) \subset \Delta^n$ for which the $\Phi(x)$ are closed convex sets. Second, the maps Φ must have a property similar to continuity; namely, they must be *upper semicontinuous*. This means that if $\lim_{i\to\infty} x_i = x_0$ and $y_i \in \Phi(x_i)$ are points such that $\lim_{i\to\infty} y_i = y_0$, then $y_0 \in \Phi(x_0)$. If Φ is a usual map, i.e., the set $\Phi(x)$ consists of a single point for each $x \in \Delta^n$, then its upper semicontinuity is equivalent to continuity.

Theorem 2.28 (Kakutani [63]). *Let Φ be an upper semicontinuous map which takes each point $x \in \Delta^n$ to a closed convex set $\Phi(x) \subset \Delta^n$. Then there exists a point $x_0 \in \Delta^n$ such that $x_0 \in \Phi(x_0)$.*

Proof. Consider the mth barycentric subdivision of the simplex Δ^n; to each of its vertices v_α we assign a point $w_\alpha \in \Phi(v_\alpha)$. Extending this map by linearity to the simplices of the mth barycentric subdivision, we obtain a continuous map $\varphi_m\colon \Delta^n \to \Delta^n$. According to the Brouwer theorem, there exists a point x_m for which $\varphi_m(x_m) = x_m$. Choose a convergent subsequence $\{x_{m_i}\}$ of the sequence $\{x_m\}$. We show that $x_0 = \lim_{i\to\infty} x_{m_i}$ has the required property, i.e., $x_0 \in \Phi(x_0)$.

For each m, take an n-simplex Δ^n_m of the mth barycentric subdivision that contains x_m (if there are several such simplices, we take any of them). Let $v_{0,m}, \dots, v_{n,m}$ be the vertices of Δ^n_m. Then $\lim_{i\to\infty} v_{j,m_i} = x_0$ and $x_m = \sum_{j=0}^n \lambda_{j,m} v_{j,m}$, where $\lambda_{j,m} \geq 0$ and $\sum_{j=0}^n \lambda_{j,m} = 1$. We set $w_{j,m} = \varphi_m(v_{j,m})$. By the definition of φ_m, we have $x_m = \varphi_m(x_m) = \sum_{j=0}^n \lambda_{j,m} w_{j,m}$ and, moreover, $w_{j,m} \in \Phi(v_{j,m})$. The expression $\sum_{j-0}^n \lambda_{j,m} w_{j,m}$ is determined by $n+1$ points in the compact space $I \times \Delta^n$. Therefore, the sequence $\{x_{m_i}\}$ has a subsequence $\{x_{m'_i}\}$ such that the sequences $\{\lambda_{j,m'_i}\}$ and $\{w_{j,m'_i}\}$ converge for all $j = 0, 1, \dots, n$. We set $\lim_{i\to\infty} \lambda_{j,m'_i} = \lambda_j$ and $\lim_{i\to\infty} w_{j,m'_i} = w_j$. Then $\sum_{j=0}^m \lambda_j w_j = x_0$. The upper semicontinuity of Φ implies $w_j \in \Phi(x_0)$, because $\lim_{i\to\infty} v_{j,m'_i} = x_0$, $w_{j,m'_i} \in \Phi(v_{j,m'_i})$, and $\lim_{i\to\infty} w_{j,m'_i} = w_j$. By assumption, the set $\Phi(x_0)$ is convex. Therefore, $x_0 = \sum_{j=0}^m \lambda_j w_j \in \Phi(x_0)$, as required. $\square$

The Kakutani theorem has many applications, but they belong mainly to geometry of convex bodies and mathematical economics, so we do not discuss them in this book.

Topological Spaces

1. Elements of General Topology

1.1. Hausdorff Spaces and Compact Spaces. A topological space X is said to be *Hausdorff* if any two different points $x, y \in X$ have disjoint neighborhoods $U \ni x$ and $V \ni y$. This separation property was introduced by F. Hausdorff in the book [**51**]. The simplest example of a non-Hausdorff space is a space X with *trivial* (*nondiscrete*) topology in which precisely two sets are open, X and $\varnothing$.

Exercise 14. Prove that a subspace of a Hausdorff space is Hausdorff.

Exercise 15. Prove that different points $x_1, \dots, x_n$ in a Hausdorff space X have disjoint neighborhoods $U_1, \dots, U_n$.

Non-Hausdorff spaces (in particular, spaces with trivial topology) often arise as spaces of orbits of group actions. Suppose that X is a set and G is a group. An *action* of the group G on the set X is a map $G \times X \to X$ (which assigns an element $g(x) \in X$ to each pair (g, x)) with the following properties:

(1) $g(h(x)) = (gh)(x)$;

(2) $e(x) = x$, where e is the identity element of G.

A *topological group* is a Hausdorff topological space G such that it is simultaneously a group and the maps $(g, h) \mapsto gh$ and $g \mapsto g^{-1}$ are continuous. Topology usually deals with continuous actions of topological groups on Hausdorff topological spaces.

For a point $x \in X$, the set $G(x) = \{g(x) \in X : g \in G\}$ is called the *orbit* of the point x under the action of the group G.

Exercise 16. Prove that the orbits $G(x)$ and $G(y)$ of any two points $x, y \in X$ are either disjoint or coincide.

Let X/G be the set consisting of all orbits under the action of G. Assigning the orbit $G(x)$ to each point $x \in X$, we obtain a map $p\colon X \to X/G$. We define a topology on the set X/G by declaring a set $U \subset X/G$ to be open if and only if $p^{-1}(U)$ is open. The resulting topological space X/G is called the *orbit space*.

Example 1. Suppose that $X = S^1 \times S^1$ (this is the 2-torus), $G = \mathbb{R}$, and a is a number. For $t \in G$ and $(e^{i\varphi}, e^{i\psi}) \in X$, we set $t(e^{i\varphi}, e^{i\psi}) = (e^{i(\varphi+t)}, e^{i(\psi+at)})$. If the number a is irrational, then the topology of the space X/G is trivial.

Proof. If a is irrational, then each orbit is everywhere dense, and therefore each orbit meets every nonempty open subset of the torus. □

Example 2. Let $X = \mathrm{Mat}_n(\mathbb{C})$ be the set of complex matrices of order n, and let $G = \mathrm{GL}_n(\mathbb{C}) \subset \mathrm{Mat}_n(\mathbb{C})$ be the group of nonsingular matrices. For $A \in X$ and $B \in G$, we set $B(A) = BAB^{-1}$. If $n \geq 2$, then the space X/G is non-Hausdorff.

Proof. We consider only the case $n = 2$. The matrices $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & \lambda \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} \lambda & 1 \\ 0 & \lambda \end{smallmatrix}\right)$ belong to different orbits O_1 and O_2, and all matrices $\left(\begin{smallmatrix} \lambda & s \\ 0 & \lambda \end{smallmatrix}\right)$ with $s \neq 0$ belong to the orbit O_2. Since

$$\lim_{s\to 0} \begin{pmatrix} \lambda & s \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix},$$

the orbit O_1 cannot be separated from O_2. □

Now, we prove some most important properties of Hausdorff spaces. First, note that for any distinct points x and y in a Hausdorff space X, there exists a neighborhood $U \ni x$ such that its closure does not contain y. Indeed, if $U \ni x$ and $V \ni y$ are disjoint neighborhoods, then $U \subset X \setminus V$. The set $X \setminus V$ is closed; therefore, $\overline{U} \subset X \setminus V$, and hence $\overline{U} \cap V = \varnothing$.

Theorem 3.1. *If C is a compact subset of a Hausdorff space X and $x \in X \setminus C$, then the point x and the set C have disjoint neighborhoods.*

Proof. Each point $c \in C$ has a neighborhood U whose closure does not contain x. Such neighborhoods form a cover of the compact space C, which has a finite subcover $U_1, \dots, U_n$. We set $V = \bigcup_{i=1}^n U_i$. Then $C \subset V$ and $x \notin \overline{V}$, i.e., V and $X \setminus \overline{V}$ are disjoint neighborhoods of the set C and the point x, respectively. □

Corollary 1. *Any compact subset C of a Hausdorff space X is closed.*

Proof. Each point $x \in X \setminus C$ has a neighborhood disjoint from C. This means that the set $X \setminus C$ is open. □

Corollary 2. *Any two disjoint compact subsets A and B of a Hausdorff space X have disjoint neighborhoods.*

Proof. Each point $a \in A$ has a neighborhood whose closure does not intersect B. Since the set A is compact, it has a finite cover $\{U_1, \dots, U_n\}$ consisting of such neighborhoods. The required neighborhoods of the sets A and B are $V = \bigcup_{i=1}^{n} U_i$ and $X \setminus \overline{V}$. □

Problem 26. Prove that any closed subset C of a compact space K is compact.

Theorem 3.2. *Let $f\colon X \to Y$ be a continuous one-to-one map from a compact space X to a Hausdorff space Y. Then f is a homeomorphism.*

Proof. The image Y of the compact space X under the continuous map f is compact because the preimage of any open cover of Y is an open cover of X (therefore, it has a finite subcover). Any closed subset C of the compact space X is compact; therefore, its image $f(C) \subset Y$ is compact as well. Since Y is Hausdorff, it follows that $f(C)$ is a closed subset of Y. But $f(C)$ is the preimage of C under f^{-1}; therefore, the map f^{-1} is continuous. □

Exercise 17. Construct a continuous one-to-one map of the half-open interval $[0,1)$ onto the circle S^1. (In Theorem 3.2, the compactness of X is essential.)

Exercise 18. Prove that a continuous map of the interval onto the square cannot be one-to-one.

In both of the above examples of non-Hausdorff orbit spaces X/G, the groups G are noncompact. Compact groups never cause such troubles.

Theorem 3.3. *If a compact group G acts on a (Hausdorff) space X, then the orbit space X/G is Hausdorff.*

Proof. Each orbit $G(x)$ is the image of the compact space G under the continuous map $G \to G \times \{x\} \to G(x)$; therefore, $G(x)$ is compact. If $G(x)$ and $G(y)$ are different orbits, then, according to Theorem 3.1, the point x has a neighborhood U whose closure is disjoint from $G(y)$. Thus, $p(U)$ and $(X/G) \setminus p(\overline{U})$ are disjoint open sets which contain the points of X/G representing the orbits $G(x)$ and $G(y)$. □

Problem 27* ([70]). (a) Prove that for any topological space X, there exists a Hausdorff space X^H and a continuous map $\sigma\colon X \to X^H$ with the following properties: if Y is a Hausdorff space and $f\colon X \to Y$ is a continuous map, then there exists a unique continuous map $f^H: X^H \to Y$ such that $f^H\sigma = f$.

(b) Let X/G be the orbit space from Example 2 on p. 88. Prove that $(X/G)^H$ is homeomorphic to $\mathbb{C}^n$.

(c) Prove that a continuous map $f\colon \mathrm{Mat}_n(\mathbb{C}) \to \mathbb{C}$ such that $f(BAB^{-1}) = A$ for all $B \in \mathrm{GL}_n(\mathbb{C})$ has a unique representation in the form $f(A) = F(c_1(A), \ldots, c_n(A))$, where $c_1(A), \ldots, c_n(A)$ are the coefficients of the polynomial $\det(A + \lambda I)$ and $F\colon \mathbb{C} \to \mathbb{C}^n$ is a continuous function.

A topological space X is said to be *locally compact* if every point $x \in X$ has an (open) neighborhood $U_x \ni x$ with compact closure.

Theorem 3.4. *Let X be a locally compact Hausdorff space. Then, for any open set $U \ni x$, there exists an open set $U_x \ni x$ such that $\overline{U}_x$ is compact and contained in U.*

Proof. Take an open neighborhood $W_x \ni x$ with compact closure and consider the compact set K that is the closure of $K' = \mathrm{int}(\overline{W}_x \cap U)$. The set $C = K \setminus K'$ is compact and does not contain x. Therefore, according to Theorem 3.1, x and C have disjoint neighborhoods U_x and U_C. The set U_x has the required properties. □

Theorem 3.5. *For any compact subspace K of a locally compact Hausdorff space X and any open set $U \subset X$ containing K, there exists an open subset V of X such that $K \subset V \subset \overline{V} \subset U$ and $\overline{V}$ is compact.*

Proof. For each point $x \in K$, take a neighborhood $U_x \ni x$ such that $\overline{U}_x \subset U$ and a neighborhood $W_x \ni x$ with compact closure $\overline{W}_x$. The closure of $V_x = U_x \cap W_x$ is compact because it is a closed subspace of the compact space $\overline{W}_x$. Using the compactness of K, we can choose a finite set of points $x_1, \ldots, x_n \in K$ such that $K \subset V = V_{x_1} \cup \cdots \cup V_{x_n}$. The set $\overline{V} = \overline{V}_{x_1} \cup \cdots \cup \overline{V}_{x_n}$ is compact, and $\overline{V} \subset \overline{U}_{x_1} \cup \cdots \cup \overline{U}_{x_n} \subset U$. □

Let X be a Hausdorff space. The *one-point compactification* of X is the topological space $X^+ = X \cup \{\infty\}$ in which the open sets are all open subsets of X and the sets $U \subset X^+$ such that their complements $X^+ \setminus U$ are compact subsets of X. (It is assumed that ∞ is a point that does not belong to X.) We must verify that the finite intersections and arbitrary unions of open subsets of X^+ are open. Clearly, the intersection with X of a finite intersection or an arbitrary union of open subsets of X^+ is open in X. Suppose that ∞ belongs to a finite intersection of sets open in X^+. Then the complement of the intersection of these sets is a finite union of compact sets; therefore, it is compact. Now, suppose that ∞ belongs to a union of a family of open subsets of X^+. Then ∞ belongs to some set U from this family. The complement of the union of such sets is a closed subset of the compact set $X \setminus U$; hence it is compact. Thus, X^+ is a topological space, and X is its subspace.

Take an arbitrary open cover $\mathcal{U}$ of X^+. We show that $\mathcal{U}$ has a finite subcover. The point ∞ belongs to one of the sets $U \in \mathcal{U}$. The set $X \setminus U$ is compact; therefore, $\mathcal{U}$ contains a finite subcover of this set.

Now, we show that if the space X is not only Hausdorff but also locally compact, then X^+ is Hausdorff. For this purpose, we must find disjoint open neighborhoods of a point $x \in X$ and the point ∞. The point x has an open neighborhood V_x with compact closure. The set $U = (X \setminus \overline{V}_x) \cup \{\infty\}$ is an open neighborhood of ∞ disjoint from V_x.

1.2. Normal Spaces. A topological space X is said to be *normal* if each point of X is a closed set and every two disjoint closed subsets A and B of X have disjoint open neighborhoods U and V.

Corollaries 1 and 2 of Theorem 3.1 show that any compact Hausdorff space is normal.

Exercise 19. Prove that any metrizable space is normal.

Urysohn's lemma, which we proved for the space $\mathbb{R}^n$ (see p. 56), remains valid for any normal space. Urysohn himself proved it for normal spaces.

Theorem 3.6 (Urysohn's lemma). *Let A and B be disjoint closed subsets of a normal space X. Then there exists a continuous function $f\colon X \to [0,1]$ such that $f(A) = 0$ and $f(B) = 1$.*

Proof. Let V be an open subset of a normal topological space X, and let U be a subset of X for which $\overline{U} \subset V$. Then there exists an open W such that $\overline{U} \subset W \subset \overline{W} \subset V$. Indeed, for W we can take an open set which contains the closed set $\overline{U}$ and does not intersect an open neighborhood of the closed set $X \setminus V$.

For $U = A$ and $V = X \setminus B$, we construct an open set A_1 such that

$$A \subset A_1 \subset X \setminus B \tag{1}$$

and $\overline{A}_1 \subset X \setminus B$. Then, we insert intermediate sets A_1' and A_2 so that

$$A \subset A_1' \subset A_1 \subset A_2 \subset X \setminus B \tag{2}$$

and the closure of each set is contained in the next set.

For the sequence of sets (1), we define the function $f_1\colon X \to [0,1]$ by

$$f_1(x) = \begin{cases} 0 & \text{if } x \in A, \\ 1/2 & \text{if } x \in A_1 \setminus A, \\ 1 & \text{if } x \in X \setminus A_1. \end{cases}$$

For the sequence of sets (2), we define the function $f_2\colon X \to [0,1]$ by

$$f_2(x) = \begin{cases} 0 & \text{if } x \in A, \\ 1/4 & \text{if } x \in A_1' \setminus A, \\ 1/2 & \text{if } x \in A_1 \setminus A_1', \\ 3/4 & \text{if } x \in A_2 \setminus A_1, \\ 1 & \text{if } x \in X \setminus A_2. \end{cases}$$

Then, we construct a third sequence of sets by inserting intermediate open sets between the neighboring terms of sequence (2) and define the function $f_3(x)$ for this sequence, and so on.

It is easy to show that $f_2(x) \geq f_1(x)$, and similarly, $f_{n+1}(x) \geq f_n(x)$. Hence the limit $\lim_{n\to\infty} f_n(x) = f(x)$ exists. Clearly, $f(x) = 0$ for $x \in A$ and $f(x) = 1$ for $x \in B$. We must only prove that the function $f(x)$ is continuous.

Suppose that at the nth step, a sequence of sets

$$A \subset A_1 \subset \cdots \subset A_r \subset X \setminus B,$$

where $\overline{A}_i \subset A_{i+1}$, is constructed. (This sequence corresponds to the function f_n.) We put $A_0 = \operatorname{int} A$ ($\operatorname{int} A$ denotes the interior of A), $A_{-1} = \varnothing$, and $A_{r+1} - X$. Consider the open sets $A_{i+1} \setminus \overline{A}_{i-1}$ for $i = 0, 1, \ldots, r$. Clearly,

$$X = \bigcup_{i=0}^{r} (\overline{A}_i \setminus \overline{A}_{i-1}) \subset \bigcup_{i=0}^{r} (A_{i+1} \setminus \overline{A}_{i-1})\,;$$

therefore, the $A_{i+1} \setminus \overline{A}_{i-1}$ cover the space X.

On each set $A_{i+1} \setminus \overline{A}_{i-1}$, the function $f_n(x)$ takes two values, which differ by $1/2^n$. Clearly, $|f(x) - f_n(x)| \leq \sum_{k=n+1}^{\infty} 1/2^k = 1/2^n$.

For each point $x \in X$, choose an open neighborhood of it in the form $A_{i+1} \setminus \overline{A}_{i-1}$. The image of the open set $A_{i+1} \setminus \overline{A}_{i-1}$ is contained in the interval $(f(x) - \varepsilon, f(x) + \varepsilon)$, where $\varepsilon > 1/2^n$. Letting n tend to infinity, we see that the function f is continuous. □

The Tietze extension theorem can be derived from Urysohn's lemma.

Theorem 3.7 (Tietze). *Suppose that Y is a normal topological space, $X \subset Y$ is a closed subset of Y, and $f\colon X \to [-1,1]$ is a continuous function. Then there exists a continuous function $F\colon Y \to [-1,1]$ such that its restriction to X coincides with f.*

The proof is the same as in the case of $Y = \mathbb{R}^n$ (see Theorem 2.4 on p. 57). The only difference is that $\mathbb{R}^n$ is replaced by Y and, instead of Urysohn's lemma for $\mathbb{R}^n$, Urysohn's lemma for normal topological spaces is applied. The corollary of Theorem 2.4 remains valid as well.

1.3. Partitions of Unity. Let φ be a continuous function on a topological space X. The *support* of φ is the closed set

$$\operatorname{supp}(\varphi) = \overline{\{x \in X : \varphi(x) \neq 0\}}.$$

Let $\{U_\alpha\}$ be an open cover of a topological space X. A *partition of unity subordinate to* $\{U_\alpha\}$ is defined as a family of continuous functions $\varphi_\alpha \colon X \to [0,1]$ with the following properties:

(1) the family φ_α is *locally finite*, i.e., each point $x \in X$ has a neighborhood $V(x)$ intersecting only finitely many supports $\operatorname{supp}(\varphi_\alpha)$;

(2) $\sum \varphi_\alpha(x) = 1$ for any $x \in X$;

(3) $\operatorname{supp}(\varphi_\alpha) \subset U_\alpha$ for all α.

Sometimes, families $\{U_\alpha\}$ and $\{\varphi_\beta\}$ with different indices are considered. In this case, it is assumed that for every β, there exists an α such that $\operatorname{supp}(\varphi_\beta) \subset U_\alpha$.

Theorem 3.8 (Stone [**123**]). *Let X be a metrizable topological space. Then, for any at most countable open cover $\{U_i\}$ of X, there exists a partition of unity $\{\varphi_i\}$ subordinate to this cover.*

Proof [**37,82**]. First, consider the case of a finite cover $U_1, \ldots, U_n$. The functions $f_i(x) = d(x, X \setminus U_i)$ are continuous (see the remark on p. 55); therefore, the function $F(x) = \sum_{i=1}^n f_i(x)$ is continuous as well. Each point $x \in X$ is covered by some set U_i. We have $f_i(x) > 0$, and hence, $F(x) > 0$ for all $x \in X$. Let

$$g_i(x) = \max\left\{ f_i(x) - \frac{1}{n+1} F(x), 0 \right\}.$$

Then

$$\begin{aligned} \operatorname{supp}(g_i) &= \overline{\{x : f_i(x) > F(x)/(n+1)\}} \subset \{x : f_i(x) \geq F(x)/(n+1)\} \\ &\subset \{x : f_i(x) \geq 0\} \subset \{x : f_i(x) > 0\} = U_i. \end{aligned}$$

Moreover,

$$\sum_{i=1}^n g_i(x) \geq \sum_{i=1}^n \left(f_i(x) - \frac{1}{n+1} F(x) \right) = F(x) - \frac{n}{n+1} F(x) = \frac{F(x)}{n+1} > 0.$$

The required partition of unity consists of the functions

$$\varphi_i(x) = g_i(x) / \Big(\sum_{i=1}^n g_i(x) \Big).$$

Now, consider the case of a countable open cover $U_0, U_1, U_2, \ldots$. This time, we define the functions $f_i \colon X \to [0, 2^{-i}]$ by

$$f_i(x) \doteq \min\{ d(x, X \setminus U_i), 2^{-i} \}.$$

We have $f_i(x) > 0$ for $x \in U_i$ and $f_i(x) = 0$ for $x \notin U_i$. The function F is defined differently too: we set $F(x) = \sum_{i=0}^{\infty} 2^{-i} f_i(x)$. Since $\{U_i\}$ is a cover, it follows that $F(x) > 0$ for all $x \in X$. The function $F(x)$ is continuous because $\sum_{i=0}^{N} 2^{-i} f_i(x)$ is continuous and, since $f_i(x) \le 2^{-i}$, for any $\varepsilon > 0$, we can choose N such that $\sum_{i=N+1}^{\infty} 2^{-i} f_i(x) < \varepsilon$.

We set $g_i(x) = \max\{f_i(x) - \frac{1}{3}F(x), 0\}$. As in the case of finite covers, we have $\operatorname{supp}(g_i) \subset U_i$.

We prove that the family of functions $\{g_i\}$ is locally finite. Take $x \in X$. The continuity of F implies the existence of a neighborhood $V(x)$ of x such that for some $\varepsilon > 0$, $F(y) > \varepsilon$ at all $y \in V(x)$. Choose i_0 for which $2^{-i_0} < \varepsilon/3$. We have $f_i(y) \le 2^{-i}$ for any point $y \in X$. Therefore, if $y \in V(x)$ and $i \ge i_0$, then

$$f_i(x) - \frac{1}{3}F(x) \le 2^{-i} - \frac{\varepsilon}{3} \le 2^{-i_0} - \frac{\varepsilon}{3} < 0;$$

hence $g_i(y) = 0$.

Finally, we prove that $\sum_{i=0}^{\infty} g_i(x) > 0$ for all $x \in X$, i.e., for any point $x \in X$, there exists a number i such that $g_i(x) > 0$. Since $f_j(x) > 0$ for some j and $f_n(x) \le 2^{-n}$, it follows that $\sup_{j \in \mathbb{N}} f_j(x) = f_{i_0}(x)$ for some i_0, and $f_{i_0}(x) > 0$. The definition of F implies

$$F(x) = \sum_{i=0}^{\infty} 2^{-i} f_i(x) \le \sum_{i=0}^{\infty} 2^{-i} f_{i_0}(x) = 2 f_{i_0}(x).$$

Therefore,

$$g_{i_0}(x) \ge f_{i_0}(x) - \frac{2 f_{i_0}(x)}{3} = \frac{f_{i_0}(x)}{3} > 0.$$

The required partition of unity consists of the functions

$$\varphi_i(x) = g_i(x) / \Bigl(\sum_{i=1}^{\infty} g_i(x)\Bigr). \qquad \square$$

1.4. Paracompact Spaces. Suppose that $\mathcal{U} = \{U_\alpha\}$ and $\mathcal{V} = \{V_\beta\}$ are open covers of a topological space X. We say that the cover $\mathcal{V}$ is a *refinement* of $\mathcal{U}$ if each set V_β is contained in some U_α.

A cover $\mathcal{V} = \{V_\beta\}$ is said to be *locally finite* if every point $x \in X$ has a neighborhood intersecting only finitely many sets V_β.

A topological space X is called *paracompact* if it is Hausdorff and any open cover $\mathcal{U}$ of X has an open locally finite refinement $\mathcal{V}$.

The most important property of paracompact spaces is that *for any open cover of a paracompact space, there exists a locally finite partition of unity subordinate to it.* This property of paracompact spaces follows from Theorems 3.9 and 3.10 below (which are of independent interest). Before proving them, we give an example which shows that the class of paracompact spaces is very large.

Example. Any set $X \subset \mathbb{R}^n$ (with induced topology) is paracompact.

Proof. Consider any open cover $\{U_\alpha\}$ of the topological space X. For each set U_α, there exists an open set $U'_\alpha \subset \mathbb{R}^n$ such that $U_\alpha = U'_\alpha \cap X$.

Let $X_k = \{x \in X : \|x\| < k\}$ for $k = 1, 2 \dots$. The sets X_k are open in X, each $\overline{X}_k$ is compact, $\overline{X}_k \subset X_{k+1}$ for all k, and $X = \bigcup_{k=1}^{\infty} X_k$.

Consider the compact set $\overline{X}_k \setminus X_{k-1}$. For every point $z \in \overline{X}_k \setminus X_{k-1}$, we choose its open neighborhood V'_z in $\mathbb{R}^n$ such that $V'_z \subset U'_\alpha$ for some α, $V'_z \cap X = V_z \subset X_{k+1}$, and $V_z \cap X_{k-2} = \varnothing$. The open sets V'_z cover the compact set $\overline{X}_k \setminus X_{k-1}$; hence there exists a finite family of sets V_z which covers $\overline{X}_k \setminus X_{k-1}$. The union of all such families over all k is a locally finite refinement of the cover $\{U_\alpha\}$. □

Theorem 3.9 (Dieudonné [32]). *Any paracompact space is normal.*

Proof. First, we prove that any paracompact space X is *regular*, i.e., every open neighborhood of a point $x \in X$ contains the closure of some other open neighborhood of x. Let U be an open neighborhood of a point $x \in X$. Since X is Hausdorff, for every point $y \in X \setminus U$, there exist disjoint open sets $U_y \ni y$ and $W_y \ni x$. The sets U_y (for all $y \in X \setminus U$), together with U, form an open cover $\mathcal{U}$ of X; it has a locally finite open refinement $\mathcal{V}$. The local finiteness of $\mathcal{V}$ means that the point x has a neighborhood W intersecting only finitely many elements of the cover $\mathcal{V}$. Let $V_1, \dots, V_n$ be those of them which are not contained in U. Each V_i is contained in U_{y_i} for some $y_i \in X \setminus U$. We put

$$Z = W \cap W_{y_1} \cap \dots \cap W_{y_n} \quad \text{and} \quad C = \overline{Z}.$$

The set Z is open and contains x because $x \in W$ and $x \in W_y$ for any point $y \in X \setminus U$. It remains to show that $C \subset U$. Consider the union T of all elements of the cover $\mathcal{V}$ not contained in U. By construction,

$$T \cap W \subset V_1 \cup \dots \cup V_n \subset U_{y_1} \cup \dots \cup U_{y_n}.$$

Clearly, $Z \subset W$; therefore,

$$Z \cap T \subset W \cap W_{y_1} \cap \dots \cap W_{y_n} \cap (U_{y_1} \cup \dots \cup U_{y_n}).$$

By construction, $W_y \cap U_y = \varnothing$, whence

$$W \cap W_{y_1} \cap \dots \cap W_{y_n} \cap (U_{y_1} \cup \dots \cup U_{y_n}) = \varnothing;$$

this means that $Z \cap T = \varnothing$, i.e., $Z \subset X \setminus T$. Since $X \setminus T$ is closed, we have $C = \overline{Z} \subset X \setminus T$.

Every point of $X \setminus U$ belongs to some element of $\mathcal{V}$ not contained in U. Hence $X \setminus U \subset T$, i.e., $X \setminus T \subset U$ (these inclusions are equivalent to $X = T \cup U$).

Now, we prove that X is normal. Suppose that A and B are disjoint closed subsets of X. Each point $a \in A$ belongs to the open set $X \setminus B$; therefore, a has an open neighborhood Z_a for which $C_a = \overline{Z}_a \subset X \setminus B$. The sets Z_a (for all $a \in A$), together with $X \setminus A$, form an open cover $\mathcal{U}$ of X; it has a locally finite open refinement $\mathcal{V}$. Let U be the union of all elements of $\mathcal{V}$ not contained in $X \setminus A$. Then U is an open set containing A. It remains to construct an open set V containing B and disjoint from U. We construct it as the union of certain sets V_b over all $b \in B$. Namely, for each point $b \in B$, we choose an open neighborhood $W_b \ni b$ intersecting only finitely many elements of $\mathcal{V}$. Let $Y_1, \dots, Y_n$ be those of them which are not contained in $X \setminus A$. By construction, we have $Y_i \subset Z_{a_i}$, where $a_i \in A$. We put

$$V_b = W_b \cap (X \setminus C_{a_1}) \cap \dots \cap (X \setminus C_{a_n}).$$

The set V_b is open, and $b \in V_b$, because $B \subset X \setminus C_a$. Moreover,

$$U \cap W_b \subset Y_1 \cap \dots \cap Y_n \subset Z_{a_1} \cap \dots \cap Z_{a_n} \subset C_{a_1} \cap \dots \cap C_{a_n};$$

therefore, $V_b \cap U = \varnothing$. Thus, if $V = \bigcup_{b \in B} V_b$, then $B \subset V$ and $V \cap U = \varnothing$. □

Remark 3.1. The reader might have noticed that we repeated twice very similar arguments. Instead, we could prove one general assertion and apply it twice in different situations. Such a proof of Theorem 3.9 was given in [**20**, Chapter IX, Section 4.4, Proposition 4].

Theorem 3.10. *For any open locally finite cover $\mathcal{U} = \{U_\alpha : \alpha \in \mathcal{A}\}$ of a normal space X, there exists a partition of unity subordinate to this cover.*

Proof. First, we construct an open cover $\mathcal{V} = \{V_\alpha : \alpha \in \mathcal{A}\}$ of X such that $\overline{V}_\alpha \subset U_\alpha$ for all $\alpha \in \mathcal{A}$. The construction uses *transfinite induction*; let us briefly recall what this means (the details, including the proof of Zermelo's theorem, can be found in [**134**]).

A set $\mathcal{A}$ is said to be *well ordered* if it is ordered and any nonempty subset of $\mathcal{A}$ has the least element, which precedes all the other elements of $\mathcal{A}$. According to *Zermelo's well-ordering theorem*, any set $\mathcal{A}$ can be well ordered. Suppose that $\mathcal{A}$ is a well-ordered set and P is a property such that if all elements preceding some $\alpha \in \mathcal{A}$ have this property P, then the element α has property P (in particular, the least element of $\mathcal{A}$ must have property P). Then all elements of the set $\mathcal{A}$ have property P. Indeed, if the set of elements of $\mathcal{A}$ not having property P is nonempty, then it has a least element α_0. All elements preceding α_0 have property P; therefore, the element α_0 also has property P, which contradicts its definition.

Suppose that for each $\alpha < \alpha_0$, there exist open sets V_α such that $\overline{V}_\alpha \subset U_\alpha$ and, for every $\alpha_1 < \alpha_0$, the family $\{V_\alpha : \alpha \leq \alpha_1\} \cup \{U_\alpha : \alpha > \alpha_1\}$ is a cover of X. It is required to construct an open set V_{α_0} such that $\overline{V}_{\alpha_0} \subset U_{\alpha_0}$ and $\{V_\alpha : \alpha \leq \alpha_0\} \cup \{U_\alpha : \alpha > \alpha_0\}$ is a cover of X.

First, we show that $\{V_\alpha : \alpha < \alpha_0\} \cup \{U_\alpha : \alpha \geq \alpha_0\}$ is a cover of X. To this end, we apply the local finiteness of the cover $\mathcal{U}$. Each point $x \in X$ belongs to only finitely many sets $U_{\beta_1}, \dots, U_{\beta_n}$; therefore, we can choose the largest element among $\beta_1, \dots, \beta_n$. For definiteness, we assume that $\beta_n > \beta_i$ for $i = 1, \dots, n-1$. If $\beta_n \geq \alpha_0$, then $x \in U_\alpha$, where $\alpha = \beta_n \geq \alpha_0$. If $\beta_n < \alpha_0$, then, by assumption, the family $\{V_\alpha : \alpha \leq \beta_n\} \cup \{U_\alpha : \alpha > \beta_n\}$ is a cover of X. But $x \notin U_\alpha$ for $\alpha > \beta_n$; therefore, $x \in \bigcup_{\alpha \leq \beta_n} V_\alpha \subset \bigcup_{\alpha < \alpha_0} V_\alpha$.

Thus, the open set $W = \left(\bigcup_{\alpha<\alpha_0} V_\alpha\right) \cup \left(\bigcup_{\alpha>\alpha_0} U_\alpha\right)$, together with U_{α_0}, covers the whole space X; therefore, $X \setminus U_{\alpha_0} \subset W$. The closed sets $X \setminus U_{\alpha_0}$ and $X \setminus W$ do not intersect; since X is normal, they have disjoint open neighborhoods $Z \supset X \setminus U_{\alpha_0}$ and $T \supset X \setminus W$. Clearly,

$$X \setminus U_{\alpha_0} \subset Z \subset \overline{Z} \subset X \setminus T \subset W.$$

Let $V_{\alpha_0} = X \setminus \overline{Z}$. The set V_{α_0} has the required properties. Indeed, $\overline{V}_{\alpha_0} = X \setminus Z \subset U_{\alpha_0}$ and $V_{\alpha_0} \cup W = X$, because $\overline{Z} \subset W$.

Now we construct the partition of unity subordinate to the cover $\mathcal{U}$. Instead of $\mathcal{U}$, we take the cover $\mathcal{V}$ constructed above, for which $\overline{V}_\alpha \subset U_\alpha$; clearly, it is locally finite. The normality of X implies the existence of an open set W_α such that $\overline{V}_\alpha \subset W_\alpha \subset \overline{W}_\alpha \subset U_\alpha$. By the Tietze extension theorem, there exists a continuous map $g_\alpha \colon X \to [0,1]$ for which $g_\alpha(X \setminus W_\alpha) = 0$ (i.e., $\operatorname{supp} g_\alpha \subset \overline{W}_\alpha \subset U_\alpha$) and $g_\alpha(\overline{V}_\alpha) = 1$. The sets $V_\alpha \subset \overline{V}_\alpha$ cover X; hence $\sum_{\alpha\in\mathcal{A}} g_\alpha(x) > 0$ for each $x \in X$. Since the cover $\mathcal{V}$ is locally finite, it follows that the function $\sum_{\alpha\in\mathcal{A}} g_\alpha(x)$ is continuous. To obtain the required partition of unity, we set $\varphi_\alpha(x) = g_\alpha(x)\left(\sum_{\alpha\in\mathcal{A}} g_\alpha(x)\right)$. □

Corollary of Theorems 3.9 and 3.10. *For any open cover of a paracompact space X, there exists a locally finite partition of unity subordinate to it.*

Proof. Let $\mathcal{U} = \{U_\alpha : \alpha \in \mathcal{A}\}$ be an open cover of a paracompact space X, and let $\mathcal{V} = \{V_\beta : \beta \in \mathcal{B}\}$ be its locally finite refinement. Then there exists a map $A \colon \mathcal{B} \to \mathcal{A}$ such that $V_\beta \subset U_{A(\beta)}$. According to Theorem 3.9, the space X is normal; therefore, Theorem 3.10 implies the existence of a partition of unity $\{\varphi_\alpha\}$ subordinate to $\mathcal{V}$. For each $\alpha \in \mathcal{A}$, we set $\psi_\alpha = \sum_{A(\beta)=\alpha} \varphi_\beta$. This sum is well defined and continuous because $\operatorname{supp} \varphi_\beta \subset V_\beta$ and the cover $\mathcal{V}$ is locally finite. Let $C_\alpha = \bigcup_{A(\beta)=\alpha} \operatorname{supp} \varphi_\beta$. The set C_α is closed, being the union of a locally finite family of closed sets. Clearly, $C_\alpha \subset U_\alpha$ and $\psi_\alpha(x) = 0$ for $x \notin C_\alpha$. Thus, $\operatorname{supp} \psi_\alpha \subset C_\alpha \subset U_\alpha$.

It is easy to verify that the family of sets $\{C_\alpha\}$ is locally finite. Indeed, for any point $x \in X$, there exists a neighborhood W intersecting only finitely many elements of the cover $\mathcal{V}$; we denote them by $V_{\beta_1}, \ldots, V_{\beta_k}$. The neighborhood W does not intersect C_α if $\alpha \notin \{A(\beta_1), \ldots, A(\beta_k)\}$. Thus, the families of sets $\{\operatorname{supp}\varphi_\beta\}$ and $\{\operatorname{supp}\psi_\alpha\}$ are locally finite; therefore,

$$\sum_{\alpha \in \mathcal{A}} \psi_\alpha(x) = \sum_{\alpha \in \mathcal{A}} \Bigl(\sum_{A(\beta)=\alpha} \varphi_\beta(x) \Bigr) = \sum_{\beta \in \mathcal{B}} \varphi_\beta(x) = 1. \qquad \square$$

We have proved (see p. 93) that for an arbitrary at most countable cover of a metrizable space, there exists a partition of unity subordinate to it. Below we prove a somewhat stronger assertion.

Theorem 3.11 (Stone [123]). *Any metrizable space is paracompact.*

Proof (see [114]). Let $\mathcal{U} = \{U_\alpha : \alpha \in \mathcal{A}\}$ be an open cover of a metric space X with metric d. We again use the fact that the set $\mathcal{A}$ can be well ordered. For $x \in X$ and $r > 0$, consider the open disk $D_{x,r} = \{y \in X : d(x,y) < r\}$. For $\alpha \in \mathcal{A}$ and $n \in \mathbb{N}$, we define $V_{\alpha,n}$ as the union of the sets $D_{x,2^{-n}}$ over all points $x \in X$ satisfying the following three conditions:

(1) $D_{x,3\cdot 2^{-n}} \subset U_\alpha$;

(2) $x \notin U_\beta$ for $\beta < \alpha$;

(3) $x \notin V_{\beta,j}$ for $j < n$.

We first define the sets $V_{\alpha,n}$ for $n = 1$ (in this case, only the first two conditions are taken into consideration), then for $n = 2$, and so on.

We start by proving that the sets $V_{\alpha,n}$ cover X. For every point $x \in X$, consider the set $\mathcal{B} = \{\beta \in \mathcal{A} : x \in U_\beta\}$. Let α be the least element of $\mathcal{B}$. Choose n such that $D_{x,3\cdot 2^{-n}} \subset U_\alpha$. If $x \notin V_{\beta,j}$ for $j < n$, then x satisfies conditions (1)–(3), and therefore $x \in V_{\alpha,n}$. Thus, the point x belongs to some set $V_{\beta,j}$ with $j \le n$.

It remains to prove that the cover $\{V_{\alpha,n}\}$ is locally finite. For a point $x \in X$, consider the set

$$\mathcal{B} = \{\beta \in \mathcal{A} : x \in V_{\beta,n} \text{ for some } n\}.$$

Suppose that α is the least element of $\mathcal{B}$ and let $x \in V_{\alpha,n}$. Choose $j \in \mathbb{N}$ such that $D_{x,2^{-j}} \subset V_{\alpha,n}$. We show that the open set $D_{x,2^{-j-n}}$ intersects only finitely many sets $V_{\beta,i}$. For this purpose, it suffices to prove that this set does not intersect $V_{\beta,i}$ for $i \ge n+j$ and intersects at most one $V_{\beta,i}$ for $i < n+j$.

First, suppose that $i \ge n+j > n$. The set $V_{\beta,i}$ consists of open disks of radius 2^{-i} centered at points satisfying conditions (1)–(3). In particular, it follows from (3) that if y is the center of such a disk, then $y \notin V_{\alpha,n}$. But

$D_{x,2^{-j}} \subset V_{\alpha,n}$; therefore, $d(x,y) \geq 2^{-j}$. On the other hand, $n+j \geq j+1$ and $i \geq j+1$, whence $2^{-j-n}+2^{-i} \leq 2^{-j}$, which means that $D_{x,2^{-j-n}} \cap D_{y,2^{-i}} = \varnothing$.

Now, suppose that $i < n+j$, $p \in D_{x,2^{-j-n}} \cap V_{\beta,i}$, and $q \in D_{x,2^{-j-n}} \cap V_{\gamma,i}$, where $\beta \neq \gamma$. For definiteness, we assume that $\beta < \gamma$. To obtain a contradiction, it is sufficient to prove that the relations $p \in V_{\beta,i}$ and $q \in V_{\gamma,i}$ imply $d(p,q) \geq 2^{-j-n+1}$. Let y and z be the centers of disks $D_{y,2^{-i}}$ and $D_{z,2^{-i}}$ such that $p \in D_{y,2^{-i}} \subset V_{\beta,i}$ and $q \in D_{z,2^{-i}} \subset V_{\gamma,i}$. Condition (1) implies $D_{y,3\cdot 2^{-i}} \not\subset U_\beta$, and (2) implies $z \notin U_\beta$. Therefore, $d(y,z) \geq 3\cdot 2^{-i}$, and hence

$$d(p,q) \geq d(y,z) - d(p,y) - d(q,z) \geq 3\cdot 2^{-i} - 2^{-i} - 2^{-i} = 2^{-i} \geq 2^{-n-j+1}. \quad \square$$

2. Simplicial Complexes

The Euclidean space $\mathbb{R}^n$ is the most important example of a topological space. All basic classes of topological spaces (simplicial complexes, CW-complexes, and manifolds) are constructed by gluing together Euclidean simplices or disks. For purely technical reasons, in homotopic topology, it is more convenient to deal with CW-complexes than with simplicial complexes. The point is that simplicial complexes carry a great deal of geometric information, which is obviously excessive for needs of topology. Nevertheless, simplicial complexes form a fairly interesting and extensive class of topological spaces, and they are most convenient to use (at least, they are most often used) in geometric topology.

A *simplicial complex* K is a set of simplices in $\mathbb{R}^n$ satisfying the following conditions:

- all faces of simplices from K belong to K;
- the intersection of any two simplices from K is a face for each of them (for convenience, we assume that the empty set is a face of dimension -1 for any simplex);
- any point that belongs to one of the simplices from K has a neighborhood which intersects only finitely many simplices from K.

The *dimension* of a complex K is defined as the maximum dimension of its simplices.

A simplicial complex K is said to be *finite* if it consists of a finite number of simplices. In what follows, we mainly consider finite simplicial complexes,

To every simplicial complex K corresponds the topological space $|K|$, which is the union of all simplices of K with topology induced by $\mathbb{R}^n$.

On p. 81, the barycentric subdivision of a simplex was defined. Taking the barycentric subdivisions of all simplices of K, we obtain the *barycentric subdivision* of the simplicial complex K.

Problem 28. Prove that the simplices in the barycentric subdivision of the simplex Δ^n are in one-to-one correspondence with the ordered sets of vertices of Δ^n.

2.1. Rectilinear Cell Complexes. A *convex polyhedron* of dimension k is defined as a subset of $\mathbb{R}^k$ that is determined by a system of linear inequalities of the form $Ax \leq b$, contains a k-disk, and is contained in a k-disk.

A k-dimensional convex polyhedron contained in a k-dimensional (affine) subspace of $\mathbb{R}^n$, where $n \geq k$, is called a *rectilinear k-cell*.

A *rectilinear cell complex* K is a set of rectilinear cells in $\mathbb{R}^n$ satisfying the following conditions:

- all faces of rectilinear cells from K belong to K;
- the intersection of any two rectilinear cells from K is a face for each of them;
- every point of the set $|K|$ has a neighborhood intersecting only finitely many rectilinear cells from K (here $|K|$ again denotes the union of all cells of K).

Any simplicial complex is a rectilinear cell complex.

A rectilinear cell complex K' is called a *subdivision* of a rectilinear cell complex K if $|K| = |K'|$ and any cell of K' is contained in some cell of K.

The union of all cells of dimension at most l in a rectilinear cell complex K is called the *l-skeleton* of K; we denote it by K^l. If the dimension of K is not less than l, then the l-skeleton of K is an l-dimensional rectilinear cell complex.

Theorem 3.12. *Two rectilinear cell complexes K_1 and K_2 such that $|K_1| = |K_2|$ have a common subdivision L.*

Proof. Clearly, the intersection of two rectilinear cells is again a rectilinear cell. Let L be the set of all cells of the form $c_1 \cap c_2$, where c_1 is a cell from K_1 and c_2 is a cell from K_2. Then L is a rectilinear cell complex, $|L| = |K_1| = |K_2|$, and any cell $c_1 \cap c_2$ from L is contained in the cell c_1 of K_1 and in the cell c_2 of K_2. □

The following assertion shows that, topologically, rectilinear cell complexes give nothing new in comparison with simplicial complexes.

Theorem 3.13. *Any rectilinear cell complex K has a subdivision which is a simplicial complex.*

Proof. We prove this theorem by induction on $n = \dim K$. The rectilinear cells of dimension ≤ 1 are simplices; therefore, for $n \leq 1$, the assertion of the theorem is obvious. Suppose that we have already constructed a simplicial complex L which is a subdivision of the $(m-1)$-skeleton of K. Inside each m-cell c^m of K, we choose a point M. Consider the simplices such that for each of them, one of the vertices is M and the remaining vertices are the vertices of one of the simplices constituting the boundary of c^m. As a result, we obtain a simplicial subdivision of K. □

Remark 3.2. For M we can take a vertex of the cell c^m rather than its interior point. Then the constructed simplicial subdivision has the same vertices as the initial rectilinear cell complex.

2.2. Simplicial Maps. Let K_1 and K_2 be simplicial complexes. A map $f\colon |K_1| \to |K_2|$ is said to be *simplicial* if the image of any simplex Δ_1 from K_1 is a simplex Δ_2 from K_2 and the restriction of f to Δ_1 is linear, i.e.,

(1) $$f\left(\sum \lambda_i v_i\right) = \sum \lambda_i f(v_i),$$

where the v_i are the vertices of Δ_1, $\sum \lambda_i = 1$, and $\lambda_i \geq 0$. By definition, the vertices (i.e., 0-simplices) of the complex K_1 are mapped to vertices of K_2. Therefore, the map f determines a map $f^0\colon K_1^0 \to K_2^0$ of 0-skeletons. On the other hand, according to (1), f is uniquely determined by f^0. The map f^0 has the following property: if $v_0, \dots, v_n$ are the vertices of a simplex from K_1, then $f^0(v_0), \dots, f^0(v_n)$ are the vertices of a simplex from K_2 (some of the points $f^0(v_0), \dots, f^0(v_n)$ may coincide). We refer to maps of 0-skeletons with this property as *admissible*. Each admissible map $K_1^0 \to K_2^0$ of 0-skeletons determines a simplicial map $|K_1| \to |K_2|$. We usually denote simplicial maps by $K_1 \to K_2$.

Exercise 20. Prove that any simplicial map is continuous.

Exercise 21. Prove that the image of the k-skeleton of K_1 under a simplicial map is contained in the k-skeleton of K_1.

Theorem 3.14. *Suppose that $f\colon K \to K$ is a simplicial map, Δ' is a simplex from the barycentric subdivision of the complex K, and $f(\Delta') = \Delta'$. Then the restriction of f to Δ' is the identity map.*

Proof. For the simplex Δ', there is a unique simplex Δ in K which contains Δ' and has the same dimension. Moreover, Δ' uniquely determines the ordering of the vertices of Δ under which v_0 is the common vertex of Δ and Δ', $[v_0, v_1]$ is their common edge (or, to be more precise, the edge of Δ containing an edge of Δ'), $[v_0, v_1, v_2]$ is their common 2-face, etc. Conversely, any ordering of the vertices of Δ uniquely determines a simplex in the barycentric subdivision.

Since $f(\Delta') = \Delta'$, we also have $f(\Delta) = \Delta$; the map f can only permute the vertices of Δ. If this permutation were not the identity, then it would change the ordering of the vertices of Δ, and the new ordering would correspond to a different simplex of the barycentric subdivision, i.e., we would have $f(\Delta') \neq \Delta'$. Therefore, the restriction of f to $\Delta \supset \Delta'$ is the identity map. $\square$

2.3. Abstract Simplicial Complexes. From the point of view of topology, of interest are the topological spaces $|K|$ determined by simplicial complexes K rather than the complexes themselves. Each simplicial complex K determines not only the space $|K|$ but also its embedding in $\mathbb{R}^n$; this is a superfluous information, which often complicates dealing with simplicial complexes. To get rid of a particular embedding in $\mathbb{R}^n$, we define an *abstract simplicial complex* K as a vertex set $\{v_\alpha\}$ and a family of its subsets, called simplices (a set of $k+1$ vertices is a k-simplex); any subset of a simplex from K must be a simplex from K.

To each abstract simplicial complex K we assign a topological space $|K|$ as follows. To every abstract k-simplex $\{v_{i_1}, \dots, v_{i_{k+1}}\}$ corresponds the geometric simplex $[v_{i_1}, \dots, v_{i_{k+1}}]$, which we treat as a topological space. In the disjoint union of these topological spaces, we identify the respective points of the simplex $[v_1, \dots, v_p]$ and of the face $[v_1, \dots, v_p]$ in the simplex $[v_1, \dots, v_p, v_{p+1}, \dots, v_q]$. The topology on the coset space $|K|$ thus obtained is defined as follows: a set U is open in $|K|$ if and only if its intersection with each simplex is open in the topology of this simplex.

Suppose that K is an abstract simplicial complex and $\sigma\colon K^0 \to L^0$, where L^0 is the 0-skeleton of a simplicial complex L in $\mathbb{R}^n$, is a one-to-one map with the following property: a set of vertices $v_1, \dots, v_k$ is a simplex in K if and only if L contains a simplex with vertices $\sigma(v_1), \dots, \sigma(v_k)$. Such a map σ admits a natural extension to a homeomorphism $|K| \to |L|$. The simplicial complex L together with this homeomorphism is called a *realization* of the simplicial complex K.

Theorem 3.15. *Any finite n-dimensional abstract simplicial complex has a realization in the Euclidean $2n+1$-space.*

Proof. Let K be an abstract simplicial complex with vertices $v_1, \dots, v_k$. Choose pairwise different numbers $t_1, \dots, t_k$ and consider the points $\sigma(v_i) = (t_i, t_i^2, t_i^3, \dots, t_i^{2n+1})$ $(i = 1, \dots, k)$ in $\mathbb{R}^{2n+1}$. To each simplex $v_{i_1}, \dots, v_{i_m}$ from K we assign the geometric simplex with vertices $\sigma(v_{i_1}), \dots, \sigma(v_{i_m})$. We must only verify that the geometric simplices having no common vertices do not intersect.

By assumption, the dimensions of the geometric simplices under consideration are not larger than n, i.e., each of them has at most $n+1$ vertices. The number of vertices in two such simplices is at most $2n+2$. Therefore, it is sufficient to show that any $2n+2$ or fewer different points on the curve

$(t, t^2, t^3, \dots, t^{2n+1})$ are the vertices of a (nondegenerate) simplex. If the number of points is precisely $2n+2$, then the volume of this simplex is equal to

$$\pm\frac{1}{(2n+1)!}\begin{vmatrix} 1 & \tau_1 & \dots & \tau_1^{2n+1} \\ \dots & \dots & \dots & \dots \\ 1 & \tau_{2n+2} & \dots & \tau_{2n+2}^{2n+1} \end{vmatrix} \neq 0.$$

Smaller sets can be extended to $(2n+2)$-point sets in an arbitrary way. □

Remark 3.3. Points $x_1, \dots, x_k$ in the space $\mathbb{R}^N$ are said to be *generic*, or *in general position*, if no $m+1$ of these points belong to the same $(m-1)$-dimensional affine subspace for $m \le N$. To construct a realization of an n-dimensional abstract simplicial complex (with vertices $v_1, \dots, v_k$) in $\mathbb{R}^{2n+1}$, it suffices to specify generic points $x_1, \dots, x_k$ in $\mathbb{R}^{2n+1}$. An explicit construction of generic points was given in the proof of Theorem 3.15; yet another construction is as follows. First, take two different points x_1 and x_2 in $\mathbb{R}^N$. Then choose a point x_3 not belonging to the straight line x_1x_2. After that take a point x_4 not belonging to the plane $x_1x_2x_3$, and so on. In this way, we construct a set $\{x_1, \dots, x_{N+1}\}$. Then, we draw hyperplanes through all of its N-point subsets, take a point x_{N+2} belonging to none of these hyperplanes, draw hyperplanes through all N-point subsets of the new set, choose a point belonging to none of them, and so on.

A simplicial subcomplex $L \subset K$ is said to be *complete* if any set of vertices of L that spans a simplex of the complex K spans a simplex of the complex L.

Problem 29. Prove that a simplicial subcomplex $L \subset K$ is complete if and only if every simplex from K with boundary contained in L is itself contained in L.

Problem 30. Let $L \subset K$ be a simplicial subcomplex, and let L' and K' be the barycentric subdivisions of L and K. Prove that the subcomplex $L' \subset K'$ is complete.

2.4. Simplicial Approximations. Simplicial maps are much simpler than continuous maps. For example, for any two simplicial complexes K and L, there are only finitely many simplicial maps $K \to L$. Nevertheless, any continuous map can be approximated by simplicial maps (although, such an approximation may require passing from the complexes K and L to their subdivisions). It is most important for homotopic topology that any continuous map of simplicial complexes is homotopic to a certain simplicial map. This fact substantially facilitates studying homotopy classes of maps, but its proof requires a certain effort.

Suppose that K and L are simplicial complexes and $f\colon |K| \to |L|$ is a continuous map. For each point $x \in |K|$, consider the point $f(x) \in |L|$. It is an interior point of precisely one simplex from L. We say that a simplicial map $\varphi\colon K \to L$ is a *simplicial approximation* of the map f if, for each $x \in |K|$, the point $\varphi(x)$ belongs to the simplex whose interior contains $f(x)$.

Theorem 3.16. *Any simplicial approximation φ of the map f is homotopic to f.*

Proof. Let $f_t(x)$ be the point dividing the segment with endpoints $\varphi(x)$ and $f(x)$ in the ratio $t : (1-t)$. Then f_t is a homotopy between $f_0 = \varphi$ and $f_1 = f$. □

When dealing with simplicial approximations, it is more convenient to use another definition of simplicial approximation, which employs the notion of a star. Let K be a simplicial complex, and let Δ be a simplex in K. The *star* of the simplex Δ is defined as the union of the interiors of all simplices from K containing Δ. The *star* of a point $x \in |K|$ is the star of the simplex from K that contains x in its interior. The star of a simplex that has maximum dimension is the interior of this simplex. We denote the star of a simplex Δ by $\operatorname{st}\Delta$ and the star of a point x by $\operatorname{st} x$.

Theorem 3.17. *A simplicial map $\varphi\colon K \to L$ is a simplicial approximation for a continuous map $f\colon |K| \to |L|$ if and only if $f(\operatorname{st} v) \subset \operatorname{st}\varphi(v)$ for each vertex v of K.*

Proof. First, suppose that φ is a simplicial approximation of f and v is a vertex of K. Take $x \in \operatorname{st} v$. Consider the simplex Δ_K with vertex v whose interior contains x and the simplex Δ_L whose interior contains $f(x)$. On one hand, $\varphi(x)$ is an interior point of the simplex $\varphi(\Delta_K)$ with vertex $\varphi(v)$, and on the other hand, it belongs to the simplex Δ_L. Therefore, $\Delta_L \supset \varphi(\Delta_K) \ni \varphi(v)$, which means that $f(x) \in \operatorname{int}\Delta_L \subset \operatorname{st}\varphi(v)$.

Now, suppose that $f(\operatorname{st} v) \subset \operatorname{st}\varphi(v)$ for any vertex v of K. Take $x \in |K|$ and let $v_0, \dots, v_n$ be the vertices of the simplex Δ from K whose interior contains x. We have

$$f(x) \in f\left(\bigcap_{i=0}^{n} \operatorname{st} v_i\right) \subset \bigcap_{i=0}^{n} f(\operatorname{st} v_i) \subset \bigcap_{i=0}^{n} \operatorname{st}\varphi(v_i) = \operatorname{int}\varphi(\Delta).$$

Therefore, $\varphi(\Delta)$ is the simplex that contains $f(x)$ in its interior. It remains to note that $\varphi(x) \in \varphi(\Delta)$, because $x \in \Delta$. □

Corollary. *If $\varphi\colon K \to L$ and $\psi\colon L \to M$ are simplicial approximations of continuous maps $f\colon |K| \to |L|$ and $g\colon |L| \to |M|$, then $\psi\varphi$ is a simplicial approximation of the map gf.*

Let K be a finite simplicial complex, and let $K^{(n)}$ be its nth barycentric subdivision. Note that the maximal diameter of simplices from $K^{(n)}$ tends to zero as $n \to \infty$ (see p. 81).

Theorem 3.18 (simplicial approximation theorem). (a) *Suppose that K and L are simplicial complexes, the complex K is finite, and $f: |K| \to |L|$ is a continuous map. Then, for some $n \geq 0$, there exists a simplicial map $\varphi: K^{(n)} \to L$ which is a simplicial approximation for f.*

(b) *If, in addition, the restriction of f to a subcomplex $K_1 \subset K$ is simplicial, then the simplicial approximation φ can be chosen so that it coincide with f on K_1.*

Proof. (a) The stars of the vertices of the complex L form an open cover of the topological space $|L|$. The preimage of this cover under f is an open cover $\mathcal{U}$ of the compact subset $|K|$ of Euclidean space. According to Lebesgue's theorem on open covers (Theorem 2.6 on p. 59), there exists a number $\delta > 0$ such that any subset $B \subset |K|$ of diameter less than δ is contained in an element of the cover $\mathcal{U}$.

Choose n such that the diameter of any simplex from $K^{(n)}$ is less than $\delta/2$. The required simplicial map $\varphi: K^{(n)} \to L$ is defined as follows. Let v be a vertex of $K^{(n)}$. Then the diameter of $\operatorname{st} v$ is less than δ; therefore, the set $f(\operatorname{st} v)$ is contained entirely in the star $\operatorname{st} w$ of some vertex w of L. We set $\varphi(v) = w$ (if there are several suitable vertices w, we take any one of them). We have defined a map of 0-skeletons. We must verify that this map is admissible, i.e., $\varphi(v_1), \dots, \varphi(v_k)$ are the vertices of a simplex from L whenever $v_1, \dots, v_k$ are the vertices of a simplex from $K^{(n)}$. Note that vertices $v_1, \dots, v_k$ form a simplex Δ if and only if $\bigcap_{i=1}^k \operatorname{st} v_i = \operatorname{st} \Delta \neq \varnothing$. Suppose that $v_1, \dots, v_k$ are the vertices of a simplex from $K^{(n)}$. Then $\bigcap_{i=1}^k \operatorname{st} v_i \neq \varnothing$, and hence $\bigcap_{i=1}^k f(\operatorname{st} v_i) \neq \varnothing$. But $\bigcap_{i=1}^k \operatorname{st} \varphi(v_i) \supset \bigcap_{i=1}^k f(\operatorname{st} v_i) \neq \varnothing$; therefore, the vertices $\varphi(v_1), \dots, \varphi(v_k)$ form a simplex in L.

By Theorem 3.17, the simplicial map $\varphi: K^{(n)} \to L$ is a simplicial approximation of f.

(b) Let v be a vertex of K_1. Then $f(v) = w$ is a vertex L. If the subdivision $K^{(n)}$ is sufficiently fine (i.e., the number n is sufficiently large), then $f(\operatorname{st} v) \subset \operatorname{st} w$ for this subdivision, and we can set $\varphi(v) = w$. □

The simplicial approximation theorem enables us to prove the following assertion.

Theorem 3.19. *Any continuous map $f: S^n \to S^m$, where $n < m$, is homotopic to a constant map (which takes S^n to one point).*

Proof. It is sufficient to prove that f is homotopic to a nonsurjective map φ. Indeed, if $\varphi(x) \neq \xi_0 \in S^m$ for all $x \in S^n$, then

$$\varphi_t(x) = \frac{t\varphi(x) - (1-t)\xi_0}{\|t\varphi(x) - (1-t)\xi_0\|}$$

is a homotopy between φ and the constant map $S^n \to -\xi_0 \in S^m$.

The sphere S^n can be represented as a simplicial complex K that is the n-skeleton of an $(n+1)$-simplex. Similarly, the sphere S^m can be represented as a simplicial complex L. The continuous map $f\colon |K| \to |L|$ has a simplicial approximation $\varphi\colon K^{(N)} \to L$. The map φ is not surjective because its image is contained in the n-skeleton of the complex L, and it is homotopic to f by Theorem 3.16. □

Example. Let K be a triangulation of an n-simplex L with vertices $v_0, v_1, \dots, v_n$. Suppose that the vertices of K are labeled by the numbers $0, 1, \dots, n$. We define a simplicial map $\varphi\colon |K| \to |L|$ by assigning the vertex v_i to each vertex $a \in K$ labeled by i. The map φ is a simplicial approximation of the identity map $|K| \to |K| = |L|$ if and only if the set of labels satisfies the conditions of Sperner's lemma, i.e., the label of every vertex a belonging to some face of L coincides with the number of one of the vertices of this face.

The following theorem is derived from Sperner's lemma (to be more precise, from its refined version given in Theorem 2.27 on p. 83); we could formulate it without employing the notion of simplicial maps, but the statement would then look unnatural.

Theorem 3.20 (combinatorial Lefschetz formula [73]). *Suppose that K is a triangulation of an n-simplex L, $\varphi\colon K \to L$ is a simplicial map, and φ_i is the signed[1] number of i-simplices $\Delta^i \subset K$ for which $\Delta^i \subset \varphi(\Delta^i)$. Then $\varphi_0 - \varphi_1 + \varphi_2 - \dots + (-1)^n\varphi_n = 1$.*

Proof. Suppose that $v_0, v_1, \dots, v_n$ are the vertices of the simplex L, m is its barycenter, and points a_i are chosen so that m belongs to the segments $[a_i, v_i]$ and $|a_i m| = k|mv_i|$, where $k > 0$ is a fixed number. If k is sufficiently large, then the simplex $L = [v_0, \dots, v_n]$ is contained inside the simplex $[a_0, \dots, a_n]$.

Consider the simplicial complex K_1 whose vertex set consists of the points $a_0, \dots, a_n$ and the vertices of the complex K (recall that K is a triangulation of L); the simplices of K_1 are those of K and all simplices for which one of the vertices is a_i and the other vertices are those of a simplex from K contained in the face $[v_0, \dots, \hat{v}_i, \dots, v_n]$. An example of such a complex K_1 for $n = 2$ is given in Figure 1. We define a simplicial map $\psi\colon K_1 \to L$

[1] If the simplices Δ^i and $\varphi(\Delta^i)$ have the same orientation, then we take the plus sign, and if they have opposite orientations, then we take the minus sign. Note that if $\Delta^i \subset \varphi(\Delta^i)$, then the simplex $\varphi(\Delta^i)$ has the same dimension as Δ^i.

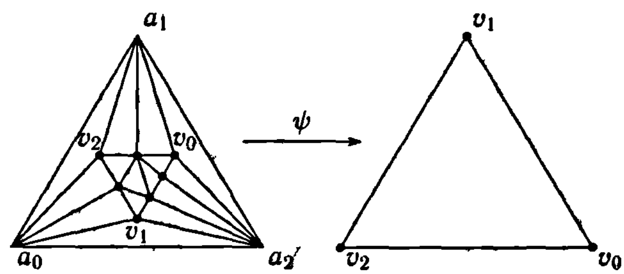

Figure 1. The construction of the complex K_1

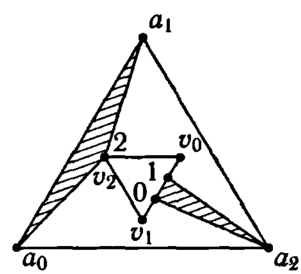

Figure 2. Orientation of simplices

in such a way that it coincides with φ on $K' \subset K_1$ and takes a_i to v_i. We label each vertex a of K_1 by the number i of the vertex $v_i = \psi(a)$. This set of labels satisfies the conditions of Sperner's lemma; therefore, $\psi_n = 1$. It remains to prove that $\psi_n = \varphi_0 - \varphi_1 + \varphi_2 - \cdots + (-1)^n \varphi_n$.

First, consider the case $n = 2$, which is the easiest to represent visually (see Figure 2). Each vertex $v_i \in K_1$ labeled by i corresponds to a completely labeled simplex $[v_i, a_j, a_k]$ in K_1. This simplex has positive orientation (i.e., the same as the simplex $[a_0, a_1, a_2]$). If a one-dimensional simplex $[x, y]$ from K_1 with labels p and q contributes to φ_1, then, by assumption, it is contained in the segment $[v_p, v_q]$. The edge $[x, y]$ corresponds to a completely labeled simplex $[x, y, a_r]$, where $r \neq p, q$. The orientation of this simplex is opposite to the orientation of $[x, y]$ induced by the edge $[v_p, v_q]$, because the simplices $[a_0, a_1]$ and $[v_0, v_1]$ have opposite orientations (these simplices lie on parallel straight lines, so their orientations can be compared). Finally, if a simplex $[x, y, z]$ from K_1 contributes to φ_2, then it is completely labeled. Moreover, the orientations of the simplex $[x, y, z]$ with respect to $[v_0, v_1, v_2]$ and with respect to $[a_0, a_1, a_2]$ coincide, because the simplices $[v_0, v_1, v_2]$ and $[a_0, a_1, a_2]$ have the same orientations.

For an arbitrary n, the argument is similar. The signs alternate because the orientations of the simplices $[v_{i_0}, v_{i_1}, \dots, v_{i_k}]$ and $[a_{i_0}, a_{i_1}, \dots, a_{i_k}]$ are the same for even k and opposite for odd k. □

2.5. Nerves of Covers. An arbitrary family $\mathcal{U} = \{U_\alpha\}$ of subsets of a set X determines a simplicial complex $N = N(\mathcal{U})$ whose vertices $\{v_\alpha\}$ are in one-to-one correspondence with the sets $\{U_\alpha\}$; a set $v_{\alpha_0}, \dots, v_{\alpha_k}$ is a simplex if and only if $U_{\alpha_0} \cap \dots \cap U_{\alpha_k} \neq \varnothing$. If X is a topological space and $\mathcal{U}$ is its cover (not necessarily open), then N is called the *nerve* of the cover $\mathcal{U}$.

Example. Let K be a simplicial complex with vertices $\{v_\alpha\}$, and let $U_\alpha = \operatorname{st} v_\alpha$ be the star of the vertex v_α, i.e., the union of the interiors of all simplices containing v_α. Then the nerve of the cover $\{U_\alpha\}$ coincides with K.

Proof. Vertices $v_{\alpha_0}, \dots, v_{\alpha_k}$ form a simplex Δ^k if and only if $\operatorname{st} v_{\alpha_0} \cap \dots \cap \operatorname{st} v_{\alpha_k} = \operatorname{st} \Delta^k \neq \varnothing$. □

We say that an open cover $\mathcal{U}$ of a space X is *contractible* if all nonempty finite intersections $U_{\alpha_0} \cap \dots \cap U_{\alpha_k}$ are contractible. The nerve of a contractible cover carries a great deal of information on the homotopy structure of X. For example, the following assertion is valid.

Theorem 3.21. *If $\mathcal{U} = \{U_\alpha\}$ is a contractible locally finite cover of a paracompact space X, then the nerve $N = N(\mathcal{U})$ is homotopy equivalent to X.*

Proof. The contractibility of the cover and the paracompactness of the space are used at different places in the proof. Thus, first we assume $\mathcal{U}$ to be an arbitrary locally finite open cover of an arbitrary space X. Let us construct an auxiliary space $X_{\mathcal{U}}$. For each nonempty intersection $U_{\alpha_0} \cap \dots \cap U_{\alpha_n} = U_{\alpha_0 \dots \alpha_n}$, consider the direct product $U_{\alpha_0 \dots \alpha_n} \times \Delta^n_{\alpha_0 \dots \alpha_n}$, where $\Delta^n_{\alpha_0 \dots \alpha_n}$ is the simplex with vertices $\alpha_0, \dots, \alpha_n$. In the disjoint union of the topological spaces thus obtained, we identify each point (x, y), where $x \in U_{\alpha_0 \dots \alpha_n}$ and $y \in [\alpha_0 \dots \hat{\alpha}_i \dots \alpha_n] \subset \Delta^n_{\alpha_0 \dots \alpha_n}$, with the corresponding point of the space $U_{\alpha_0 \dots \hat{\alpha}_i \dots \alpha_n} \times \Delta^{n-1}_{\alpha_0 \dots \hat{\alpha}_i \dots \alpha_n}$ by using the inclusion $U_{\alpha_0 \dots \alpha_n} \subset U_{\alpha_0 \dots \hat{\alpha}_i \dots \alpha_n}$.

Step 1. If the space X is paracompact, then $X \sim X_{\mathcal{U}}$.

Let $p\colon X_{\mathcal{U}} \to X$ be the map induced by the natural projections $U_{\alpha_0 \dots \alpha_n} \times \Delta^n_{\alpha_0 \dots \alpha_n} \to U_{\alpha_0 \dots \alpha_n}$. This map is continuous, because the cover $\mathcal{U}$ is open. Every point of the set $p^{-1}(x)$ can be represented as $\sum_\alpha t_\alpha x_\alpha$, where $t_\alpha \geq 0$, $\sum t_\alpha = 1$, and the summation is over all α such that $x \in U_\alpha$; by x_α we denote the point x together with the fact that $x \in U_\alpha$. This sum is finite, because the cover is locally finite.

Since X is paracompact, there exists a partition of unity $\{\varphi_\alpha\}$ subordinate to the cover $\{U_\alpha\}$, i.e., such that $\operatorname{supp} \varphi_\alpha \subset U_\alpha$. Consider the map

$s\colon X \to X_{\mathcal{U}}$ defined by $s(x) = \sum \varphi_\alpha(x)x_\alpha$ (if $\varphi_\alpha(x) = 0$, then the corresponding term is zero, and if $\varphi_\alpha(x) \neq 0$, then $x \in U_\alpha$; by x_α we denote the point x together with the fact that $x \in U_\alpha$). Clearly, $p \circ s = \mathrm{id}_X$. We must only verify that $s \circ p \sim \mathrm{id}_{X_{\mathcal{U}}}$. Suppose that a point x belongs to sets $U_{\alpha_0}, \dots, U_{\alpha_n}$ and does not belong to any other U_α. Then the points $y = \sum t_\alpha x_\alpha$ and $s(p(y)) = \sum \varphi_\alpha x_\alpha$ belong to the simplex with vertices $x_{\alpha_0}, \dots, x_{\alpha_n}$. The required homotopy uniformly moves $s(p(y))$ to y along the segment joining these points.

Step 2. If the cover $\mathcal{U}$ is contractible, then $X_{\mathcal{U}} \sim |N(\mathcal{U})|$.

Consider the space $X_{\mathcal{U}}$. Let $p\colon X_{\mathcal{U}} \to |N(\mathcal{U})|$ be the map induced by the natural projections $U_{\alpha_0\dots\alpha_n} \times \Delta^n_{\alpha_0\dots\alpha_n} \to \Delta_{\alpha_0\dots\alpha_n}$. Then the preimage of a vertex α is U_α, the preimage of an interior point of an edge $[\alpha, \beta]$ is $U_{\alpha,\beta}$, the preimage of an interior point of a simplex $[\alpha, \beta, \gamma]$ is $U_{\alpha,\beta,\gamma}$, etc. Let us successively contract these preimages to points (first, we contract the preimage U_α of each vertex, then the preimages of all interior points of edges, and so on). As a result, we obtain the space $|N(\mathcal{U})|$. □

2.6. Pseudomanifolds. A finite simplicial complex K is called an n-dimensional *pseudomanifold* if the following conditions hold:

- K is *homogeneous*, i.e., each simplex of K is a face of some n-simplex;
- K is *unramified*, i.e., each $(n-1)$-simplex of K is a face of no more than two n-simplices;
- K is *strongly connected*, i.e., for any two n-simplices Δ^n_a and Δ^n_b, there exists a sequence of simplices $\Delta^n_1 = \Delta^n_a, \Delta^n_2, \dots, \Delta^n_k = \Delta^n_b$ in which every two neighboring terms Δ^n_i and Δ^n_{i+1} have a common $(n-1)$-face.

The union of all $(n-1)$-simplices in an n-pseudomanifold M^n such that each of them is a face of precisely one n-simplex is called the *boundary* of M^n and denoted by ∂M^n. If $\partial M^n = \varnothing$, then the pseudomanifold M^n is said to be *closed*. Any $(n-1)$-simplex in a closed pseudomanifold M^n is a face for exactly two n-simplices.

An *orientation* of a simplex $\Delta^n \subset \mathbb{R}^n$ is one of the two possible families of all reference frames in $\mathbb{R}^n$ of the same orientation with origins at the points of the simplex Δ^n. For $n > 0$, each simplex admits exactly two orientations. If a simplex is endowed with an orientation, then this orientation is said to be *positive*, and the opposite orientation is said to be *negative*.

An orientation of a simplex Δ^n induces an orientation of its face $\Delta^{n-1} \subset \Delta^n$ as follows. Take a positively oriented frame with origin at a point $x \in \Delta^{n-1}$ such that its first $n-1$ vectors belong to Δ^{n-1} and the last

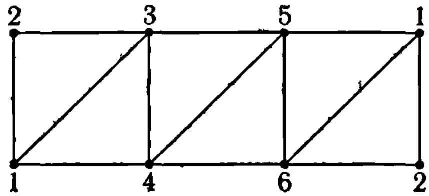

Figure 3. A triangulation of the Möbius band

vector is directed inside Δ^n. The orientation of Δ^{n-1} determined by the first $n-1$ vectors is declared to be positive.

A pseudomanifold M^n is *orientable* if all of its n-simplices can be oriented in such a way that any two simplices having a common $(n-1)$-face induce opposite orientations on this face. It follows from the strong connectedness condition that if a pseudomanifold is orientable, then it admits exactly two orientations. An orientable pseudomanifold M^n with a fixed orientation is said to be *oriented*.

Example. Let us represent the Möbius band as an abstract simplicial complex with six vertices (see Figure 3). Any realization of this abstract simplicial complex, including a realization in $\mathbb{R}^5$ (which exists according to Theorem 3.15 on p. 102), is a nonorientable pseudomanifold.

Example. Suppose that $M^n \subset \mathbb{R}^m$ is a pseudomanifold (possibly, with boundary). Let us embed $\mathbb{R}^m$ into $\mathbb{R}^{m+1}$ and take a point a in $\mathbb{R}^{m+1} \setminus \mathbb{R}^m$. The union of all segments of the form $[a, x]$, where $x \in M^n$, is an $(n+1)$-dimensional pseudomanifold. It is called the *suspension* of M^n and denoted by ΣM^n.

Remark 3.4. The suspension of a usual (topological or smooth) closed manifold M^n can be a manifold only if M^n is a homology sphere. Thus, the class of pseudomanifolds is larger than that of manifolds. On the other hand, if M^n is a pseudomanifold and $(M^n)^{n-2}$ is its $(n-2)$-skeleton, then $M^n \setminus (M^n)^{n-2}$ is a manifold; thus, any pseudomanifold becomes a manifold after removing a set of codimension 2.

Example. The suspension over the two-dimensional pseudomanifold shown in Figure 4 is a three-dimensional pseudomanifold whose boundary is not a pseudomanifold (the strong connectedness condition is violated).

2.7. The Degree of a Map to a Euclidean Space. Let M^n be an n-pseudomanifold. We say that a map $f\colon M^n \to \mathbb{R}^m$ is *simplicial* if its restriction to each simplex is affine. Any simplicial map $M^n \to \mathbb{R}^m$ is completely determined by its restriction to the 0-skeleton $(M^n)^0$, and any map $(M^n)^0 \to \mathbb{R}^m$ can be extended to a simplicial map $M^n \to \mathbb{R}^m$.

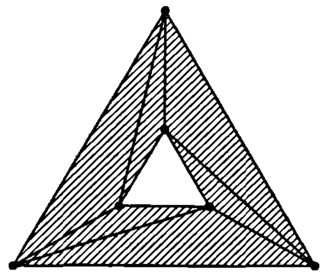

Figure 4. The boundary of the suspension of this pseudomanifold is not a pseudomanifold

Consider a simplicial map $f\colon M^n \to \mathbb{R}^n$ (the dimensions are equal). We say that a point $y \in \mathbb{R}^n$ is a *regular value* of f if y does not belong to the image of the $(n-1)$-skeleton of M^n. In particular, any point $y \in \mathbb{R}^n$ that does not belong to the image of K is a regular value of f. The set of regular values of f is everywhere dense in $\mathbb{R}^n$.

Suppose that M^n is a compact oriented pseudomanifold, $f\colon M^n \to \mathbb{R}^n$ is a simplicial map, and y is a regular value of f. The *degree* of the map f with respect to the point y is defined as

$$\deg(f, y) = \sum_{x \in f^{-1}(y)} \operatorname{sign} J_f(x),$$

where $\operatorname{sign} J_f(x)$ is the sign of the determinant of the linear part of the affine map f at x; thus, $\operatorname{sign} J_f(x) = 1$ if f preserves the orientation of the simplex Δ^n whose interior contains x and $\operatorname{sign} J_f(x) = -1$ otherwise. The sum under consideration is finite because the number of n-simplices in M^n is finite and each of these simplices contains at most one point of the set $f^{-1}(y)$.

Theorem 3.22. *Suppose that M^n is a compact oriented pseudomanifold, $f\colon M^n \to \mathbb{R}^n$ is a simplicial map, and y_1 and y_2 are regular values of f. If y_1 and y_2 belong to the same connected component of $\mathbb{R}^n \setminus f(\partial M^n)$, then* $\deg(f, y_1) = \deg(f, y_2)$.

Proof. The image of the $(n-2)$-skeleton of M^n does not separate $\mathbb{R}^n$, because no $(n-2)$-dimensional subspace of $\mathbb{R}^n$ separates $\mathbb{R}^n$ (and the union of finitely many $(n-2)$-dimensional subspaces of $\mathbb{R}^n$ does not separate $\mathbb{R}^n$ either). Hence there exists a finite-sided polygonal curve L in $\mathbb{R}^n$ with endpoints y_1 and y_2 which does not intersect the image of the $(n-2)$-skeleton of M^n, does not intersect $f(\partial M^n)$, and intersects the image of the $(n-1)$-skeleton in only finitely many points $a_1, \dots, a_k$. The set $f^{-1}(a_i)$ contains no points belonging to simplices of dimension $\leq n-2$; therefore, $f^{-1}(a_i)$ is a finite union of sets each of which either consists of one interior point of an

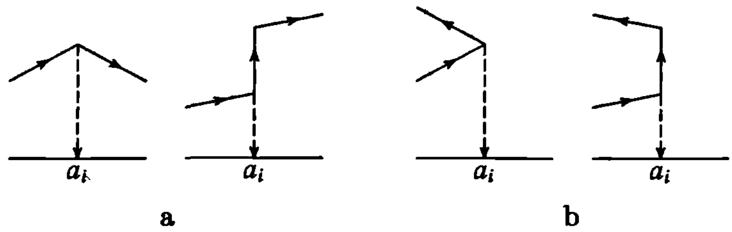

Figure 5. Passing through a critical value

$(n-1)$-face of M^n or is a straight line segment contained inside an n-face and joining interior points of two $(n-1)$-faces.

By assumption, $f^{-1}(a_i) \cap \partial M^n = \varnothing$; therefore, to each of the sets (points and segments) forming $f^{-1}(a_i)$ correspond two n-simplices. Namely, to each interior point of an $(n-1)$-simplex correspond the two n-simplices for which the given simplex is a common face, and to each segment inside an n-simplex correspond the n-simplices adjacent to the $(n-1)$-faces which contain the endpoints of this segment. If the images of these simplices have the same orientation, then the number of preimages, as well as the signs of the determinants at the preimages, do not change in passing through the point a_i (see Figure 5a; we mean only the preimages belonging to the two simplices under consideration). If the images of the simplices have opposite orientations, then two preimages at which the determinants are of opposite signs either appear or disappear (see Figure 5b). Thus, the sum of the signs of the determinants does not change. □

Theorem 3.23. *Let M^n be an oriented pseudomanifold, and let $f, g\colon M^n \to \mathbb{R}^n$ be simplicial maps with coinciding restrictions to ∂M^n. If y is a regular value of f and g, then* $\deg(f, y) = \deg(g, y)$.

Proof. Consider the family of maps $f_t = (1-t)f + tg$. Clearly, $f_0 = f$, $f_1 = g$, and the restriction of f_t to ∂M^n does not depend on t.

Let X be the connected component of $\mathbb{R}^n \setminus f_t(\partial M^n)$ that contains y (X does not depend on t, because $f_t(\partial M^n)$ does not depend on t). The set X is open; therefore, the regular values of each map f_t are everywhere dense in it. In particular, each map f_t has a regular value $y_t \in X$. According to Theorem 3.22, $\deg(f_t, y_t)$ does not depend on the choice of a regular value $y_t \in X$; hence the function $\varphi(t) = \deg(f_t, y_t)$ is well defined.

For any regular value $y_t \in X$ of f_t, there exists an $\varepsilon > 0$ with the following properties: for all $\tau \in (t-\varepsilon, t+\varepsilon) \cap [0,1]$, the point y_t is a regular value for the maps f_τ, the preimages $f_\tau^{-1}(y_t)$ contain the same number of points, and the determinants at the respective points have the same sign. Indeed, if y_t is an interior point of the image of an n-simplex under an

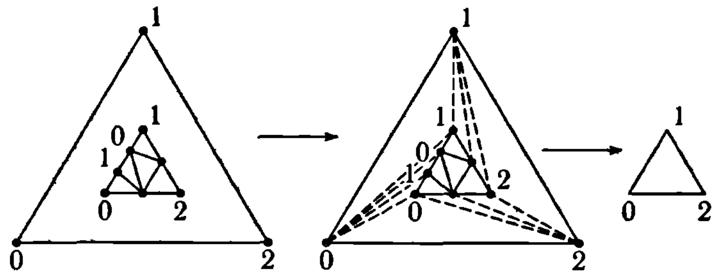

Figure 6. The supplementary triangulation of a simplex

affine map to $\mathbb{R}^n$, then it remains an interior point of the image when the affine map slightly changes. Thus, the function $\varphi(t)$ is constant on the set $U_t = (t-\varepsilon, t+\varepsilon) \cap [0,1]$. The family of sets $\{U_t\}$, where $t \in [0,1]$, forms an open cover of the interval $[0,1]$. Considering a finite subcover of this cover, we see that the function $\varphi(t)$ is constant on the entire interval $[0,1]$; therefore, $\varphi(0) = \varphi(1)$, i.e., $\deg(f, y_1) = \deg(f, y_0)$. □

Theorem 3.23 readily implies Sperner's lemma (and even its refined version given in Theorem 2.27 on p. 83). We consider only the case of two-dimensional simplices; this is sufficient to illustrate the general idea of the proof (a detailed proof for n-simplices is given in [**130**]). We embed a triangulated simplex in a larger simplex and triangulate the new object so as to prevent the appearance of new completely labeled simplices (see Figure 6); this can easily be done. The large triangulated simplex is an oriented pseudomanifold. Consider a simplicial map f from this pseudomanifold to a fixed completely labeled simplex Δ in $\mathbb{R}^n$ (each vertex is mapped to a vertex with the same label). On the boundary, the map f coincides with the identity map; therefore, its degree (with respect to the interior points of Δ) is 1. But the degree of f equals the difference between the numbers of completely labeled simplices with positive and negative orientations. By construction, all completely labeled simplices are contained in the given triangulation.

2.8. The Borsuk–Ulam Theorem. In 1933, K. Borsuk [**17**] proved the following assertion, which had been conjectured earlier by S. Ulam.

Theorem 3.24 (Borsuk Ulam). *Let $f\colon S^n \to \mathbb{R}^n$ be a continuous map. Then $f(x) = f(-x)$ for some point $x \in S^n$.*

The points x and $-x$ are said to be *antipodal*; for this reason, the Borsuk–Ulam theorem is also known as the *antipodal theorem*.

A map $g\colon S^n \to \mathbb{R}^n$ is called *odd*, or *antipodal*, if $g(-x) = -g(x)$ for any x. It is easy to see that the Borsuk–Ulam theorem is equivalent to the following statement.

Theorem 3.25. *If $g: S^n \to \mathbb{R}^n$ is an odd map, then $g(x) = 0$ for some $x \in S^n$.*

Indeed, if $f: S^n \to \mathbb{R}^n$ is an arbitrary map, then the map $g(x) = f(x) - f(-x)$ is odd, and the equality $f(x) = f(-x)$ is equivalent to $g(x) = 0$. Conversely, if $g: S^n \to \mathbb{R}^n$ is an odd map, then the Borsuk Ulam theorem implies $g(x) = g(-x)$ for some $x \in S^n$. On the other hand, $g(-x) = -g(x)$; hence $g(x) = 0$.

We derive Theorem 3.25 from the following assertion.

Theorem 3.26. *Let $h: D^n \to \mathbb{R}^n$ be a map such that its restriction to $S^{n-1} = \partial D^n$ is odd. Then $h(x) = 0$ for some point $x \in D^n$.*

To derive Theorem 3.25 from Theorem 3.26, it suffices to represent D^n as the section of the disk D^{n+1} (with boundary S^n) by a plane passing through its center and define h to be the composition of the projection of D^n onto the hemisphere and the map g.

Proof of Theorem 3.26. Instead of D^n, we shall consider the n-cube I^n, where $I = [-1, 1]$. This cube is symmetric with respect to the origin. Suppose that the restriction of $h: I^n \to \mathbb{R}^n$ to ∂I^n is odd and $0 \notin h(I^n)$. The set $h(I^n)$ is compact and, hence, is disjoint from some n-disk centered at 0. Let r be the radius of this disk.

Since the map h is uniformly continuous, the image of any simplex from a sufficiently fine triangulation of the cube I^n is contained in a disk of diameter $\varepsilon < r$. For each vertex v of such a triangulation, we set $h_\varepsilon(v) = h(v)$ and extend the map h_ε to the entire simplex by linearity. Every point $x \in I^n$ belongs to some simplex of the triangulation; therefore, the points $h(x)$ and $h_\varepsilon(x)$ belong to the same disk of diameter ε, which means that $\|h(x) - h_\varepsilon(x)\| \leq \varepsilon$.

We can assume that the triangulation of the cube I^n is symmetric with respect to the origin. In this case, the restriction of $h_\varepsilon(x)$ to ∂I^n is odd. Moreover, the relation $\|h(x) - h_\varepsilon(x)\| \leq \varepsilon$ implies $0 \notin h_\varepsilon(I^n)$.

To obtain a contradiction, it is sufficient to construct a simplicial map $\varphi: I^n \to \mathbb{R}^n$ such that its restriction to ∂I^n coincides with h_ε and $\deg(\varphi, 0)$ is odd. Indeed, on the one hand, $\deg(h_\varepsilon, 0) = 0$; on the other hand, $\deg(\varphi, 0) = \deg(h_\varepsilon, 0)$ by Theorem 3.23.

Clearly, if $\varphi: I^n \to \mathbb{R}^n$ is an odd map and 0 is its regular value, then $\deg(\varphi, 0)$ is odd. Indeed, $\varphi^{-1}(0)$ consists of 0 and pairs of the form $(x, -x)$, and the parity of the sum $\sum \pm 1$ depends only on the number of terms.

The required odd map φ is very easy to construct. We take an arbitrary map from the interior vertices of the symmetric triangulation defined in constructing h_ε to $\mathbb{R}^n$ such that it takes symmetric vertices to symmetric points. For the vertices $v \in \partial I^n$, we set $\varphi(v) = h_\varepsilon(v)$. Then, we extend φ by linearity.

It remains to overcome the last, purely technical, difficulty: the point 0 is a vertex of the triangulation, and therefore $0 = \varphi(0)$ is not a regular value. It can be rendered regular as follows. Let W be the union of all simplices with vertex 0. We can assume that $W \cap \partial I^n = \varnothing$. For each vertex $v \in W$, we set $\varphi(v) = v$. The restriction $\varphi \restriction W$ is the identity map; therefore, it remains simplicial under any triangulation of W. Now, we can turn $0 \in \mathbb{R}^n$ into a regular value of φ by slightly moving the vertices that do not belong to ∂I^n. $\square$

The following theorem is equivalent to the Borsuk–Ulam theorem.

Theorem 3.27. *If $m > n \geq 1$, then there exists no odd map $f\colon S^m \to S^n$.*

Indeed, if $m > n \geq 1$, then any odd map $S^m \to S^n$ is also an odd map $S^m \to S^n \subset \mathbb{R}^{n+1} \setminus \{0\} \subset \mathbb{R}^m \setminus \{0\}$, and Theorem 3.27 follows from the Borsuk–Ulam theorem. Conversely, given an odd map $S^m \to \mathbb{R}^m \setminus \{0\}$, it is easy to construct an odd map $S^m \to S^{m-1}$.

Remark 3.5. A very simple proof of the Borsuk–Ulam theorem is given in [**84**]. The proof of Theorem 3.26 given above has much in common with [**43**].

Problem 31*. (a) (Tucker's lemma [**126**]) Suppose that the n-cube I^n is triangulated so that the triangulation of its boundary ∂I^n is symmetric with respect to the center; suppose also that the vertices of the triangulation are labeled by the numbers $\pm 1, \pm 2, \dots, \pm n$ in such a way that for any vertex $v \in \partial I^n$, the labels of v and $-v$ add to zero. Prove that this triangulation has a pair of adjacent (i.e., joined by an edge) vertices such that their labels add to zero.

(b) Derive the Borsuk–Ulam theorem from Tucker's lemma.

Remark 3.6. A purely algebraic proof of the Borsuk–Ulam theorem for polynomial maps is given in [**104**].

2.9. Corollaries and Generalizations of the Borsuk–Ulam Theorem. The Borsuk–Ulam theorem has many interesting corollaries. Borsuk included one of them in the same paper [**17**] where he proved the Borsuk–Ulam theorem, although this corollary had already been proved by Lyusternik and Shnirelman [**76**, p. 26].

Theorem 3.28 (Lyusternik–Shnirelman). *If the sphere S^n is covered by closed sets $F_1, \dots, F_{n+1}$, then one of these sets contains a pair of antipodal points of the sphere.*

Proof. We denote the set symmetric to F_i with respect to the center of the sphere by $-F_i$. Let us show that if $F_i \cap (-F_i) = \varnothing$ for $i = 1, \ldots, n$, then $F_{n+1} \cap (-F_{n+1}) \neq \varnothing$.

Applying Urysohn's lemma (see p. 56) to the disjoint closed sets F_1 and $-F_1$ in $\mathbb{R}^{n+1}$, we can construct a continuous function $\varphi_1 \colon S^n \to [0,1]$ such that $\varphi_1(F_1) = 0$ and $\varphi_1(-F_1) = 1$. (Urysohn's lemma gives a function f for which $f(F_1) = -1$ and $f(-F_1) = 1$; we set $\varphi_1 = (1+f)/2$.) Similarly, for $F_2, \ldots, F_n$, we construct functions $\varphi_2, \ldots, \varphi_n$. Consider the map $\varphi \colon S^n \to \mathbb{R}^n$ defined by $\varphi(x) = (\varphi_1(x), \ldots, \varphi_n(x))$. According ~~to~~ the Borsuk Ulam theorem, there exists a point $x_0 \in S^n$ for which $\varphi(x_0) = \varphi(-x_0)$. If $x \in \pm F_i$ for $i = 1, \ldots, n$, then $\varphi_i(x) - \varphi_i(-x) = \pm 1$; therefore, $\varphi_i(x) \neq \varphi_i(-x)$. Thus, $x_0 \notin \bigcap_{i=1}^n F_i$ and $x_0 \notin \bigcap_{i=1}^n (-F_i) = -\bigcap_{i=1}^n F_i$. Hence $x_0 \in F_{n+1}$ and $-x_0 \in F_{n+1}$. □

Theorem 3.29. *Let $F_1, \ldots, F_n$ be measurable subsets of $\mathbb{R}^n$. Then there exists a hyperplane which cuts each set F_i into two parts of the same volume.*

Proof. Take $x \in S^{n-1} \subset \mathbb{R}^n$. Suppose that the sphere S^{n-1} is centered at the origin. For $c \in \mathbb{R}$, we set[2]

$$\Pi_c(x) = \{y \in \mathbb{R}^n : (y, x) = c\}.$$

It is easy to verify that for each vector $x \in S^{n-1}$, there exists a unique number $c \in \mathbb{R}$ such that the hyperplane $\Pi_c(x)$ cuts F_1 into two parts of equal volume. We set $\varphi_1(x) = c$. The hyperplane cutting F_1 into two equal-volume parts is the same for x and $-x$. Clearly, $\Pi_{-c}(-x) = \Pi_c(x)$, whence $\varphi_1(-x) = -c$. Let $\varphi_2, \ldots, \varphi_n$ be similar functions for $F_2, \ldots, F_n$ and consider the map $\varphi \colon S^{n-1} \to \mathbb{R}^{n-1}$ defined by

$$\varphi(x) = \big(\varphi_n(x) - \varphi_1(x), \ldots, \varphi_n(x) - \varphi_{n-1}(x)\big).$$

Clearly, $\varphi(x) = -\varphi(-x)$. Therefore, by the Borsuk Ulam theorem, there exists a point $x_0 \in S^{n-1}$ for which $\varphi(x_0) = 0$, i.e., $\varphi_1(x_0) = \varphi_2(x_0) = \cdots = \varphi_n(x_0) = c$. The hyperplane $\Pi_c(x_0)$ has the required property. □

It is easy to prove that the length of any closed centrally symmetric curve on the unit sphere S^m is at least 2π (any centrally symmetric curve contains two antipodal points, and the length of any arc joining two such points is at least π). This assertion has the following generalization.

Problem 32* ([13]). Let S^n and S^m be the unit spheres, and let $\varphi \colon S^n \to S^m$ be an odd map. Prove that the n-dimensional volume of $\varphi(S^n)$ is not less than the n-dimensional volume of S^n.

[2]Here, (y, x) is the inner product in $\mathbb{R}^n$.

The Borsuk Ulam theorem implies also a nonlinear generalization of the well-known *Radon theorem*: *If a set $A \subset \mathbb{R}^n$ contains at least $n+2$ points, then A has disjoint subsets B and C such that their convex hulls have a common point.* To prove the Radon theorem, it suffices to consider the case where A consists of exactly $n+2$ points; in this case, the theorem can be stated as follows: *If $f\colon \Delta^{n+1} \to \mathbb{R}^n$ is a linear map, then Δ^{n+1} has two disjoint faces with intersecting images.* The nonlinear generalization of this theorem consists in replacing the linear map f by an arbitrary continuous map f. Namely, the following assertion is valid; we state it as a problem.

Problem 33* ([12]). (a) Given a nondegenerate (i.e., containing an $(n+1)$-disk) convex polyhedron P in $\mathbb{R}^{n+1}$ and a continuous map $f\colon \partial P \to \mathbb{R}^n$, prove that P has disjoint faces[3] with intersecting images.

(b) Prove that if $f\colon \partial\Delta^{n+1} \to \mathbb{R}^n$ is a continuous map and $\Delta_1^n, \dots, \Delta_{n+2}^n$ are the n-faces of the simplex Δ^{n+1}, then $\bigcap_{i=1}^{n+2} f(\Delta_i^n) \neq \varnothing$.

3. CW-Complexes

In homotopy topology, the CW-complexes, which were introduced by Whitehead [141], are most convenient in many respects. The CW-complexes are obtained from the closed disks D^n by attaching their boundaries $\partial D^n = S^{n-1}$ to each other. For this reason, we start with the general operation of attaching via a map.

3.1. Attaching via Maps. The operation of *attaching* a space X to a space Y via a map $\varphi\colon A \to Y$, where $A \subset X$, is defined as follows. Consider the disjoint union $X \sqcup Y$ of X and Y. On $X \sqcup Y$, we introduce an equivalence relation by declaring that $a \sim \varphi(a)$ for all $a \in A$. We denote the quotient space modulo this equivalence by $Y \cup_\varphi X$.

The natural projection $Y \to Y \cup_\varphi X$ is always injective, while the natural projection $X \to Y \cup_\varphi X$ is injective only if the map $\varphi\colon A \to Y$ is injective; the restriction of the natural projection to $X \setminus A$ is injective.

A set $U \subset Y \cup_\varphi X$ is open (closed) if and only if its preimages under the natural projection $p\colon X \sqcup Y \to Y \cup_\varphi X$ are open (closed) in X and Y.

Example. Suppose that $X = \mathbb{R}$, $A = \{x \in \mathbb{R} : x < 0\}$, $Y = \mathbb{R}$, and $\varphi\colon A \to Y$ is the identity map, i.e., $\varphi(x) = x$ for all $x \in A$ (see Figure 7). Then the space $Y \cup_\varphi X$ is not Hausdorff: the images of $0 \in X$ and $0 \in Y$ in $Y \cup_\varphi X$ are different, but they have no disjoint neighborhoods.

The space from this example is not Hausdorff because the attachment is performed along a nonclosed set.

[3] The faces may have dimensions less than n.

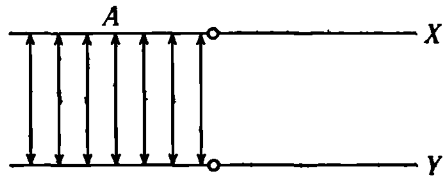

Figure 7. Attaching via a map

Theorem 3.30. *If X and Y are normal topological spaces, $A \subset X$ is a closed set, and $\varphi: A \to Y$ is a continuous map, then the space $Y \cup_\varphi X$ is normal.*

Proof. First, we prove that every point $c \in Y \cup_\varphi X$ is a closed set. If $c \in p(X \setminus A)$ or $c \in p(Y) \setminus p(A)$, then $p^{-1}(c)$ consists of one point (belonging to X or Y). If $c \in p(A)$, then the preimage of c in Y consists of one point $\bar{c}$, and the preimage of c in X is the set $\varphi^{-1}(\bar{c})$, which is closed because the map φ is continuous.

Let C_1 and C_2 be disjoint closed subsets of $Y \cup_\varphi X$. Then the set $C = C_1 \cup C_2$ is closed, and the function $f: C \to I$ which takes the value 0 on C_1 and the value 1 on C_2 is continuous. Therefore, it is sufficient to show that any continuous function $f: C \to I$, where C is a closed subset of $Y \cup_\varphi X$, can be extended to the entire space $Y \cup_\varphi X$.

Let $C \subset Y \cup_\varphi X$ be a closed set, and let $f: C \to I$ be a continuous function. Consider the closed sets $C_X = p^{-1}(C) \cap X$ and $C_Y = p^{-1}(C) \cap Y$. On these sets, f determines the functions $f_X: C_X \xrightarrow{p} C \xrightarrow{f} I$ and $f_Y: C_Y \xrightarrow{p} C \xrightarrow{f} I$. By the Tietze extension theorem, the function f_Y can be extended to a continuous function $F_Y: Y \to I$. On the set A, F_Y determines the function $f_A: A \xrightarrow{\varphi} Y \xrightarrow{F_Y} I$. The continuous functions f_X and f_A coincide on $C_X \cap A$; therefore, they determine a continuous function $f_{XA}: C_X \cup A \to I$.

Now, it is time to use the closedness of A. We have to construct an extension to $Y \cup_\varphi X$ of the function f_{XA}, which is defined on $C_X \cup A$, where C_X is a closed set. By assumption, the set A is closed; hence $C_X \cup A$ is closed as well. By the Tietze extension theorem, the function f_{XA} can be extended to a function $F_X: X \to I$. Moreover, if $x \in A$, then $F_X(x) = F_Y(\varphi(x))$. Hence the functions F_X and F_Y determine a function F on $Y \cup_\varphi X$; it is continuous, because F_X and F_Y are continuous. By construction, $F \restriction C = f$; thus, F is the required extension of f. $\square$

3.2. Definition of CW-Complexes. A topological space X is called a *CW-complex* if $X = \bigcup_{i=0}^{\infty} X^i$, where X^0 is a discrete space and X^{i+1} is obtained by attaching the disjoint union $\bigsqcup_{\alpha \in A} D_\alpha^{i+1}$ of $(i+1)$-disks to X^i

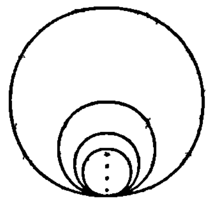

Figure 8. A space not homeomorphic to a CW-complex

via a continuous map $\varphi\colon \bigsqcup_{\alpha\in A} S^i_\alpha \to X^i$, where $S^i_\alpha = \partial D^{i+1}_\alpha$. Moreover, X must satisfy conditions (c) and (w) specified below.

The images of D^{i+1}_α and $\operatorname{int} D^{i+1}_\alpha$ under the natural projection onto $X^{i+1} \subset X$ are called, respectively, *closed* and *open cells* of dimension $i+1$. The conditions (c) and (w) are as follows:

(c) each closed cell intersects only finitely many open cells;

(w) a set $C \subset X$ is closed if and only if all intersections of C with closed cells are closed.

Note that if the number of cells is finite, then conditions (c) and (w) hold automatically. Condition (w) determines the topology of the space X if the dimensions of cells in X are unbounded.

The symbols (c) and (w) are abbreviations for "closure finite" and "weak topology."

The open cells are pairwise disjoint and cover the space X.

The space X^i is called the i-dimensional *skeleton* of the CW-complex X. If a CW-complex X has cells of dimension n and no cells of larger dimensions, then X is said to be an *n-dimensional* CW-complex.

The natural map $\chi^{i+1}_\alpha\colon D^{i+1}_\alpha \to X^{i+1} \subset X$ is called the *characteristic map* of the cell.

Example. Suppose that X^0 is $S^1 = \partial D^2$ with discrete topology, $X^1 = X^0$, and $X^2 = X$ is obtained by attaching D^2 to X^0 via the identity map $S^1 \to S^1$. Then the space X satisfies (w) but does not satisfy (c).

Example. Let S^1_n be the circle of radius $1/n$ centered at $(0, 1/n)$. Consider the union $X = \bigcup_{n=1}^\infty S^1_n$ (see Figure 8) with topology induced by $\mathbb{R}^2$. The natural one-to-one map $f\colon X \to Y$, where Y is the CW-complex consisting of one 0-cell and the cells D^1_n attached to it at both endpoints $(n = 1, 2, \dots)$, is not a homeomorphism.

Proof. Take a point x_n different from the origin on each circle S^1_n. Let F be the subset of X consisting of the points x_n, where $n = 1, 2, \dots$. It is not closed, because $\lim_{n\to\infty} x_n = (0,0) \notin F$. On the other hand, the set $f(F)$ is closed, because it intersects each closed 1-cell in precisely one point. $\square$

One of the most important advantages of CW-complexes is that continuous maps of a CW-complex can be constructed by induction on skeletons; extending maps defined on the boundaries of cells inside the cells by continuity, we necessarily obtain a continuous map $f: X \to Y$ of the entire CW-complex, because $f^{-1}(C)$ is closed if and only if its intersections with all closed cells are closed.

Let X be a CW-complex. A *subcomplex* of X is a closed subspace which is a union of open cells.

Below, we give some of the most important examples of CW-complexes. The sphere S^n is a CW-complex with one 0-cell and one n-cell. On S^n, it is easy to define the structure of a CW-complex with two cells of each dimension from 0 to n by induction: it suffices to attach the northern and southern hemispheres to the equator $S^{n-1} \subset S^n$.

The *real projective space* $\mathbb{R}P^n$ is the quotient of $\mathbb{R}^{n+1} \setminus \{0\}$ modulo the equivalence relation defined by $x \sim \lambda x$ for all $\lambda \in \mathbb{R} \setminus \{0\}$. Replacing $\mathbb{R}$ by $\mathbb{C}$, we obtain the definition of the *complex projective space* $\mathbb{C}P^n$.

To every point $(x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} \setminus \{0\}$ corresponds the point $(x_1 : \dots : x_{n+1}) \in \mathbb{R}P^n$; the numbers $x_1, \dots, x_{n+1}$ are called the *homogeneous coordinates* of this point in $\mathbb{R}P^n$. For $\mathbb{C}P^n$, the notation is similar. The map $(x_1 : x_2) \mapsto x_1/x_2$ is a homeomorphism between $\mathbb{R}P^1 \setminus \{(1:0)\}$ and $\mathbb{R}^1$; therefore, $\mathbb{R}P^1 \approx S^1$. Similarly, $\mathbb{C}P^1 \approx S^2$.

To introduce the structure of a CW-complex on $\mathbb{R}P^n$, consider the map $f: D^n \to \mathbb{R}P^n$ defined by

$$f(x_1, \dots, x_n) = \left(x_1 : \dots : x_n : \sqrt{1 - x_1^2 - \dots - x_n^2}\right).$$

The image of the boundary $S^{n-1} \subset D^n$ is contained in

$$\mathbb{R}P^{n-1} = \{(x_1 : \dots : x_n : x_{n+1}) \in \mathbb{R}P^n : x_{n+1} = 0\}.$$

The map f is a homeomorphism of $\operatorname{int} D^n$ to $\mathbb{R}P^n \setminus \mathbb{R}P^{n-1}$; the inverse map is given by

$$(x_1 : \dots : x_{n+1}) \mapsto (\lambda^{-1} x_1 x_{n+1}, \dots, \lambda^{-1} x_n x_{n+1}),$$

where $\lambda > 0$ and $\lambda^2 = x_{n+1}^2(x_1^2 + \dots + x_n^2)$. Thus, $\mathbb{R}P^n$ is obtained from $\mathbb{R}P^{n-1}$ by attaching one cell of dimension n.

Similarly, $\mathbb{C}P^n$ is obtained from $\mathbb{C}P^{n-1}$ by attaching one cell of dimension $2n$. We assume that

$$D^{2n} = \{(z_1, \dots, z_n) \in \mathbb{C}^n : |z_1|^2 + \dots + |z_n|^2 \leq 1\}.$$

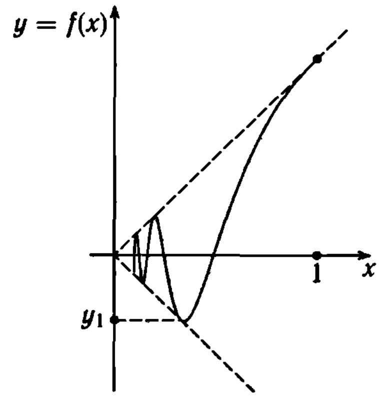

Figure 9. The graph of the function f

Consider the map $f\colon D^{2n} \to \mathbb{C}P^n$ defined by

$$f(z_1, \dots, z_n) = \left(z_1 : \dots : z_n : \sqrt{1 - |z_1|^2 - \dots - |z_n|^2}\right).$$

On $\mathbb{C}P^n \setminus \mathbb{C}P^{n-1}$, the inverse map has the form

$$(z_1 : \dots : z_{n+1}) \mapsto (\lambda^{-1} z_1 \bar{z}_{n+1}, \dots, \lambda^{-1} z_n \bar{z}_{n+1}),$$

where $\lambda > 0$ and $\lambda^2 = |z_{n+1}|^2(|z_1|^2 + \dots + |z_n|^2)$. Thus, $\mathbb{C}P^n$ admits the structure of a CW-complex with cells of dimensions $2i$ for $i = 0, 1, \dots, n$.

Problem 34. Prove that $\mathbb{C}P^n$ is obtained from $D^{2n} \subset \mathbb{C}^n$ by identifying each point x of $\partial D^{2n} = S^{2n-1}$ with λx for all $\lambda \in \mathbb{C}$ such that $|\lambda| = 1$.

The same constructions as those used to obtain the CW-complexes S^n, $\mathbb{R}P^n$, and $\mathbb{C}P^n$ give the CW-complexes S^∞, $\mathbb{R}P^\infty$, and $\mathbb{C}P^\infty$.

Problem 35. Prove that the space S^∞ is contractible.

CW-complexes resemble simplicial complexes in many respects. It can be even proved that any CW-complex is homotopy equivalent to a simplicial complex (the proof of this assertion is given in, e.g., [**99, 112**]). But there exist also CW-complexes which are not homeomorphic to simplicial complexes. To construct an example of such a CW-complex, consider the continuous function on the interval $I = [0, 1]$ defined by $f(x) = x \sin(\pi/2x)$ for $x > 0$ and $f(0) = 0$ (see Figure 9); the image of I under f is the interval $[y_1, 1]$.

Consider the map $I^2 \to \mathbb{R}^3$ given by $(x, y) \mapsto (x, xy, f(y))$ (see Figure 10). In the plane $x = 1$, its graph coincides with that of f. In the plane $x = c$, where $0 < c \leq 1$, this is the same graph shrank by a factor of c in the y-direction. Finally, in the plane $x = 0$, the graph is the segment consisting of the points $(0, 0, z)$, where $z \in [y_1, 1]$.

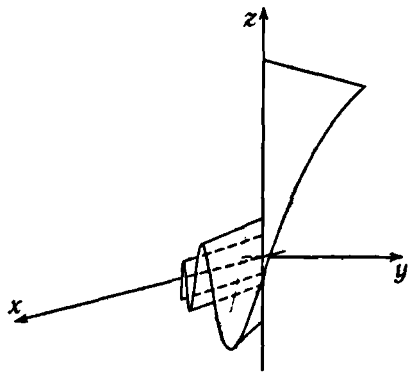

Figure 10. The graph of the map of the square

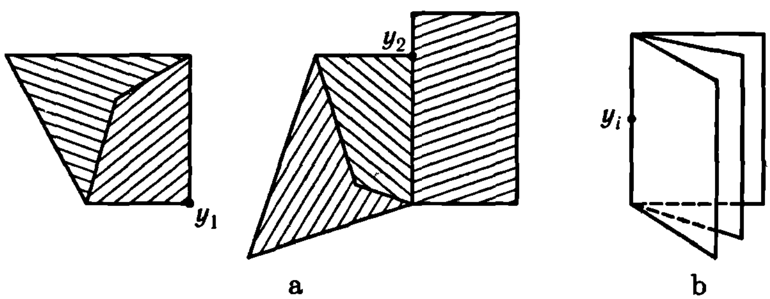

Figure 11. The structure of neighborhoods of the points y_k

Consider the CW-complex X whose 0-cells are the images of the vertices of the square I^2 and the point $(0, 0, y_1)$ under this map, 1-cells are the images of the sides of the square and the segment of the z-axis between 0 and y_1, and 2-cell is the image of the square.

It is easy to show that this CW-complex X is homeomorphic to no simplicial complex, i.e., X is a *nontriangulable* CW-complex. Indeed, X is a compact topological space; therefore, a simplicial complex homeomorphic to X cannot have infinitely many vertices. On the other hand, all the points $(0, 0, y_i)$, where y_i is the value of f at a point of local maximum or minimum, must be vertices of any simplicial complex homeomorphic to X because of the structure of small neighborhoods of these points. The two simplest cases of these neighborhoods are shown in Figure 11a. In the other cases, several half-planes are added; the additional half-planes are shown in Figure 11b.

3.3. Topological Properties. CW-complexes have many good topological properties. Thus, any CW-complex is Hausdorff (and even normal); in CW-complexes, connectedness is equivalent to path-connectedness; any

CW-complex is locally contractible; any CW-complex is paracompact. In this section, we prove these and other properties of CW-complexes.

Theorem 3.31. *Any CW-complex is a normal topological space.*

Proof. First, we prove that every skeleton X^n of a CW-complex X is normal. For $n = 0$, this assertion is obvious: each point in the discrete space X^0 is both open and closed. The induction step is Theorem 3.30.

Now, we prove the normality of X. Let $C \subset X$ be a closed set, and let $f: C \to I$ be a continuous function. The function f determines a continuous function f_0 on $C \cap X^0$, which can be extended to a function F_0 on X^0. The functions f and F_0 determine a continuous function f_1 on the closed set $(C \cap X^1) \cup X^0$, which can be extended to X^1. Continuing this procedure, we obtain a function $F: X \to I$ which is continuous on each skeleton and, in particular, on each closed cell. Condition (w) implies the continuity of F. □

Problem 36. Prove that any compact subset of a CW-complex intersects only finitely many open cells.

In CW-complexes, connectedness coincides with path-connectedness; moreover, there is a fairly simple criterion for a CW-complex to be connected.

Theorem 3.32. (a) *A CW-complex X is connected if and only if its 1-skeleton X^1 is connected.*

(b) *A CW-complex is connected if and only if it is path-connected.*

Proof. (a) If $n \geq 2$, then attaching the disk D^n to the skeleton X^{n-1} via a map $S^{n-1} \to X^{n-1}$ does not change the number of connected components. Indeed, for $n \geq 2$, the image of S^{n-1} under a continuous map is connected; therefore, it is entirely contained in one connected component. Moreover, any space obtained by attaching D^n to a connected space is connected.

Clearly, if the skeletons X^n of a CW-complex X are connected for $n \geq 1$, then X itself is connected, and if the X^n are disconnected, then X is disconnected as well.

(b) For 1-dimensional CW-complexes, there is no difference between connectedness and path-connectedness. Thus, in the proof of assertion (a), we can replace "connectedness" by "path-connectedness" because the sphere S^{n-1} and the disk D^n with $n \geq 2$ are both connected and path-connected. □

A topological space X is said to be *locally contractible* if, for each point $x \in X$ and every open set $U \ni x$, there exists a contractible open set V such that $x \in V \subset U$ (the contractibility of V means that the identity map

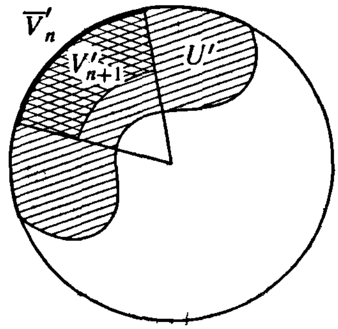

Figure 12. The construction of the set V'_{n+1}

$V \to V$ is homotopic to the constant map $V \to x$). Local contractibility is very useful in the theory of coverings.

Theorem 3.33. *Any CW-complex is locally contractible.*

Proof. Let X be a CW-complex. We construct a contractible neighborhood V of a given point x by induction on the dimension of the skeletons; this V will also satisfy the condition $\overline{V} \subset U$.

Each point $x \in X$ is contained in a unique open cell $\operatorname{int} e^m_\alpha \approx \operatorname{int} D^m_\alpha$. The set $\operatorname{int} e^m_\alpha \cap U$ is open in the topology of X^m. Let V_m be the open disk centered at x of radius so small that $\overline{V}_m \subset \operatorname{int} e^m_\alpha \cap U$, and let $f^m_t \colon V_m \to V_m$ be a homotopy between the identity map and the map $V_m \to V_m$ that takes the entire disk V_m to the point x.

Suppose that $n \geq m$ and the required neighborhood V_n of x in X^n, together with the homotopy f^n_t, has already been constructed. Let us construct the neighborhood V_{n+1} of x in X^{n+1} and the homotopy f^{n+1}_t. Suppose that $\chi \colon D^{n+1} \to X^{n+1}$ is the characteristic map of some cell. Then $\overline{V}'_n = \chi^{-1}(\overline{V}_n)$ is a closed subset of $S^n \subset D^{n+1}$ and $U' = \chi^{-1}(U)$ is an open (in the topology of D^{n+1}) subset of D^{n+1}; moreover, $\overline{V}'_n \subset U'$, because $\overline{V}_n \subset U$. The set $\overline{V}'_n$ is compact; therefore, for some $\varepsilon \in (0,1)$, the set

$$\overline{V}'_{n+1} = \{tv : 1 - \varepsilon \leq t \leq 1, v \in \overline{V}'_n\}$$

is contained in U' (see Figure 12). The set

$$V'_{n+1} = \{tv : 1 - \varepsilon < t \leq 1, v \in V'_n\}$$

is open in D^{n+1}, and its closure coincides with $\overline{V}'_{n+1}$. It is easy to construct a homotopy between the identity map $V'_{n+1} \to V'_{n+1}$ and the natural projection from V'_{n+1} to V'_n treated as a map $V'_{n+1} \to V'_{n+1}$ such that it is the identity on V'_n and its image is contained in V'_n. After constructing such neighborhoods V'_{n+1} and the corresponding homotopies for all $(n+1)$-cells, we obtain a neighborhood V_{n+1} of X in X^{n+1} such that $\overline{V}_{n+1} \subset U$; in addition, we obtain a homotopy between the identity map $V_{n+1} \to V_{n+1}$ and some map $V_{n+1} \to V_{n+1}$ such that its image is contained in V_n and its restriction to V_n is the identity map. Now, using the homotopy f_t^n, we can construct the required homotopy f_t^{n+1}.

The set $V = \bigcup_{n=0}^{\infty} V_n$ is open in X, and the homotopies f_t^n determine a homotopy between the identity map $V \to V$ and the map $V \to V$ taking the entire set V to the point x. □

Theorem 3.34. *Any CW-complex is paracompact.*

Proof. Let $\mathcal{U} = \{U_\alpha : \alpha \in \mathcal{A}\}$ be an open cover of a CW-complex X. We construct a locally finite refinement $\mathcal{V} = \{V_\beta : \beta \in \mathcal{B}\}$ of $\mathcal{U}$ by induction on the skeletons X^n, where $n = 0, 1, \dots$. Namely, at the nth step, we construct an open cover $\{V_{\beta,n}\}$ of X^n so that $V_{\beta,0} \subset V_{\beta,1} \subset \cdots$ and $V_\beta = \bigcup_{n=0}^{\infty} V_{\beta,n}$. Simultaneously, also by induction on n, we construct the index set $\mathcal{B}$: at the nth step, we add the indices $\mathcal{B}_n$ which correspond to the sets $V_{\beta,n}$ needed to completely cover $X^n \setminus X^{n-1}$.

For $n = 0$, we put $\mathcal{B}_0 = X^0$; for each $\beta \in \mathcal{B}_0$, the set $V_{\beta,0}$ consists of one point $\beta \in X^0$. For every $\beta \in \mathcal{B}_0$, we choose $\alpha(\beta) \in \mathcal{A}$ such that $\beta \in U_{\alpha(\beta)}$.

Suppose that $n \geq 0$ and the index sets $\mathcal{B}_0, \dots, \mathcal{B}_n$, as well as the sets $V_{\beta,n}$ for $\beta \in \mathcal{B}_0 \cup \cdots \cup \mathcal{B}_n$, have already been constructed. We construct sets $V_{\beta,n+1}$ of two different types. First, for each $\beta \in \mathcal{B}_0 \cup \cdots \cup \mathcal{B}_n$, we extend $V_{\beta,n}$ to a set $V_{\beta,n+1}$ which is open in X^{n+1} and lies in $U_{\alpha(\beta)}$. Some part of $X^{n+1} \setminus X^n$ may be not covered by the sets $V_{\beta,n+1}$. To completely cover X^{n+1}, we construct additional sets $V_{\beta,n+1}$, where $\beta \in \mathcal{B}_{n+1}$, so that each of them is contained in some $U_{\alpha(\beta)}$. We assume that $V_{\beta,0} = V_{\beta,1} = \cdots = V_{\beta,n} = \varnothing$ for $\beta \in \mathcal{B}_{n+1}$.

We begin with the extension of $V_{\beta,n}$. Suppose that $\chi\colon D^{n+1} \to X^{n+1}$ is the characteristic map of some $(n+1)$-cell, $V'_{\beta,n} = \chi^{-1}(V_{\beta,n})$, and $U'_{\alpha(\beta)} = \chi^{-1}(U_{\alpha(\beta)})$. The set $U' = \bigcup U'_{\alpha(\beta)}$ is open in the topology of D^{n+1}; therefore, $D^{n+1} \setminus U'$ is closed. Moreover, $S^n = \partial D^{n+1} \subset U'$. Hence, for some $\varepsilon \in (0,1)$, the set $\{tv : 1 - \varepsilon \leq t \leq 1, v \in S^n\}$ is contained in U'. Changing the map χ if necessary, we can assume that $\varepsilon = 1/2$. We put

$$V'_{\beta,n+1} = \{tv : 1/2 < t \leq 1, v \in V'_{\beta,n}\}.$$

The set $V_{\beta,n+1}$ is defined as the union of all $\chi(V'_{\beta,n+1})$ over all $(n+1)$-cells. Clearly, $V_{\beta,n} \subset V_{\beta,n+1} \subset U_{\alpha(\beta)}$ and $V_{\beta,n+1}$ is open in X^{n+1}.

Now, let us construct the additional sets $V_{\beta,n+1}$. We again change the characteristic map χ in such a way that $\varepsilon = 1/2$. This means, in particular, that if $B = \{tv : 0 \leq t < 3/4, v \in S^n\}$, then the set $D^{n+1} \setminus B$ is already covered by the sets $V'_{\beta,n+1}$ extending the $V'_{\beta,n}$. It remains to cover B. The open sets $\operatorname{int} \chi^{-1}(U_\alpha)$ cover the compact set $\overline{B}$; hence we can choose finitely many indices $\alpha_1, \ldots, \alpha_k$ such that the sets $\operatorname{int} \chi^{-1}(U_{\alpha_i})$ cover $\overline{B}$. We put $V_{\alpha_i,n+1} = \chi(B \cap \chi^{-1}(U_{\alpha_i}))$ for $i = 1, \ldots, k$. We construct such sets for all $(n+1)$-cells.

Clearly, $V_\beta \cap X^n = V_{\beta,n}$, so V_β is open. Moreover, if $\beta \in \mathcal{B}_n$ and $V_{\beta,n} \subset U_{\alpha(\beta)}$, then $V_\beta \subset U_{\alpha(\beta)}$. Thus, it only remains to prove that the cover $\mathcal{V} = \{U_\beta : \beta \in \mathcal{B}\}$ is locally finite. First, we prove (by induction on n) that each $\mathcal{V}_n = \{U_{\beta,n} : \beta \in \mathcal{B}_0 \cup \cdots \cup \mathcal{B}_n\}$ is a locally finite cover of the skeleton X^n. For $n = 0$, this is obvious. Suppose that the required assertion is proved for the skeletons of dimensions at most n. Take any point $x \in X^{n+1}$. Let $\chi \colon D^{n+1} \to X^{n+1}$ be the characteristic map of the cell containing x.

First, suppose that x belongs to the boundary of the $(n+1)$-cell $\chi(D^{n+1})$, i.e., $x \in X^n$. By the induction hypothesis, there exists an open set $W_n \ni x$ in X^n which intersects only finitely many sets $V_{\beta,n}$. We put

$$W'_{n+1} = \{tv : 3/4 < t \leq 1, v \in \chi^{-1}(W_n)\}.$$

The set W_{n+1} is defined as the union of $\chi(W'_{n+1})$ over all $(n+1)$-cells containing x. The set W_{n+1} is open in X^{n+1}, and it is disjoint from all of the sets $V_{\beta,n+1}$ with $\beta \in \mathcal{B}_{n+1}$ (such sets correspond to $t < 3/4$).

Now, suppose that x is inside the $(n+1)$-cell $\chi(D^{n+1})$. If x is the image of the center of the disk D^{n+1}, then we put $W'_{n+1} = \{tv : 0 < t < 1/2, v \in S^n\}$ and $W_{n+1} = \chi(W'_{n+1})$; the set W_{n+1} is disjoint from all of the $V_{\beta,n+1}$ with $\beta \in \mathcal{B}_0 \cup \cdots \cup \mathcal{B}_n$. Now, suppose that x is not the image of the center. Let x' be the projection of the preimage of x onto S^n from the center. By the induction hypothesis, there exists an open set $W_n \ni x'$ in X^n which intersects only finitely many sets $V_{\beta,n}$. We put $W'_{n+1} = \{tv : 0 < t < 1, v \in \chi^{-1}(W_n)\}$ and $W_{n+1} = \chi(W'_{n+1})$. The set W_{n+1} intersects only those $V_{\beta,n+1}$ with $\beta \in \mathcal{B}_0 \cup \cdots \cup \mathcal{B}_n$ for which $W_n \cap V_{\beta,n} \neq \varnothing$. By construction, each $(n+1)$-cell intersects only finitely many sets $V_{\beta,n+1}$ with $\beta \in \mathcal{B}_{n+1}$. Therefore, W_{n+1} intersects only finitely many sets $V_{\beta,n+1}$.

Now, take a point x inside an n-cell of X. Using the construction described above, we can obtain a sequence of sets $W_n \subset W_{n+1} \subset \cdots$ such that $x \in W_m$ and W_m is open in X^m for any $m \geq n$. Moreover, each set W_m intersects only those $V_{\beta,m}$ for which $W_n \cap V_{\beta,n} \neq \varnothing$. Therefore, the set $W = \bigcup_{m=n}^{\infty} W_m$ is open in X and intersects only those V_β for which $W_n \cap V_{\beta,n} \neq \varnothing$. There are only finitely many such sets. ☐

3.4. Cellular Approximation. Let X and Y be CW-complexes. A continuous map $f \colon X \to Y$ is said to be *cellular* if $f(X^n) \subset Y^n$ for all n.

Theorem 3.35 (cellular approximation theorem). *Suppose that X and Y are CW-complexes, $A \subset X$ is a subcomplex (possibly, $A = \varnothing$), and $f\colon X \to Y$ is a continuous map whose restriction to A is cellular. Then there exists a cellular map $g\colon X \to Y$ homotopic to f; moreover, the homotopy is the identity on A.*

Proof. We construct the map g and the homotopy by induction on the dimensions of cells σ_α^n in X, considering each cell separately and extending maps already defined on the boundaries of the cells. To construct the map g, it is sufficient to consider each cell e_β^m in Y, where $m > n$, and "squeeze" the image of σ_α^n from this cell onto the boundary ∂e_β^m in such a way that the map f does not change outside $\operatorname{int} e_\beta^m$. Thus, the problem is as follows. Given a continuous map $f\colon D^n \to Y$ and the characteristic map $\chi\colon D^m \to Y$, where $m > n$ and $f(S^{n-1}) \subset Y \setminus \operatorname{int}\chi(D^m)$, construct a continuous map $g\colon D^n \to Y$ with the following properties:

(1) if $f(x) \notin \operatorname{int}\chi(D^m)$, then $g(x) = f(x)$;

(2) the map g is homotopic to f, and the homotopy is the identity outside $\operatorname{int}\chi(D^m)$;

(3) $g(D^n) \subset Y \setminus \operatorname{int}\chi(D^m)$.

Step 1. There exists a map $g\colon D^n \to Y$ such that it has properties (1) and (2) and its image does not contain at least one point $y \in \operatorname{int}\chi(D^m)$.

Let $D_\varepsilon^m = \{x \in \mathbb{R}^m \ ; \ \|x\| \leq \varepsilon\}$ (we assume that $D^m = D_1^m$). For $0 < \varepsilon < 1$, the disk D_ε^m is homeomorphic to $\chi(D_\varepsilon^m) \subset Y$. To abbreviate the notation, we identify D_ε^m with $\chi(D_\varepsilon^m) \subset Y$. On the compact set $f^{-1}(D_{3/4}^m)$, the map f is uniformly continuous; hence we can choose $\delta > 0$ such that if $x, y \in f^{-1}(D_{3/4}^m) \subset D^n$ and $\|x - y\| < \delta$, then $\|f(x) - f(y)\| < 1/4$. Consider a sufficiently fine triangulation of the disk D^n (identified with an n-simplex), in which the diameter of each simplex is less than δ. If the image of a simplex from this triangulation under f intersects $S_{1/2}^{m-1} = \partial D_{1/2}^m$, then this image is entirely contained in $D_{3/4}^m \setminus D_{1/4}^m$. All the simplices of this triangulation are divided into three disjoint classes:

(a) the simplices whose images are disjoint from $S_{1/2}^{m-1}$;

(b) the simplices whose images are entirely contained in $S_{1/2}^{m-1}$;

(c) the simplices whose images intersect $S_{1/2}^{m-1}$.

We construct the map g and the homotopy separately for each simplex from the triangulation. In case (a), we set $g(x) = f(x)$ for all points of the simplex. In case (b), we set $g(v) = f(v)$ for all vertices of the simplex and extend the obtained map by linearity. For a simplex whose image intersects $S_{1/2}^{m-1}$, the situation is most complicated, because the map is already defined

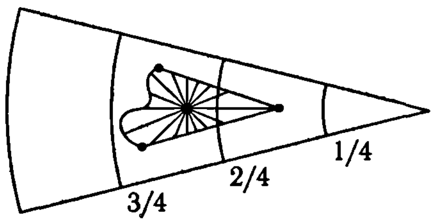

Figure 13. The first step of cellular approximation

on some of its faces (namely, on those satisfying (a) or (b)), and we must consistently extend this map over the entire simplex. For the vertices, we set $g(v) = f(v)$. On a 1-face, the map is either already defined or not. In the latter case, we extend the map from the endpoints to the entire 1-face by linearity. If the map g is not yet defined on a 2-face, then we define it as follows. The 2-face Δ^2 can be covered by segments of the form $[m, x]$, where m is the barycenter of the simplex Δ^2 and x is a point on the boundary $\partial\Delta^2$. At the boundary points x, the map g is already defined. We set $g(m) = f(m)$ and extend the obtained map to $[m, x]$ by linearity (see Figure 13). Then, the same construction is performed for 3-faces, and so on.

Let Δ^k be a simplex in the triangulation of D^n. Clearly, $g(\Delta^k)$ is contained in the convex hull of $f(\Delta^k)$. In case (c), this convex hull does not intersect $D^m_{1/4}$. Indeed, if $y_0 \in f(\Delta^k) \cap S^{m-1}_{1/2}$, then $f(\Delta^k)$ is contained in the disk of radius $1/4$ centered at y_0, and this disk does not intersect $D^m_{1/4}$.

The homotopy f_t between f and g is defined as follows. If $f(x) = g(x)$, then we set $f_t(x) = f(x)$ for all x. If $f(x) \neq g(x)$, then both points $f(x)$ and $g(x)$ belong to the disk D^m, and we set $f_t(x) = (1-t)f(x) + tg(x)$.

The intersection of $D^m_{1/4}$ with the image of g is contained in a finite union of affine planes of dimension $n < m$; therefore, the disk $D^m_{1/4}$ contains a point y not belonging to the image of g, as required.

Step 2. There exists a map $g_1\colon D^n \to Y$ with properties (1), (2), and (3).

According to Step 1, the map f can be replaced by a map g_0 whose image does not contain some point $y \in \operatorname{int}\chi(D^m)$. Consider the composition g_1 of the map g_0 and the projection from y onto the boundary of the disk (see Figure 14). It has property (3) and is homotopic to g_0; the homotopy between g_0 and g_1 is defined by $g_t = (1-t)g_0 + tg_1$. □

3.5. Geometric Realization of CW-Complexes. Let X be a CW-complex. A continuous map $i\colon X \to \mathbb{R}^n$ is called an *embedding* if it is a homeomorphism of X onto $i(X)$.

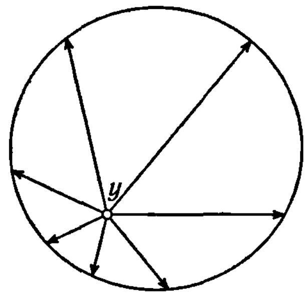

Figure 14. The second step of cellular approximation

Theorem 3.36. *For any finite CW-complex X of dimension n, there exists an embedding of X into $\mathbb{R}^{(n+1)(n+2)/2}$.*

Proof. The finite CW-complex X is compact; therefore, by Theorem 3.2 (see p. 89), any injective map $X \to \mathbb{R}^N$ is an embedding.

We prove the theorem by induction on $n = \dim X$. For $n = 0$, the assertion is obvious. Suppose that we have constructed an embedding $i_{n-1}\colon X^{n-1} \to \mathbb{R}^N$ for the $(n-1)$-skeleton X^{n-1}. Applying a translation if necessary, we can assume that $0 \notin i_{n-1}(X^{n-1})$. Let us define an embedding $i_n\colon X^{n-1} \to \mathbb{R}^N \oplus \mathbb{R}^n \oplus \mathbb{R}$ by setting

$$i_n(x) = (i_{n-1}(x), 0, 0) \in \mathbb{R}^N \oplus \mathbb{R}^n \oplus \mathbb{R}$$

for $x \in X^{n-1}$. Consider the n-cells $\chi_\alpha(D^n_\alpha)$, where $\alpha = 1, \dots, k$. Each point of the disk D^n_α can be represented in the form tx_α, where $0 \le t \le 1$ and $x_\alpha \in S^{n-1}_\alpha = \partial D^n_\alpha$, i.e., $\|x_\alpha\| = 1$. At the points x_α, we have already defined the values $i_{n-1}(\chi_\alpha(x_\alpha))$; we denote them by $i_{n-1}(x_\alpha)$ for short. At $tx_\alpha \in D^n_\alpha$, we set

$$i_n(tx_\alpha) = \begin{cases} (0, tx_\alpha, \alpha) & \text{if } t \le 1/2, \\ ((2t-1)i_{n-1}(x_\alpha), (1-t)x_\alpha, 2\alpha(1-t)) & \text{if } t \ge 1/2; \end{cases}$$

these expressions coincide at $t = 1/2$.

Let us show that the map i_n is injective. Suppose that $i_n(t_1x_\alpha) = i_n(t_2x_\beta)$. For $t_1 \le 1/2$ and $t_2 > 1/2$, we have $i_{n-1}(x_\beta) \ne 0$. If $t_1, t_2 \le 1/2$, then, obviously, $t_1 = t_2$, $x_\alpha = x_\beta$, and $\alpha = \beta$. If $t_1, t_2 \ge 1/2$, then the equality $(1-t_1)x_\alpha = (1-t_2)x_\beta$ implies $t_1 = t_2$ (recall that $\|x_\alpha\| = \|x_\beta\| = 1$); therefore, $x_\alpha = x_\beta$ and $\alpha = \beta$. $\square$

Remark 3.7. The bound for the dimension in Theorem 3.36 can be made more precise; namely, any finite CW-complex of dimension n is embedded in $\mathbb{R}^{2n+1}$. The proof of this assertion is given in [**60**].

4. Constructions

We have already encountered some constructions on topological spaces, such as direct product, wedge product, and the operation of attaching via a map. In this section, we discuss these and other constructions, as well as relations between them, in more detail. We are also interested in rendering these constructions *simplicial* (or *cellular*), i.e., such that applying them to simplicial complexes (or CW-complexes) yields simplicial complexes (or CW-complexes).

4.1. Cartesian Product. Recall that if X and Y are topological spaces, then the base of the topology of $X \times Y$ consists of the Cartesian products of open subsets of X and Y. Both projections $p_X(x, y) = x$ and $p_Y(x, y) = y$ are then continuous.

A product of two simplices of positive dimensions is not a simplex, but it is a rectilinear cell. Theorem 3.13 (see p. 100) shows that any product of two simplices can always be triangulated, i.e., represented as a simplicial complex. The remark after this theorem allows us to construct this simplicial complex without employing additional vertices.

If X and Y are CW-complexes, then the space $X \times Y$ can be partitioned into cells in a natural way. Namely, consider the cells $\varphi\colon (D^p, S^{p-1}) \to (X^p, X^{p-1})$ and $\psi\colon (D^q, S^{q-1}) \to (Y^q, Y^{q-1})$. Clearly, $D^p \times D^q \approx D^{p+q}$ and

$$S^{p+q-1} \approx \partial D^{p+q} \approx (\partial D^p \times D^q) \cup (D^p \times \partial D^q).$$

Thus, the maps φ and ψ determine a map

$$(D^{p+q}, S^{p+q-1}) \to (X^p \times Y^q, X^p \times Y^{q-1} \cup X^{p-1} \times Y^q).$$

The set $X \times Y$ is endowed with the product topology. If X and Y are finite CW-complexes, then the partitioning of $X \times Y$ into cells described above satisfies conditions (c) and (w), i.e., $X \times Y$ is a CW-complex. However, for infinite CW-complexes, condition (w) may not hold.

Problem 37. Let $S^p \vee S^q = (S^p \times \{*\}) \cup (\{*\} \times S^q) \subset S^p \times S^q$. Prove that $S^p \times S^q / S^p \vee S^q \approx S^{p+q}$.

4.2. Cylinders, Cones, and Suspensions. Let $I = [0, 1]$. For a topological space X, the space $X \times I$ is called the *cylinder* over X.

The *cone* over X is defined as the quotient $X \times I/(X \times \{1\})$ modulo the equivalence relation defined by declaring that $x_1 \times \{1\} \sim x_2 \times \{1\}$ for any $x_1, x_2 \in X$. The cone over X is denoted by CX.

The *suspension* over X is the quotient space

$$\Sigma X = X \times I / (X \times \{1\} \cup X \times \{0\}) = CX/(X \times \{0\}).$$

Exercise 22. Prove that $CS^n \approx D^{n+1}$ and $\Sigma S^n \approx S^{n+1}$.

If X is a CW-complex and A is its subcomplex, then X/A is a CW-complex. Therefore, CX and ΣX are CW-complexes. Thus, the constructions of cylinders, cones, and suspensions are cellular.

4.3. Joins. The *join* $X * Y$ of topological spaces X and Y is the quotient of $X \times I \times Y$ modulo the equivalence relation defined by declaring that $(x_1, t_1, y_1) \sim (x_2, t_2, y_2)$ if either $t_1 = t_2 = 0$ and $x_1 = x_2$ or $t_1 = t_2 = 1$ and $y_1 = y_2$. If $X, Y \subset \mathbb{R}^n$ and the segments $[x, y]$, where $x \in X$ and $y \in Y$, have no common interior points, then the join $X * Y$ has a very simple geometric description: in this case, $X * Y$ is the union of all segments $[x, y]$.

Exercise 23. Prove that $D^p * D^q \approx D^{p+q+1}$.

Exercise 24. Prove that $D^0 * X \approx CX$ and $S^0 * X \approx \Sigma X$. (Here D^0 is a singleton and S^0 is a two-point set.)

Exercise 25. Given $x \in S^p$, $y \in S^q$, and $t \in [0, 1]$, prove that the map

$$(x, t, y) \mapsto \left(\cos \frac{\pi t}{2} x, \sin \frac{\pi t}{2} y\right)$$

is a homeomorphism of $S^p * S^q$ onto S^{p+q+1}.

Problem 38. Prove that the natural embedding $f\colon X \to X * Y$ defined by $f(x) = (x, 0, y)$ (the point $(x, 0, y) \in X * Y$ does not depend on y) is homotopic to a constant map.

If $a_0, \dots, a_p, b_0, \dots, b_q$ are points in general position in $\mathbb{R}^{p+q+1}$, then the join of the simplices with vertices $a_0, \dots, a_p$ and $b_0, \dots, b_q$ is a simplex with vertices $a_0, \dots, a_p, b_0, \dots, b_q$. This remark shows that the join of disjoint abstract simplicial complexes A and B consists of the simplices of the form $\alpha \cup \beta$, where α is a simplex from A and β is a simplex from B. (Recall that a simplex of an abstract simplicial complex is merely a set of vertices.)

Theorem 3.37. *If X, Y, and Z are finite simplicial complexes, then $(X * Y) * Z \approx X * (Y * Z)$.*

Proof. Consider realizations of X, Y, and Z in $\mathbb{R}^n$ such that their vertices $\{x_\alpha\}$, $\{y_\beta\}$, and $\{z_\gamma\}$ are generic points. Then both spaces $(X * Y) * Z$ and $X * (Y * Z)$ are homeomorphic to the union of all simplices with vertices x_{i_0}, $\dots$, x_{i_p}, y_{j_0}, $\dots$, y_{j_q}, z_{k_0}, $\dots$, z_{k_r}, where $p, q, r \geq 0$. $\square$

Suppose that K is a simplicial complex, $p \in \mathbb{N}$, and $2 \leq j \leq p$. The *deleted join* $J^p_j(K)$ is defined as follows. Consider the p-fold join

$J^p(K) = K * \cdots * K$; a simplex of $J^p(K)$ is an ordered set $(\sigma_1, \ldots, \sigma_p)$ of simplices from K. Among all such sets we select those in which any j simplices have empty intersection. These sets are the simplices of the complex $J^p_j(K)$.

We denote the k-skeleton of the simplex Δ^n by $\mathrm{sk}_k\,\Delta^n$; the simplex Δ^n has the natural structure of a simplicial complex.

There is a construction based on the deleted join J^2_2 which reduces the proof of the impossibility of embedding the simplicial complex $\mathrm{sk}_n\,\Delta^{2n+2}$ in $\mathbb{R}^{2n}$ (i.e., the nonexistence of a homeomorphism of $\mathrm{sk}_n\,\Delta^{2n+2}$ onto a subset of $\mathbb{R}^{2n}$) to the Borsuk Ulam theorem. Recall that, according to Theorem 3.15 on p. 102, any finite n-dimensional simplicial complex can be embedded in $\mathbb{R}^{2n+1}$. The impossibility of embedding $\mathrm{sk}_n\,\Delta^{2n+2}$ in $\mathbb{R}^{2n}$ was proved independently by van Kampen [65] and Flores [38]. Our proof mainly follows [46].

Suppose that K is a simplicial complex and $f\colon K \to \mathbb{R}^{2n}$ is an embedding (we denote the space $|K|$ by K for short). Let CK be the cone over K. The map f determines a map $\bar{f}\colon CK \to \mathbb{R}^{2n+1}$ such that its restriction to K is one-to-one and $\bar{f}(K) \cap \bar{f}(CK \setminus K) = \varnothing$.

Consider the subspace $\hat{K}$ of $CK \times CK$ consisting of the products of all pairs of disjoint simplices in CK such that at least one of them belongs to K. It is easy to construct a homomorphism $\varphi\colon \hat{K} \to J^2_2(K)$. Indeed, each point of $\hat{K}$ has a unique representation in the form $\big(t_1x_1+(1-t_1)v, t_2x_2+(1-t_2)v\big)$, where v is the vertex of the cone CK, x_1 and x_2 are points belonging to disjoint simplices $\sigma_1, \sigma_2 \subset K$, and at least one of the numbers t_1 and t_2 equals 1. We set

$$\varphi\big(t_1x_1 + (1-t_1)v, t_2x_2 + (1-t_2)v\big) = \begin{cases} \langle x_1, \frac{t_2}{2}, x_2\rangle & \text{if } t_1 = 1; \\ \langle x_1, 1-\frac{t_2}{2}, x_2\rangle & \text{if } t_2 = 1. \end{cases}$$

On the spaces $\hat{K}$ and $J^2_2(K)$, the natural involutions $(a,b) \leftrightarrow (b,a)$ and $\langle x_1, t, x_2\rangle \leftrightarrow \langle x_2, 1-t, x_1\rangle$ are defined. The homeomorphism φ commutes with these involutions.

For the map $\bar{f}$, we define a map $\hat{f}\colon \hat{K} \to \mathbb{R}^{2n+1}$ by setting $\hat{f}(a,b) = \bar{f}(a) - \bar{f}(b)$. This map anticommutes with the involution, i.e., $\hat{f}(a,b) = -\hat{f}(b,a)$. Moreover, $\hat{f}(a,b) \neq 0$ for all $(a,b) \in \hat{K}$. Indeed, suppose that $(a,b) \in \hat{K}$ and $\bar{f}(a) = \bar{f}(b)$. Then $\bar{f}(a) = \bar{f}(b) \in \bar{f}(K)$, because one of the points a and b belongs to K. Hence both a and b belong to K, because $\bar{f}(K) \cap \bar{f}(CK \setminus K) = \varnothing$. Finally, $a = b$, because the restriction of $\bar{f}$ to K is one-to-one. On the other hand, by assumption, a and b belong to disjoint simplices.

Thus, if the deleted join $J^2_2(K)$ is homeomorphic to S^{2n+1} and the natural involution becomes the symmetry with respect to the center of the sphere under the homeomorphism, then K cannot be embedded in $\mathbb{R}^{2n}$. Indeed,

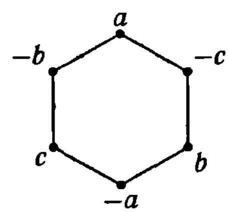

Figure 15. The deleted join of three points

if there were an embedding of K into $\mathbb{R}^{2n}$, then we could construct a map $g\colon S^{2n+1} \to \mathbb{R}^{2n+1} \setminus \{0\}$ such that $g(-x) = -g(x)$ for all $x \in S^{2n+1}$, which contradicts the Borsuk Ulam theorem.

Theorem 3.38. *The space $J_2^2(\operatorname{sk}_n \Delta^{2n+2})$ is homeomorphic to S^{2n+1}, and the homeomorphism transforms the natural involution into the symmetry with respect to the center of the sphere.*

Proof. For $n = 0$, the proof is seen in Figure 15. The points a, b, and c constitute a copy of $\operatorname{sk}_0 \Delta^2$; for the other copy we take $-a$, $-b$, and $-c$. To obtain the deleted join, we connect a to $-b$ and $-c$, and so on. It is easy to verify that the natural involution becomes the symmetry with respect to the center.

This construction is generalized to an arbitrary n as follows. For one copy of $\operatorname{sk}_n \Delta^{2n+2}$ we take the n-skeleton of a simplex with vertices $v_0, \ldots, v_{2n+2}$ in $\mathbb{R}^{2n+2}$ and place the origin at the barycenter of this simplex. The vertices $v_0, \ldots, v_{2n+2}$ can be treated as vectors with sum zero. For the second copy of $\operatorname{sk}_n \Delta^{2n+2}$ we take the n-skeleton of the simplex with vertices $-v_0, \ldots, -v_{2n+2}$.

The required space $J_2^2(\operatorname{sk}_n \Delta^{2n+2})$ is obtained as follows. In each of the sets $v_0, \ldots, v_{2n+2}$ and $-v_0, \ldots, -v_{2n+2}$, we choose $n+1$ points in such a way that all the chosen points together have pairwise different numbers. The convex hull of these points is a $(2n+1)$-dimensional simplex. If the simplices thus obtained have no common interior point, then their union is the required space.

First, note that none of the convex hulls under consideration contains 0. Indeed, the equality $\sum \lambda_i v_i = 0$ can hold only if all numbers λ_i are equal, whereas each convex hull misses one of the vectors v_i.

Let e denote the unit vector in $\mathbb{R}^{2n+2}$. We show that the ray $\{\lambda e\colon \lambda > 0\}$ intersects the set under consideration in one point. Renumbering the vectors if necessary, we can assume that $e = \sum \alpha_i v_i$, where $\alpha_0 \le \alpha_1 \le \cdots \le \alpha_{2n+2}$. Then $e = e - \alpha_{n+1} \sum v_i = \sum(\alpha_i - \alpha_{n+1}) v_i = \sum \beta_i v_i$, where $\beta_i \le 0$ for $0 \le i \le n$, $\beta_{n+1} = 0$, and $\beta_i \ge 0$ for $n+2 \le i \le 2n+2$. We set

$\beta' = -\sum_{i=0}^{n} \beta_i$, $\beta'' = \sum_{i=n+2}^{2n+2} \beta_i$, and $\beta = \beta' + \beta''$; at least one of the numbers β' and β'' is nonzero, because $e \neq 0$. The point

$$\frac{\beta'}{\beta} \sum_{i=0}^{n} \left(-\frac{\beta_i}{\beta'}\right)(-v_i) + \frac{\beta''}{\beta} \sum_{i=n+2}^{2n+2} \frac{\beta_i}{\beta'} v_i = \frac{e}{\beta}$$

belongs to one of the convex hulls under consideration (if $\beta' = 0$ or $\beta'' = 0$, then the corresponding term is assumed to be zero).

After renumbering the vectors v_i, we can represent any point y from the convex hull under consideration in the form $y = \sum \alpha_i v_i$, where $\alpha_i \leq 0$ for $0 \leq i \leq n$, $\alpha_{n+1} = 0$, $\alpha_i \geq 0$ for $n+2 \leq i \leq 2n+2$, and $\sum |\alpha_i| = 1$. Suppose that some point λy with $\lambda > 0$ belongs to one of the convex hulls. Then $\lambda y = \sum \beta_i v_i$, where $\sum |\beta_i| = 1$ and among the numbers β_i there are at most $n+1$ positive numbers and at most $n+1$ negative numbers. Clearly,

$$\sum_{i=0}^{2n+2} \left(\alpha_i - \frac{\beta_i}{\lambda}\right) v_i = y - \frac{\lambda y}{\lambda} = 0;$$

thus, all numbers $\alpha_i - \beta_i/\lambda$ are equal to the same number γ, i.e., $\beta_i = \lambda(\alpha_i - \gamma)$. If $\gamma > 0$, then $\beta_0 \leq \cdots \leq \beta_{n+1} = -\lambda\gamma < 0$, and if $\gamma < 0$, then $\beta_{2n+2} \geq \cdots \geq \beta_{n+1} = -\lambda\gamma > 0$. This contradicts the assumption that no more than $n+1$ of the β_i can be negative and no more than $n+1$ of them can be positive. Therefore, $\gamma = 0$, i.e., $\beta_i = \lambda\alpha_i$. By assumption, $\lambda > 0$ and $\sum |\alpha_i| = \sum |\beta_i| = 1$; hence $\lambda = 1$.

Thus, each ray $\{\lambda e : \lambda > 0\}$ intersects the set under consideration in precisely one point. Therefore, this set is homeomorphic to S^{2n+1}. Moreover, we have proved that each point has a unique representation in the form $\sum \alpha_i v_i$, where $\sum |\alpha_i| = 1$, at most $n+1$ of the numbers α_i are positive, and at most $n+1$ of them are negative. This means that the $(2n+1)$-simplices under consideration have no common interior points. □

Now, we give yet another example of calculation of deleted joins.

Theorem 3.39 (see [118]). $J_j^p(\Delta^n) \approx J^{n+1}(\mathrm{sk}_{j-2}\,\Delta^{p-1})$.

Proof. If $n = 0$, then $\Delta^n = *$ (one point), and the simplices of the complex $J_j^p(\Delta^n)$ have the form $(\sigma_1, \ldots, \sigma_p)$, where at most $j-1$ simplices σ_i are singletons and all of the remaining simplices are empty. Thus, $\mathrm{sk}_{j-2}\,\Delta^{p-1} = J^1(\mathrm{sk}_{j-2}\,\Delta^{p-1})$.

It is easy to verify that if A and B are arbitrary simplicial complexes, then $J_j^p(A * B) \approx J_j^p(A) * J_j^p(B)$. Indeed, in the definition of a join, all simplices from A and B are considered different (even if $A = B$). Clearly, if $a_\alpha \cap b_\beta = \varnothing$ for all α and β, then the intersection of $a_{k_1} \cup b_{l_1}, \ldots, a_{k_j} \cup b_{l_j}$ is empty if and only if $a_{k_1} \cap \cdots \cap a_{k_j} = \varnothing$ and $b_{l_1} \cap \cdots \cap b_{l_j} = \varnothing$.

Using the homeomorphism $\Delta^n \approx J^{n+1}(\Delta^0)$, we obtain

$$J_j^p(\Delta^n) \approx J_j^p(\Delta^0 * \cdots * \Delta^0) \approx J^{n+1}\left(J_j^p(\Delta^0)\right) \approx J^{n+1}(\mathrm{sk}_{j-2}\,\Delta^{p-1}). \qquad \square$$

Corollary 1. $J_p^p(\Delta^n) \approx J^{n+1}(S^{p-2}) \approx S^{(n+1)(p-1)-1}$.

Corollary 2. *The space $J_j^p(\Delta^n)$ is homotopy equivalent to a wedge of $((n+1)(j-1)-1)$-spheres.*

4.4. Symmetric Product. Let X be a topological space. The symmetric group S_n acts of the space $X^n = X \times \cdots \times X$ by the rule $\sigma(x_1, \dots, x_n) = (x_{\sigma(1)}, \dots, x_{\sigma(n)})$. The quotient space of X^n by this action is called the nth *symmetric product* of the space X and denoted by $\mathrm{SP}^n(X)$.

Exercise 26. Prove that $\mathrm{SP}^2(\mathbb{R}) \approx \{(x, y) \in \mathbb{R}^2 : y \geq 0\}$.

Theorem 3.40. $\mathrm{SP}^n(S^2) \approx \mathrm{SP}^n(\mathbb{C}P^1) \approx \mathbb{C}P^n$.

Proof. Suppose that $(a_1 : b_1), \dots, (a_n : b_n) \in \mathbb{C}P^1$ and

$$\prod_{i=1}^{n}(a_i x - b_i y) = \sum_{k=0}^{n} c_k x^k y^{n-k}.$$

When a pair $(a_i : b_i)$ is replaced by $(\lambda a_i : \lambda b_i)$, all coefficients c_k are multiplied by λ^n; hence the formula

$$((a_1 : b_1), \dots, (a_n : b_n)) \mapsto (c_0 : \cdots : c_n)$$

determines a map $\mathbb{C}P^1 \times \cdots \times \mathbb{C}P^1 \to \mathbb{C}P^n$. A permutation of $(a_1 : b_1)$, $\dots$, $(a_n : b_n)$ does not change the point $(c_0 : \cdots : c_n)$; thus, we have a map $h\colon \mathrm{SP}^n(\mathbb{C}P^1) \to \mathbb{C}P^n$.

Any polynomial in one variable decomposes into linear factors over the field $\mathbb{C}$. Therefore, any polynomial $\sum_{k=0}^{n} c_k x^k y^{n-k}$ where not all the numbers c_k are zero, can be represented in the form $\prod_{i=1}^{n}(a_i x - b_i y)$, where for each i at least one of the numbers a_i and b_i is nonzero. This means that the map h is surjective. This map is also injective because the coefficients of any polynomial determine its roots up to permutation. Clearly, h is continuous.

Thus, $h\colon \mathrm{SP}^n(S^2) \to \mathbb{C}P^n$ is a continuous one-to-one map. The space $\mathrm{SP}^n(S^2)$, being a continuous image of the compact space $S^2 \times \cdots \times S^2$, is compact. The space $\mathbb{C}P^n$ is Hausdorff because it is a CW-complex. Therefore, by Theorem 3.2 (see p. 89), the map h is a homeomorphism. $\square$

Theorem 3.41. (a) $\mathrm{SP}^n(\mathbb{C}) \approx \mathbb{C}^n$.

(b) $\mathrm{SP}^n(\mathbb{C} \setminus \{0\}) \approx \mathbb{C}^{n-1} \times (\mathbb{C} \setminus \{0\})$.

Proof. (a) The map $b \mapsto (1 : b)$ is an embedding of $\mathbb{C}$ into $\mathbb{C}P^1 \approx S^2$, and it determines an embedding of $\mathrm{SP}^n(\mathbb{C})$ into $\mathrm{SP}^n(\mathbb{C}P^1) \approx \mathbb{C}P^n$. To each point $(b_1,\dots,b_n) \in \mathrm{SP}^n(\mathbb{C})$ it assigns the coefficients of the polynomial $\prod_{i=1}^n(x-b_i)$; therefore, the image of $\mathrm{SP}^n(\mathbb{C})$ is $\mathbb{C}^n \subset \mathbb{C}P^n$. The map $\mathrm{SP}^n(\mathbb{C}) \to \mathbb{C}^n$ is a homeomorphism because the map $\mathrm{SP}^n(\mathbb{C}P^1) \to \mathbb{C}P^n$ is a homeomorphism. (A direct proof that the map $\mathrm{SP}^n(\mathbb{C}) \to \mathbb{C}^n$ is a homeomorphism is given in [49].)

(b) We must prove that if $b_1,\dots,b_n \in \mathbb{C}\setminus\{0\}$, then the coefficients of all polynomials of the form $\prod_{i=1}^n(x-b_i)$ form a set that is homeomorphic to $\mathbb{C}^{n-1}\times(\mathbb{C}\setminus\{0\})$. Clearly, the polynomial $x^n + c_{n-1}x^{n-1}+\dots+c_0$ has no zero roots if and only if $c_0 \neq 0$. Hence all but the last coefficient of the polynomial $\prod_{i=1}^n(x-b_i)$ can be arbitrary. □

Theorem 3.42. $\mathrm{SP}^n(\mathbb{R}P^2) \approx \mathbb{R}P^{2n}$.

Proof. To each point $\mathbb{R}P^2$ correspond two antipodal points of the sphere S^2. The stereographic projection takes a pair of antipodal points of S^2 to z and $-\bar{z}^{-1}$ (or to 0 and ∞), i.e., to points of the form $(a:b)$ and $(-\bar{b}:\bar{a})$ in $\mathbb{C}P^1$.

To every unordered set of n points in $\mathbb{R}P^2$ we assign the polynomial

$$f(x,y) = \prod_{i=1}^n (a_ix - b_iy)(-\bar{b}_ix - \bar{a}_iy). \tag{1}$$

When a point $(a_i:b_i)$ is replaced by $(\lambda a_i : \lambda b_i)$ or $(-\bar{b}_i:\bar{a}_i)$, the polynomial f is multiplied by $|\lambda|^2$ or -1, respectively. Thus, f is determined uniquely up to multiplication by a nonzero real number.

It is easy to verify that

$$f(-\bar{y},\bar{x}) = (-1)^n\overline{f(x,y)}. \tag{2}$$

Clearly, every homogeneous polynomial of degree $2n$ satisfying (2) can be represented in the form (1) because its roots can be arranged into pairs $(a:b)$, $(-\bar{b}:\bar{a})$.

For a polynomial of the form $\sum_{k=-n}^n c_kx^{n-k}y^{n+k}$, relation (2) is equivalent to $c_{-k} = (-1)^k\bar{c}_k$. Thus, the polynomial $f(x,y)$ is completely determined by the real coefficient c_0 and the complex coefficients $c_1,\dots,c_n$. These coefficients are arbitrary (except that they cannot vanish simultaneously). Thus, the space of all polynomials f considered up to multiplication by a nonzero real number is homeomorphic to $\mathbb{R}P^{2n}$.

We have constructed a one-to-one continuous map $h : \mathrm{SP}^n(\mathbb{R}P^2) \to \mathbb{R}P^{2n}$. The space $\mathrm{SP}^n(\mathbb{R}P^2)$ is compact, and $\mathbb{R}P^{2n}$ is Hausdorff; therefore, h is a homeomorphism. □

Remark 3.8. An interesting discussion of the properties of the homeomorphism $h : \mathrm{SP}^n(\mathbb{R}P^2) \to \mathbb{R}P^{2n}$ is contained in [9].

In conclusion, we describe, without proof, the structure of the space $\mathrm{SP}^n(S^1)$. Let $S^1 \tilde{\times} D^n$ be the space obtained from $I \times D^n$ by identifying the points $(0, x)$ and $(1, h(x))$, where $h : D^n \to D^n$ is the symmetry about a hyperplane passing through the center of the disk (for h we can take any orientation-reversing homeomorphism). Then $\mathrm{SP}^n(S^1) \approx S^1 \times D^{n-1}$ for odd n and $\mathrm{SP}^n(S^1) \approx S^1 \tilde{\times} D^{n-1}$ for even n. The proof of this assertion is given in [90].

Exercise 27. Prove that $\mathrm{SP}^2(S^1)$ is the Möbius band.

Chapter 4

Two-Dimensional Surfaces, Coverings, Bundles, and Homotopy Groups

1. Two-Dimensional Surfaces

1.1. Basic Definitions. Let M^2 be a two-dimensional pseudomanifold without boundary in which every point has a neighborhood homeomorphic to the open disk D^2. Then any topological space X homeomorphic to M^2 is called a *closed two-dimensional surface*, or a *two-dimensional surface without boundary*.

Exercise 28. Prove that the two-dimensional simplicial complex shown in Figure 1 can be supplemented to a closed two-dimensional pseudomanifold which is not a closed two-dimensional surface.

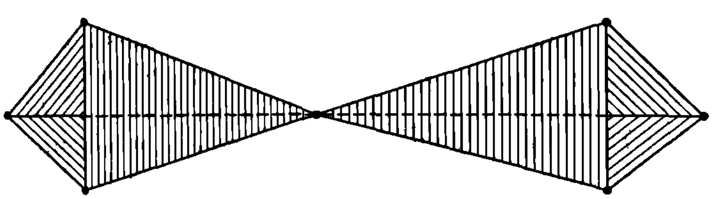

Figure 1. A pseudomanifold which is not a surface

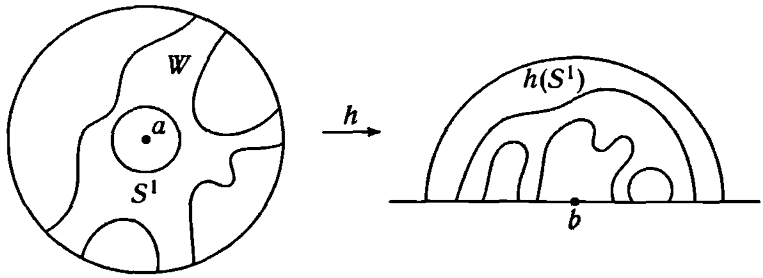

Figure 2. Invariance of the boundary

A *two-dimensional surface with boundary* is a topological space homeomorphic to a two-dimensional pseudomanifold M^2 in which every nonboundary point has a neighborhood homeomorphic to the open disk D^2 and every boundary point a has a neighborhood homeomorphic to

$$D^2_+ = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1, y \geq 0\};$$

in the latter case, the homeomorphism must take a to $(0, 0) \in D^2_+$.

To show that the boundary of a two-dimensional surface is well defined, we have to prove the following assertion.

Theorem 4.1. *If $h: M^2 \to N^2$ is a homeomorphism of two-dimensional surfaces, then $h(\partial M^2) = \partial N^2$.*

Proof. It is sufficient to verify that if $a \in M^2 \setminus \partial M^2$, then $h(a) \notin \partial N^2$. By assumption, the point $a \in M^2 \setminus \partial M^2$ has a neighborhood U homeomorphic to D^2. Suppose that $b = h(a) \in \partial N^2$. Then b has a neighborhood V homeomorphic to D^2_+, and the homeomorphism takes b to $(0, 0) \in D^2_+$. Let $W = U \cap h^{-1}(V)$ be a neighborhood of a. We identify U with D^2 and V with D^2_+ and assume that h is a homeomorphism of $W \subset D^2 \subset \mathbb{R}^2$ onto $h(W) \subset D^2_+ \subset \mathbb{R}^2$ such that $h(0, 0) = (0, 0)$.

For sufficiently small $\varepsilon > 0$, the open set W contains all points $z \in \mathbb{R}^2$ for which $\|z - a\| \leq \varepsilon$. Let S^1 be the circle of radius ε centered at a (see Figure 2). According to the Jordan theorem, the curve $h(S^1)$ divides the plane $\mathbb{R}^2 \supset D^2_+$ into two connected components, one bounded and one unbounded. On the one hand, the point $b = h(a)$ belongs to the image of the disk bounded by S^1, and the image of this disk is the bounded component. On the other hand, the set $\mathbb{R}^2_- = \{(x, y) \in \mathbb{R}^2 : y < 0\}$ does not intersect D^2_+; therefore, $\mathbb{R}^2_-$ does not intersect $h(S^1)$ either, and hence b belongs to the unbounded connected component. $\square$

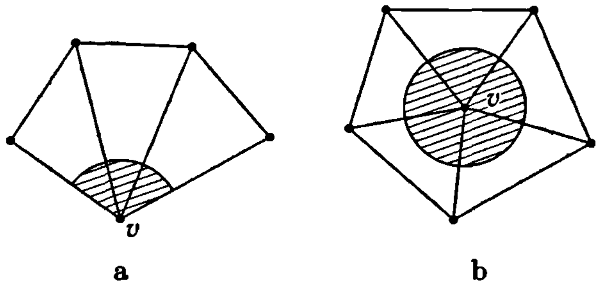

Figure 3. The structure of a neighborhood of a point on a two-dimensional surface

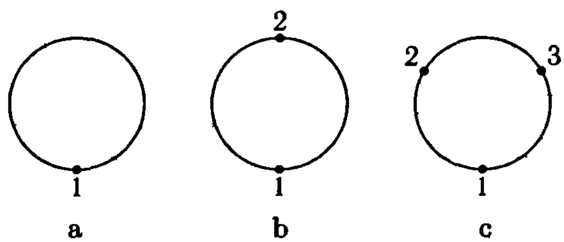

Figure 4. A triangulation of the circle

Any point of a pseudomanifold M^2 contained inside a 2-simplex has a neighborhood homeomorphic to D^2, and any point contained inside a 1-simplex has a neighborhood homeomorphic to D^2 or D^2_+. Thus, to determine whether a given pseudomanifold M^2 is a two-dimensional surface, it is sufficient to consider its vertices. Let v be a vertex of M^2. Clearly, the union of all simplices with vertex v in M^2 includes m sets of the form shown in Figure 3a and n sets of the form shown in Figure 3b. Any sufficiently small punctured neighborhood of v has $n+m$ connected components. On the other hand, puncturing the sets D^2 and D^2_+ does not violate their connectedness. Therefore, the pseudomanifold M^2 is a two-dimensional surface if and only if $m+n=1$ for any vertex v, i.e., the union of all simplices with vertex v is either of the form shown in Figure 3a or of the form shown in Figure 3b.

1.2. Reduction of Two-Dimensional Surfaces to Simplest Forms. A *triangulation* of a topological space X is a homeomorphism $X \to |K|$, where K is a simplicial complex. The complex K is also called a triangulation of X.

As a rule, triangulations are very hard to construct, because every simplex of a triangulation must have no coinciding vertices and different simplices must have different sets of vertices. For example, the partitions of the circle shown in Figures 4a and 4b are not triangulations. A simplest triangulation of the circle is shown in Figure 4c; it has three vertices.

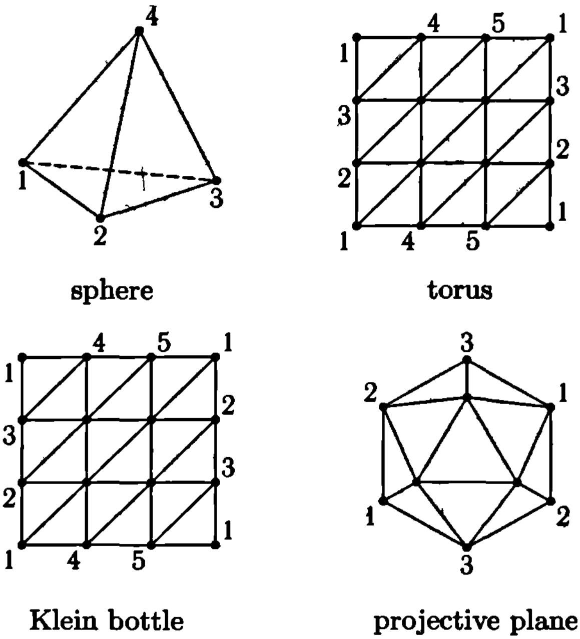

Figure 5. Triangulations of some surfaces

Figure 5 shows triangulations of the simplest two-dimensional surfaces (vertices with equal numbers are identified). We regard these triangulations as abstract simplicial complexes (although, according to Theorem 3.15 on p. 102, any two-dimensional abstract simplicial complex can be realized in the Euclidean 5-space).

Note that a triangulation of the projective plane cannot be constructed in the same way as those of the torus and the Klein bottle (such a triangulation would contain two different triangles with coinciding vertices; they are hatched in Figure 6a). The situation can easily be improved by taking the secondary diagonal in one of the corner squares, as shown in Figure 6b.

Looking at Figure 7, it is easy to see that removing a disk from the projective plane, we obtain the Möbius band.

Let T^2, K^2, and P^2 be the triangulations of the torus, Klein bottle, and projective plane shown in Figure 5, and let p, q and r be nonnegative integers. We define a two-dimensional surface $S^2 \# pT^2 \# qK^2 \# rP^2$ as follows. Consider a triangulation of the sphere S^2 so fine as to contain $p+q+r$ disjoint 2-simplices. We remove these simplices and identify p of the resulting triangles with the boundaries of p copies of $T^2 \setminus \Delta^2$, where Δ^2 is one of the simplices in T^2; then, we attach q copies of $K^2 \setminus \Delta^2$ and r copies of $P^2 \setminus \Delta^2$ in the same manner. Clearly, the two-dimensional surface obtained does

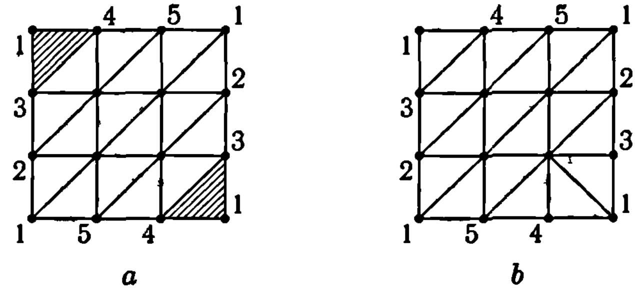

Figure 6. A triangulation of the projective plane

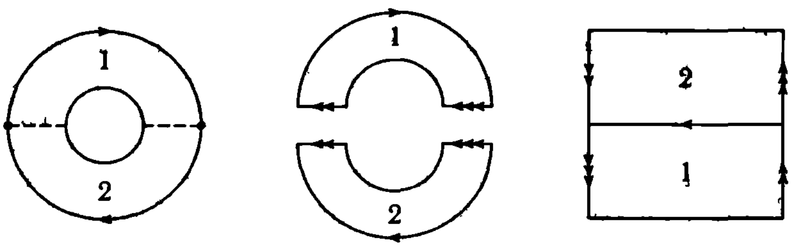

Figure 7. Möbius band

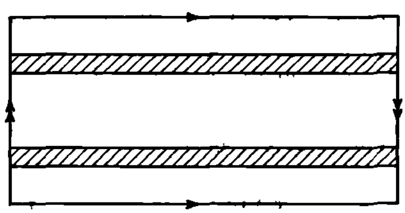

Figure 8. Klein bottle

not depend on the triangulation of the sphere S^2 (up to homeomorphism). In a similar way, we can define $M^2\#N^2$ for any two-dimensional pseudomanifolds M^2 and N^2. Clearly, $S^2\#M^2 \approx M^2$ for any two-dimensional pseudomanifold M^2.

Theorem 4.2. $S^2\#2P^2 \approx K^2$.

Proof. The surface $S^2\#2P^2$ is the cylinder $S^2 \times I$ to both ends of which Möbius bands are attached. The Klein bottle K^2 can also be represented as a cylinder with two Möbius bands attached to its boundaries; the cylinder is hatched in Figure 8. □

Theorem 4.3. $T^2\#P^2 \approx K^2\#P^2$.

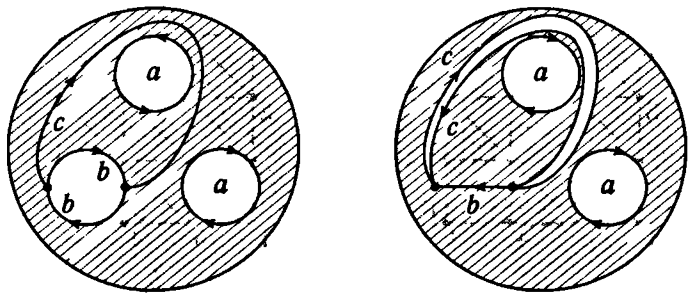

Figure 9. The surfaces $T^2\#P^2$ and $K^2\#P^2$ are homeomorphic

Proof. The surfaces $T^2\#P^2$ and $K^2\#P^2$ are shown in Figure 9. (The hatched disk represents the sphere with cuts.) The homeomorphism between them is the composition of cutting along the arrows c and gluing along b. □

It follows from Theorems 4.2 and 4.3 that the two-dimensional surface $S^2\#pT^2\#qK^2\#rP^2$ is homeomorphic to $S^2\#mT^2$ or $S^2\#nP^2$; for brevity, we denote these surfaces by mT^2 and nP^2, respectively (it is assumed that $m > 0$ and $n > 0$).

Let M^2 be a two-dimensional pseudomanifold with v vertices, e edges, and f faces. The *Euler characteristic* of M^2 is defined as $\chi(M^2) = v-e+f$.

Theorem 4.4. *Any closed two-dimensional surface M^2 is homeomorphic to mT^2 or nP^2. The numbers m and n are determined by the relations $\chi(M^2) = 2-2m$ and $\chi(M^2) = 2-n$.*

Proof (see [125]). The edges of M^2 form a graph. In this graph, we sequentially delete edges; after an edge is deleted, the faces incident to it merge into one domain, which we also call a face. We delete edges in such a way that the graph will remain connected and the number of vertices will not change. After the deletion of an edge, the number of faces either does not change or decreases by 1. At the end, we obtain a maximal tree with v vertices and $v-1$ edges (and one face). Under the deletions, the quantity

$$\text{number of vertices} - \text{number of edges} + \text{number of faces}$$

does not increase, and it equals 2 in the end; therefore, $v-e+f \leq 2$.

Suppose that there exists a closed two-dimensional surface which gives a counterexample to the theorem. Among all such surfaces we choose those for which the number $2-v+e-f \geq 0$ is minimal and among those, select the surfaces with minimal v. Finally, among the selected surfaces we choose a surface for which the minimal degree of a vertex is minimal. We denote this surface by M^2.

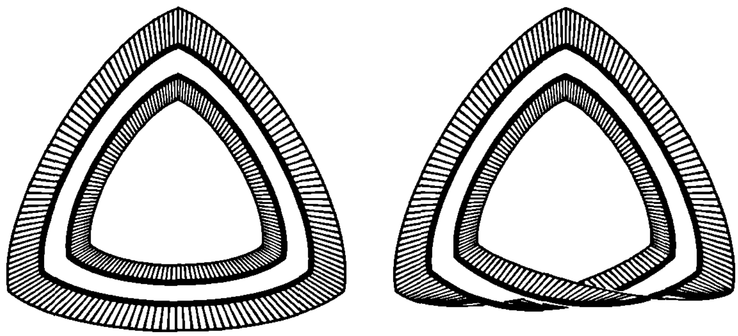

Figure 10. The boundary of the surface

Let A be a vertex of minimal degree p in M^2, and let AA_1A_2, AA_2A_3, ..., AA_pA_1 be the faces incident to it. If $p = 3$, then either M^2 is the surface of a tetrahedron or there exists a closed two-dimensional surface $\widetilde{M}^2$ obtained from M^2 by removing A and replacing the faces AA_1A_2, AA_2A_3, and AA_3A_1 by one face $A_1A_2A_3$. If M^2 is the surface of a tetrahedron, then $M^2 \approx S^2$ and $\chi(M^2) = 2$. Thus, both assumptions contradict the choice of M^2; hence $p \geq 4$.

Suppose that for some i, the vertices A_i and A_{i+2} are not joined by an edge. Then, deleting the edge AA_{i+1} and adding the edge A_iA_{i+2}, we obtain a triangulated surface for which the numbers $2 - v + e - f$ and v are the same but the minimal degree of a vertex is smaller. This contradicts the choice of the surface M^2. Therefore, for each i, the vertices A_i and A_{i+2} are joined by an edge.

The surface M^2 contains the edges AA_1, A_1A_3, and AA_3, but it cannot contain the face AA_1A_3, because $p \geq 4$. Let us cut M^2 along AA_1, A_1A_3, and AA_3. As a result, we obtain a surface bounded by a graph with six edges and six vertices in which every vertex has degree 2. This graph has no multiple edges; therefore, it either consists of two triangles or is a hexagon (see Figure 10).

To the boundary of the obtained surface we attach either two triangles or a (triangulated) hexagon. As a result, we obtain a two-dimensional surface $\widetilde{M}^2$. In the former case, we have $\tilde{f} = f + 2$, $\tilde{e} = e + 2$, and $\tilde{v} = v + 3$, whence

$$\tilde{v} - \tilde{e} + \tilde{f} = v - e + f + 2; \tag{1}$$

in the latter case, we have

$$\tilde{v} - \tilde{e} + \tilde{f} = v - e + f + 1. \tag{2}$$

In both cases, the quantity $2 - v + e - f$ decreases under the passage from M^2 to $\widetilde{M}^2$; since it is minimal for M^2, the assertion of the theorem must hold for $\widetilde{M}^2$, i.e.,

$$\widetilde{M}^2 \approx m'T^2 \quad \text{or} \quad \widetilde{M}^2 \approx n'P^2. \tag{3}$$

The surface M^2 is obtained from $\widetilde{M}^2$ by a fairly simple transformation. In the former case, $M^2 \approx \widetilde{M}^2\#T^2$ or $M^2 \approx \widetilde{M}^2\#K^2$ (either a handle or a twisted handle is attached). In the latter case, $M^2 \approx \widetilde{M}^2\#P^2$. Thus,

$$M^2 \approx mT^2 \quad \text{or} \quad M^2 \approx nP^2. \tag{4}$$

It is easy to verify that if $\chi(\widetilde{M}^2) = 2 - 2m'$ or $\chi(\widetilde{M}^2) = 2 - n'$ (m' and n' are defined by (3)), then $\chi(M^2) = 2-2m$ or $\chi(M^2) = 2-n$ (m and n are defined by (4)); it suffices to apply equalities (1) and (2) and the relations $\chi(\widetilde{M}^2\#T^2) = \chi(\widetilde{M}^2\#K^2) = \chi(\widetilde{M}^2) - 2$ and $\chi(\widetilde{M}^2\#P^2) = \chi(\widetilde{M}^2) - 1$. □

1.3. Completion of the Classification of Two-Dimensional Surfaces. To complete the classification of two-dimensional closed surfaces, it remains to prove the following assertion.

Theorem 4.5. *The surfaces S^2, mT^2 ($m = 1, 2, \dots$), and nP^2 ($n = 1, 2, \dots$) are pairwise nonhomeomorphic.*

Proof. Suppose that M_1^2 and M_2^2 are two-dimensional closed surfaces and $h: M_1^2 \to M_2^2$ is a homeomorphism. We show that $\chi(M_1^2) = \chi(M_2^2)$.

There are two graphs on M_2^2, namely, the graph G_2 formed by the edges of M_2^2 and the graph G_1, the image of the graph formed by the edges of M_1^2. Suppose that the graph G_1 has v_1 vertices, e_1 edges, and f_1 faces on M_2^2, and the graph G_2 has v_2 vertices, e_2 edges, and f_2 faces on M_2^2. We can change G_1 so that the resulting graph have the same numbers of vertices, edges, and faces and be piecewise linear, and its edges intersect the edges of G_2 transversally. Consider the graph $G = G_1 \cup G_2$. Suppose that it has v vertices, e edges, and f faces on M_2^2. We show that $v - e + f = v_2 - e_2 + f_2$. Consider any face of G_2 (i.e., a 2-simplex of the pseudomanifold M_2^2). Let v_α, e_α, and f_α be the numbers of vertices, edges, and faces of G contained in this face. According to the Euler formula, $v_\alpha - e_\alpha + f_\alpha = 1$ (the Euler formula has 2 rather than 1, but it takes into account the unbounded domain, which is absent in the case under consideration). We write v_α and e_α as $v_\alpha = v'_\alpha + v''_\alpha$ and $e_\alpha = e'_\alpha + e''_\alpha$, where v'_α and e'_α are the numbers of vertices and edges contained in the boundary of the face under consideration and v''_α and e''_α are the numbers of interior vertices and edges. Clearly, $v'_\alpha = e'_\alpha$, and hence $v''_\alpha - e''_\alpha + f_\alpha = 1$. This means that the deletion of all interior vertices, edges, and faces does not change the Euler characteristic. The equality $v'_\alpha = e'_\alpha$ implies that the deletion of all vertices belonging to the edges of G_2 does not change the Euler characteristic either. Thus, $v - e + f = v_2 - e_2 + f_2$.

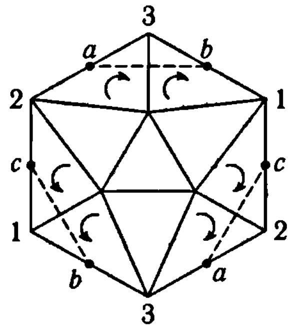

Figure 11. The orientation-reversing path on the projective plane

The graph G can be regarded as a graph on the surface M_1^2; therefore, $v-e+f = v_1-e_1+f_1$, whence $v_1-e_1+f_1 = v_2-e_2+f_2$, i.e., $\chi(M_1^2) = \chi(M_2^2)$.

So, the Euler characteristic of a two-dimensional surface does not depend on the triangulation, and to obtain the Euler characteristics for arbitrary triangulations of S^2, T^2, and P^2, it suffices to calculate the Euler characteristics for their simplest triangulations. It is easy to show that $\chi(S^2) = 2$, $\chi(T^2) = 0$, and $\chi(P^2) = 1$ for the simplest triangulations. It is also easy to verify that $\chi(M^2 \# N^2) = \chi(M^2) + \chi(N^2) - 2$. Therefore, $\chi(mT^2) = 2-2m$ and $\chi(nP^2) = 2-n$. The Euler characteristics are equal only for the surfaces mT^2 and $2mP^2$. Thus, it remains to prove that these surfaces are not homeomorphic (for $m \geq 1$).

The surfaces S^2 and T^2 with simplest triangulations are orientable pseudomanifolds; hence mT^2 with some triangulation is an orientable pseudomanifold. On the other hand, the surface P^2 with simplest triangulation is a nonorientable pseudomanifold. Indeed, there is a closed path on P^2 (namely, the path $abca$ shown by the dashed lines in Figure 11) such that after it is traversed, any compatible orientations of the 2-simplices sharing the edge 23 become incompatible. On the surface nP^2 ($n \geq 1$) with some triangulation, there is a similar closed path; it has six edges, transversally intersects edges of six simplices, and reverses orientation. This means that nP^2 ($n \geq 1$) is a nonorientable pseudomanifold.

It remains to verify that the orientability of a two-dimensional surface is invariant under homeomorphisms; i.e., the orientability of a surface can be defined without referring to triangulations.

Suppose that M^2 is a closed two-dimensional surface and $\gamma\colon I = [0,1] \to M^2$ is a path. We cover M^2 by open sets U_i homeomorphic to $\mathbb{R}^2$. Specifying an orientation at one point $x \in U_i$, we thereby specify orientations at all points of U_i; here, by an orientation we mean a direction of going around the point x.

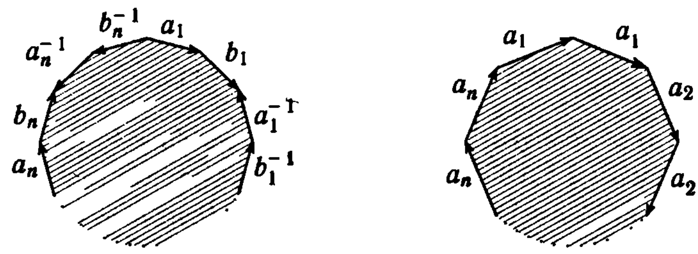

Figure 12. Constructing surfaces by gluing together polygons.

The connected components of the sets $\gamma^{-1}(U_i)$ form an open cover of the compact set I. It has a finite subcover $W_1, \dots, W_n$. We assume that $0 \in W_1$, $W_j \cap W_{j+1} \neq \varnothing$, and $1 \in W_n$. Given an orientation at a point $x \in \gamma(W_j)$, we can extend it over all points of the set $\gamma(W_j)$. Since $\gamma(W_j) \cap \gamma(W_{j+1}) \neq \varnothing$, we can also extend it over all points of $\gamma(W_{j+1})$. In this way, we can transfer orientation from $\gamma(0)$ to $\gamma(1)$ along the path γ. The result does not depend on the finite subcover of the cover of I by the connected components of the sets $\gamma^{-1}(U_i)$. Indeed, let us identify one of the domains U_i with $\mathbb{R}^2$ and consider the part of the curve γ contained in $U_i = \mathbb{R}^2$. The set $U_j \cap U_i$ is open in $\mathbb{R}^2$. The transfer of orientation along a connected component of $\gamma \cap (U_j \cap U_i)$ by means of U_j gives the same result as the transfer of orientation along γ in $\mathbb{R}^2$.[1]

We say that a two-dimensional surface M^2 is *orientable* if the transfer of orientation along any closed path is orientation-preserving, i.e., any orientation transferred along any closed path coincides with the initial orientation. Clearly, a pseudomanifold homeomorphic to a two-dimensional surface is orientable if and only if this surface is orientable. This means, in particular, that the nonorientable pseudomanifold nP^2 cannot be homeomorphic to the orientable pseudomanifold mT^2. □

Exercise 29. Prove that the surfaces nT^2 and nP^2 can be obtained from a $4n$-gon and a $2n$-gon by identifying their sides as shown in Figure 12.

Problem 39. (a) Prove that there exists a closed curve γ on the surface nP^2 such that nP^2 cut along this curve is orientable.

(b) Prove that if n is even, then a neighborhood of the curve γ is homeomorphic to the cylinder, and if n is odd, then it is homeomorphic to the Möbius band.

Problem 40. Let M_1^2 and M_2^2 be nonhomeomorphic two-dimensional surfaces with boundaries. Can the spaces $M_2^2 \times I$ and $M_1^2 \times I$ be homeomorphic?

[1] If some orientation is specified at one point of the set $U_j \approx \mathbb{R}^2$, then it is thereby specified at all points of this set.

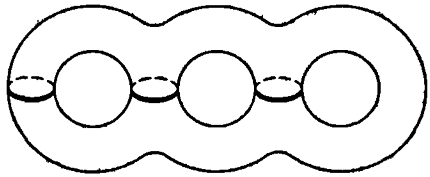

Figure 13. Curves on a two-dimensional surface

1.4. Riemann's Definition of the Surface Genus. Riemann defined the genus of a closed orientable two-dimensional surface M^2 as follows. Suppose that on the surface M^2, there exist p self-avoiding closed curves $C_1, \dots, C_p$ such that they are pairwise disjoint and the set $M^2 \setminus (C_1 \cup \dots \cup C_p)$ is connected, and let p be the maximum possible number of such curves. Then the surface M^2 has genus p.

We show that the number p thus defined indeed coincides with the genus g of the surface M^2. It is easy to give an example of curves which shows that $p \geq g$ (see Figure 13). It remains to prove that if there exist p curves $C_1, \dots, C_p$ on the surface M^2 such that the complement of their union is connected, then $p \leq g$. Let us cut M^2 along the curves $C_1, \dots, C_p$. We obtain a connected orientable surface whose boundary has $2p$ connected components. Pasting each component with a disk, we obtain a closed orientable surface $\widetilde{M}^2$ with Euler characteristic $\chi(M^2) + 2p = 2 - 2g + 2p$. But $\chi(\widetilde{M}^2) \leq 2$; therefore, $p \leq g$.

Problem 41. Prove that n pairwise disjoint Möbius bands can be arranged on the closed nonorientable surface nP^2, while $n+1$ cannot.

2. Coverings

We have considered fundamental groups and coverings in detail only for one-dimensional complexes, but we have defined them for any path-connected topological spaces. Our proofs of properties of fundamental groups and coverings do not use special features of one-dimensional complexes; the only exception is the proofs of the existence and uniqueness theorems for the lifting of a path starting at a given point and for the covering corresponding to a given subgroup of the fundamental group of the base.

For coverings of arbitrary path-connected spaces, the existence and uniqueness of the lifting of a path is not as obvious as for coverings of one-dimensional complexes, but they are fairly easy to prove. For each point of a path γ, we take the neighborhood involved in the definition of a covering. Since the closed interval is compact, we can choose a finite cover of γ by such neighborhoods. Using this finite set of neighborhoods and their preimages, we construct a lifting of γ with a given starting point. Obviously, this lifting is unique.

For coverings corresponding to a given subgroup of the fundamental group, the proof is more complicated. The construction given in the proof of Theorem 1.23 on p. 42 uses essentially the structure of a one-dimensional complex. Moreover, for general spaces, the theorem is false; it is valid only under certain assumptions. Before stating and proving the general theorem, we consider the simplest case of the universal covering of a closed orientable two-dimensional surface.

2.1. Universal Coverings of Two-Dimensional Surfaces. Recall that a covering $p\colon \widetilde{X} \to X$ is said to be *universal* if $\pi_1(\widetilde{X}) = 0$.

The torus T^2 can be obtained by identifying the points $(x+m, y+n)$ and (x, y) for all pairs of integers m, n. Therefore, the universal covering of the torus has the form $p\colon \mathbb{R}^2 \to T^2$.

The simplest construction of the universal covering of the sphere with g handles, where $g \geq 2$, uses hyperbolic geometry. Consider a regular $4g$-gon with angle $\frac{\pi}{2g}$ on the hyperbolic plane H^2. Let G be the group of motions of the hyperbolic plane generated by the translations such that each of them takes a side of the $4g$-gon under consideration to the opposite side. The images of the $4g$-gon under the action of G form a tiling of the hyperbolic plane. Therefore, the map $p\colon H^2 \to H^2/G \approx M_g^2$ is the universal covering of the sphere M_g^2 with g handles. (The facts of hyperbolic geometry used in this construction can be found in [103].)

A description of the geometric structure of the universal covering of M_g^2 which does not use hyperbolic geometry can be found in [69].

Problem 42. (a) Prove that the universal covering space of the plane $\mathbb{R}^2$ with several punctures is homeomorphic to $\mathbb{R}^2$.

(b) Let $\Delta_{ij} = \{(z_1, \dots, z_n) \in \mathbb{C}^n : z_i = z_j\}$, and let $\Sigma = \mathbb{C}^n \setminus \bigcup_{i \neq j} \Delta_{ij}$. Prove that the universal covering space of Σ is homeomorphic to $\mathbb{C}^n$.

2.2. The Existence of a Covering Space with Given Fundamental Group. Let H be a subgroup of $\pi_1(X, x_0)$. Before trying to construct a covering $p\colon \widetilde{X} \to X$ for which $p_*\pi_1(\widetilde{X}, \tilde{x}_0) = H$, we determine the properties which $\widetilde{X}$ must have. Let γ_1 and γ_2 be paths from x_0 to x in X, and let $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ be the liftings of these paths starting at $\tilde{x}_0$. The paths $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ end at the same point if and only if the class of the loop $\gamma_1\gamma_2^{-1}$ belongs to H.

This observation suggests the following construction of the space $\widetilde{X}$. Let X be a path-connected space with base point $x_0 \in X$, and let H be a subgroup of the group $\pi_1(X, x_0)$. Consider the set of all paths starting at x_0 in X. We say that paths γ_1 and γ_2 are equivalent if the class of the loop $\gamma_1\gamma_2^{-1}$

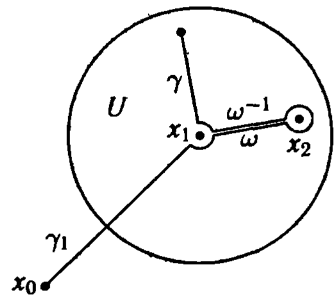

Figure 14. The path $\gamma_1\omega\omega^{-1}\gamma$

belongs to H. The points of the space $\widetilde{X}$ are classes of equivalent paths; the topology on $\widetilde{X}$ will be defined later. The projection $p\colon \widetilde{X} \to X$ takes each path γ to its end.

The map p is surjective because the space X is path-connected.

This construction does not always give the desired result. But if the space X is *locally path-connected* and *locally simply connected*, i.e., for each point $x \in X$ and any neighborhood $U \ni x$, there exists a path-connected simply connected neighborhood $V \subsetneq U$, then this construction does give the required result. In what follows, we assume that the space X is locally path-connected and locally simply connected. By a neighborhood of a point in X we mean a simply connected path-connected neighborhood. Clearly, such neighborhoods form a base for the topology of X.

The topology of the space $\widetilde{X}$. To define a topology on $\widetilde{X}$, it is sufficient to specify a base of open sets. Suppose that a point $\tilde{x} \in \widetilde{X}$ and a neighborhood $U \subset X$ are such that $p\tilde{x} \in U$. The point $\tilde{x}$ is a class of equivalent paths. Let γ be one of the paths (starting at x_0) in this class. To the pair $U, \tilde{x}$ we assign the set $(U, \tilde{x}) \subset \widetilde{X}$ consisting of the equivalence classes of all extensions of γ by paths entirely contained in U. Clearly, the set $(U, \tilde{x})$ does not depend on the choice of γ. Moreover, this set does not depend on the choice of $\tilde{x}$ in the sense that if $\tilde{x}_2 \in (U, \tilde{x}_1)$, then $(U, \tilde{x}_2) = (U, \tilde{x}_1)$. To prove this, take points $x_1 = p\tilde{x}_1$ and $x_2 = p\tilde{x}_2$ and let ω be a path between them contained in U (see Figure 14). Suppose that $\gamma_1\gamma$ is an extension of a path γ_1 between x_1 and x_0 by some path γ contained in U. Consider the path $\gamma_1\omega\omega^{-1}\gamma$, which is the extension of the path $\gamma_1\omega$ between x_0 and x_2 by the path $\omega^{-1}\gamma$ contained in U. The paths $\gamma_1\gamma$ and $\gamma_1\omega\omega^{-1}\gamma$ are homotopic; therefore, $\gamma_1\gamma \mapsto \gamma_1\omega\omega^{-1}\gamma$ is a one-to-one correspondence between $(U, \tilde{x}_1)$ and $(U, \tilde{x}_2)$.

For the base of the topology of $\widetilde{X}$ we take all sets of the form $(U, \tilde{x})$. Let us verify that any nonempty intersection of two elements of this base contains a nonempty element of the base. Suppose that $\tilde{x} \in \widetilde{U} \cap \widetilde{V}$, where

$\widetilde{U} = (U, \tilde{x}_1)$ and $\widetilde{V} = (V, \tilde{x}_2)$. We set $W = U \cap V$ and $\widetilde{W} = (W, \tilde{x})$. We have $\widetilde{W} = \widetilde{U} \cap \widetilde{V}$, and $\widetilde{W}$ belongs to the base.

The continuity of the projection p. The preimage of any (connected simply connected) neighborhood U consists of open sets from the base; therefore, it is open.

The path-connectedness of $\widetilde{X}$. Let $\tilde{x}$ be a point of $\widetilde{X}$, i.e., a class of equivalent paths. In this equivalence class, we choose an arbitrary path $\gamma(t)$ and consider the family of paths $\gamma_s(t) = \gamma(st)$, where $0 \le s \le 1$ and $0 \le t \le 1$. Each path γ_s corresponds to some point $\tilde{x}(s) \in \widetilde{X}$. Such points form a path between $\tilde{x}(0) = \tilde{x}_0$ and $\tilde{x}(1) = \tilde{x}$ in the space $\widetilde{X}$.

The projection is a local homeomorphism. Let $\bar{p}\colon (U, \tilde{x}) \to U$ be the restriction of p to the set $(U, \tilde{x})$, where U is a connected simply connected neighborhood. The path-connectedness of U implies the surjectivity of $\bar{p}$, and the simple connectedness of U implies the injectivity of p. To prove that $\bar{p}$ is continuous, consider any connected simply connected neighborhood $V \subset U$. Its preimage is $(V, \tilde{x})$.

The image of the group $\pi_i(\widetilde{X}, \tilde{x}_0)$ under the map p_* coincides with H. Let γ be a loop in X based at x_0, and let $\tilde{\gamma}$ be its lifting starting at $\tilde{x}_0$. The subgroup $p_*\pi_1(\widetilde{X}, \tilde{x}_0)$ consists of the homotopy classes of the loops γ for which the paths $\tilde{\gamma}$ are closed. By construction, $\tilde{\gamma}$ is closed if and only if the equivalence class of the path γ is $\tilde{x}_0$, i.e., the homotopy class of γ belongs to H.

2.3. The Uniqueness of a Covering Space with Given Fundamental Group. The proof of the uniqueness of a covering space with a given group $p_*\pi_1(\widetilde{X}, \tilde{x}_0) \subset \pi_1(X, x_0)$ does not use the local simple connectedness of X; it only uses the local path-connectedness of this space. The proof is based on the following lemma.

Lemma. *Suppose that $q\colon \widetilde{Y} \to Y$ is a covering, $f\colon X \to Y$ is a (continuous) map, the space X is path-connected and locally path-connected, and $f_*\pi_1(X, x_0) \subset q_*\pi_1(\widetilde{Y}, \tilde{y}_0)$. Then the map f has a unique lifting $\tilde{f}\colon X \to \widetilde{Y}$ (i.e., there exists a unique map $\tilde{f}$ such that $q\tilde{f} = f$ and $\tilde{f}(x_0) = \tilde{y}_0$).*

Proof. Consider any path γ joining the point x_0 to another point x in X. The map f takes γ to the path $f\gamma$. Consider the lifting $\tilde{\gamma}$ of $f\gamma$ starting at $\tilde{y}_0$. If the required map $\tilde{f}$ exists, then $\tilde{f}(x) = \tilde{y}$, where $\tilde{y}$ is the end of $\tilde{\gamma}$. Therefore, the map $\tilde{f}$ is unique (if it exists). We must only verify that $\tilde{y}$ does not depend on the choice of γ. In other words, we must show that if γ_1 and γ_2 are paths between x_0 and x and ω is the loop composed of γ_1 and γ_2, then the lifting of $f\omega$ starting at $\tilde{y}_0$ is a closed path in $\tilde{Y}$, i.e., the class of the loop $f\omega$ belongs to $q_*\pi_1(\tilde{Y}, \tilde{y}_0)$, which is equivalent to $f_*\pi_1(X, x_0) \subset q_*\pi_1(\tilde{Y}, \tilde{y}_0)$. This inclusion holds by assumption.

It remains to prove the continuity of the map $\tilde{f}$. At this point, we use the local path-connectedness of X. Take $x \in X$ and let $\tilde{y} = \tilde{f}(x)$. It is sufficient to prove that for any neighborhood $\widetilde{W}$ of $\tilde{y}$, there exists a neighborhood V of x such that $V \subset \tilde{f}^{-1}(\widetilde{W})$. Consider the neighborhood U of $y = q(\tilde{y})$ involved in the definition of a covering. Let $\widetilde{U}$ be the path-connected component of $q^{-1}(U) \cap \widetilde{W}$ that contains $\tilde{y}$. Since the map f is continuous, $f^{-1}(U \cap q(\widetilde{W}))$ contains some neighborhood V of x. The space X is locally path-connected; therefore, we can assume that the neighborhood V is path-connected. Then $\tilde{f}(V) \subset \widetilde{U}$, i.e., $V \subset \tilde{f}^{-1}(\widetilde{U}) \subset \tilde{f}^{-1}(\widetilde{W})$. Indeed, any point $x_1 \in V$ can be joined to x by a path γ contained in V. The image $f\gamma$ of γ is contained in $U \cap q(\widetilde{W})$; hence the lifting of $f\gamma$ is entirely contained in $\widetilde{U}$. This means that $\tilde{f}(y) = \tilde{y} \in \widetilde{U}$. □

Using this lemma, we can easily prove the uniqueness of a covering space with a given fundamental group. Indeed, let $p_i \colon \widetilde{X}_i \to X$ $(i = 1, 2)$ be coverings of a path-connected locally path-connected space X; suppose that $(p_1)_*\pi_1(\widetilde{X}_1, \tilde{x}_1) = (p_2)_*\pi_1(\widetilde{X}_2, \tilde{x}_2)$. Then there exists a homeomorphism $h \colon \widetilde{X}_1 \to \widetilde{X}_2$ such that $p_2 h = p_1$ and $h(\tilde{x}_1) = \tilde{x}_2$. The map h is defined as the lifting of p_1, and h^{-1} is the lifting of p_2.

It is seen from the proof of the lemma that there exists a unique lifting $\tilde{f}$ for any path-connected space X. But if X is not locally path-connected, then the map $\tilde{f}$ constructed in the proof of the lemma may not be continuous. This is shown by the following example.

Example (Zeeman). Consider the topological space $X \subset \mathbb{R}^2$ consisting of a circle, an arc AB, and an infinite set of intervals $I_1, I_2, \ldots$; one endpoint of each interval is A, and the other endpoints converge to B (see Figure 15). Let $\widetilde{X}_1$ and $\widetilde{X}_2$ be the topological spaces shown in the same figure. Coverings $p_i \colon \widetilde{X}_i \to X$ twice wind the circles of the respective spaces $\widetilde{X}_i$ around the circle of X, are isometries on the intervals, and homeomorphically map each arc to the arc of X. We have

$$(p_1)_*\pi_1(\widetilde{X}_1) = (p_2)_*\pi_1(\widetilde{X}_2) = 2\mathbb{Z} \subset \mathbb{Z} = \pi_1(X),$$

although the map p_1 has no continuous lifting h for which $p_2 h = p_1$.

Indeed, such a "lifting" h exists and is unique (provided that the image of one point under h is given), but it is discontinuous at the points P and Q, which belong to $p_1^{-1}(B)$.

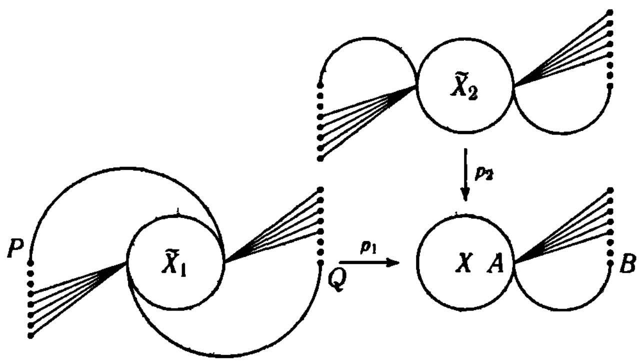

Figure 15. The Zeeman example

Problem 43. Prove that an mn-fold covering $p\colon \widetilde{X} \to X$ can be represented as a composition

$$\widetilde{X} \xrightarrow{p_1} Y \xrightarrow{p_2} X,$$

where p_1 is an m-fold covering and p_2 is an n-fold covering, if and only if the preimage $p^{-1}(x)$ of some point $x \in X$ can be partitioned into m-point sets $I_1, \dots, I_n$ so that all liftings of any closed path in X which start in the same set I_i end in the same set I_j.

The proof of the following assertion is based on the covering technique and uses the Borsuk Ulam theorem (although the Borsuk Ulam theorem can easily be derived from it).

Theorem 4.6. *If $m > n \geq 1$, then there exists no map $g\colon \mathbb{R}P^m \to \mathbb{R}P^n$ inducing an isomorphism of fundamental groups.*

Proof. Suppose that $p_m\colon S^m \to \mathbb{R}P^m$ and $p_n\colon S^n \to \mathbb{R}P^n$ are double coverings. Let us construct a map $f\colon S^m \to S^n$ for which $gp_m = p_n f$. Take a point $x_0 \in S^m$ and choose a point y_0 in the two-point set $p_n^{-1}g(x_0)$. We join x_0 to a point $x \in S^m$ by a path γ, consider the lifting of $gp_m\gamma$ starting at y_0, and define $g(x) = y$, where y is the end of this lifting. The map g is well defined, because any loop in S^m with $m \geq 2$ is contractible.

The equality $gp_m = p_n f$ implies $f(-x) = \pm f(x)$. The sign is selected as follows. Let α_m and α_n be the generators of the groups $\pi_1(\mathbb{R}P^m)$ and $\pi_1(\mathbb{R}P^n)$ (recall that $\pi_1(\mathbb{R}P^1) = \mathbb{Z}$ and $\pi_1(\mathbb{R}P^m) = \mathbb{Z}_2$ for $m \geq 2$). Suppose that $g_*\alpha_m = k\alpha_n$ (here $k \in \mathbb{Z}_2$ if $n \geq 2$ and $k \in \mathbb{Z}$ if $n = 1$). Then $f(-x) = (-1)^k f(x)$, because $k\alpha_n$ is the image of the arc joining x and $-x$ under the map gp_m.

Thus, if there exists a map $g\colon \mathbb{R}P^m \to \mathbb{R}P^n$ for which $g_*\alpha_m = \pm\alpha_n$, then there exists a map $f\colon S^m \to S^n$ for which $f(-x) = -f(x)$. $\square$

Note that if there exists a map $f\colon S^m \to S^n$ for which $f(-x) = -f(x)$, then, for the map $g\colon \mathbb{R}P^m \to \mathbb{R}P^n$ defined by $\{x, -x\} \mapsto \{f(x), -f(x)\}$, we have $g_*\alpha_m = k\alpha_n$, where k is odd. If $n \geq 2$, then $\alpha_n \in \mathbb{Z}_2$, and g_* is an isomorphism. If $n = 1$, then g_* is a nonzero homomorphism $\mathbb{Z}_2 \to \mathbb{Z}$, which does not exist. This argument gives a new proof of the Borsuk Ulam theorem for maps $S^2 \to \mathbb{R}^2$.

2.4. Local Homeomorphisms. A map $f\colon X \to Y$ is called a *local homeomorphism* if each point $x \in X$ has a neighborhood U such that the set $f(U)$ is open in Y and the restriction of f to U is a homeomorphism.

Any covering is a local homeomorphism. Under certain conditions, the converse is also true. We consider only the case of finite coverings. We say that a map $f\colon X \to Y$ is *proper* if the preimage of any compact set under this map is compact.

Theorem 4.7 (see [55]). *Suppose that X and Y are Hausdorff spaces and Y is path-connected. Then any proper surjective local homeomorphism $f\colon X \to Y$ is a finite covering.*

Proof. Take a point $y \in Y$. The set $f^{-1}(y)$ is discrete, because f is a local homeomorphism, and it is finite, because f is proper. Let $f^{-1}(y) = \{x_1, \dots, x_n\}$. The points x_i and x_j $(i \neq j)$ have disjoint neighborhoods $U_{ij} \ni x_i$ and $U_{ji} \ni x_j$ in the Hausdorff space X. Let $U_i = \bigcap_{j\neq i} U_{ij}$. Then $x_i \in U_i$ and $U_i \cap U_j = \varnothing$ for $i \neq j$.

For each point x_i, we choose a neighborhood V_i such that the set $f(V_i)$ is open in Y and the restriction of f to V_i is a homeomorphism. The neighborhoods $W_i = U_i \cap V_i$ are pairwise disjoint, and f homeomorphically maps W_i onto a neighborhood of y.

We set $W = \bigcap_{i=1}^n f(W_i)$. To prove that f is a covering, it suffices to show that the preimage of W is contained in $\bigcup_{i=1}^n f(W_i)$, and to show this, it suffices to prove that the number n does not depend on the point y, i.e., the preimages of all points of Y contain the same number of points. We use the path-connectedness of Y. Take points $y_1, y_2 \in Y$ and consider a continuous path $\gamma\colon [0,1] \to Y$ joining $\gamma(0) = y_1$ and $\gamma(1) = y_2$. We prove that the restriction of f to the preimage of γ is a covering. Consider the compact space $\tilde{Y} = \gamma([0,1]) \subset Y$. Let $\tilde{X} = f^{-1}(\tilde{Y})$, and let $\tilde{f}$ be the restriction of f to $\tilde{X}$. The spaces $\tilde{Y}$ and $\tilde{X}$ are Hausdorff, and $\tilde{f}$ is a proper surjective local homeomorphism. Thus, for any point $\tilde{y} \in \tilde{Y}$, we can construct open sets $\widetilde{W}_i$ in the same way as the open sets W_i for $y \in Y$. We put

$$\widetilde{W} = \left(\bigcap_{i=1}^n \tilde{f}(\widetilde{W}_i)\right) \setminus \tilde{f}\left(\tilde{X} \setminus \bigcup_{i=1}^n \widetilde{W}_i\right).$$

To prove that $\tilde{f}$ is a covering, it is sufficient to verify that $\widetilde{W}$ is an (open) neighborhood of $\tilde{y}$ and $\tilde{f}^{-1}\widetilde{W} \subset \bigcup_{i=1}^{n} \widetilde{W}_i$.

By construction, $\tilde{y} \in \bigcap_{i=1}^{n} \tilde{f}(\widetilde{W}_i)$ and $\tilde{f}^{-1}(\tilde{y}) \subset \bigcup_{i=1}^{n} \widetilde{W}_i$; therefore, $\tilde{y} \in \widetilde{W}$.

If $\tilde{f}(\tilde{x}) \in \widetilde{W}$, then $\tilde{x} \notin \widetilde{X} \setminus \bigcup_{i=1}^{n} \widetilde{W}_i$, i.e., $\tilde{x} \in \bigcup_{i=1}^{n} \widetilde{W}_i$.

It remains to show that the set $\widetilde{W}$ is open, i.e., $\tilde{f}\big(\widetilde{X} \setminus \bigcup_{i=1}^{n} \widetilde{W}_i\big)$ is closed. The space $\tilde{Y}$ is compact, and the map $\tilde{f}$ is proper; hence $\widetilde{X} = \tilde{f}^{-1}(\tilde{Y})$ is compact. Thus, the set $\widetilde{X} \setminus \bigcup_{i=1}^{n} \widetilde{W}_i$ is compact as a closed subset of a compact space. The set $\tilde{f}\big(\widetilde{X} \setminus \bigcup_{i=1}^{n} \widetilde{W}_i\big)$ is a compact subset of the Hausdorff space $\tilde{Y}$; hence it is closed. □

Theorem 4.7 gives a criterion for a local homeomorphism to be a global homeomorphism.

Theorem 4.8 (see [56]). *Let X and Y be path-connected Hausdorff spaces. A local homeomorphism $f: X \to Y$ is a (global) homeomorphism if and only if the map f is proper and the homomorphism $f_*: \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$ is an epimorphism for some point $x_0 \in X$.*

Proof. The "only if" part of the theorem is obvious. We prove the "if" part. It suffices to verify that the map f is one-to-one.

Step 1. The map f is surjective.

Take arbitrary points $y_0 \in f(X)$ and $y_1 \in Y$ and consider a path $\alpha: I = [0,1] \to Y$ joining y_0 and y_1. Since the map f is proper, the set $f^{-1}(\alpha(I))$ is compact; hence $f(X) \cap \alpha(I) = f\left(f^{-1}(\alpha(I))\right)$ is closed in $\alpha(I)$. On the other hand, the set $f(X)$ is open in Y, because f is a local homeomorphism, and therefore $f(X) \cap \alpha(I)$ is open in $\alpha(I)$. Thus, $f(X) \cap \alpha(I) = \alpha(I)$. In particular, $y_1 \in f(X)$.

Step 2. The map f injective.

According to Theorem 4.7, f is a covering. Therefore, $f_*: \pi_1(X, x_0) \to \pi_1(Y, f(x_0))$ is a monomorphism, and the number of elements in each fiber is equal to the index of the subgroup $f_*\pi_1(X, x_0)$ in the group $\pi_1(Y, f(x_0))$. By assumption, f_* is an epimorphism. Hence the covering f is one-fold, i.e., f is a homeomorphism. □

Corollary. *Suppose that X and Y are path-connected Hausdorff spaces, $\pi_1(Y) = 0$, and $f: X \to Y$ is a local homeomorphism. Then f is a (global) homeomorphism if and only if it is a proper map.*

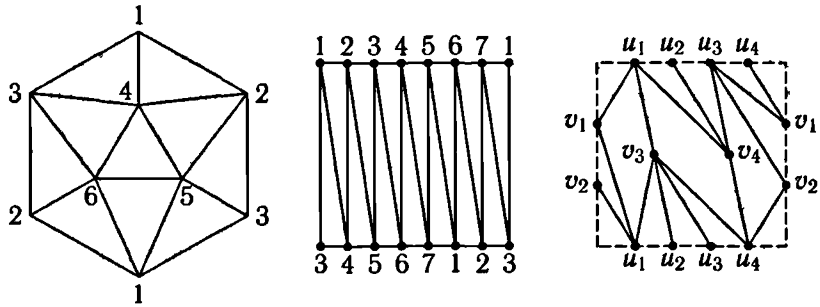

Figure 16. The graphs K_6, K_7, and $K_{4,4}$

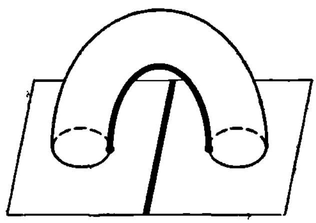

Figure 17. Eliminating the intersections of edges

3. Graphs on Surfaces and Deleted Products of Graphs

3.1. The Genus of a Graph. The graphs $K_{3,3}$ and K_5 cannot be embedded in the plane (see Theorem 1.2 on p. 7). It is easy to understand that a graph can be embedded in the plane $\mathbb{R}^2$ if and only if it can be embedded in the sphere S^2. We might consider embeddings of graphs in other surfaces. For example, the graph K_6 can be drawn in the projective plane P^2, and the graphs K_7 and $K_{4,4}$ can be drawn on the torus (see Figure 16).

Theorem 4.9 (König [**71**]). (a) *Any finite graph G can be embedded in a closed orientable two-dimensional surface M^2.*

(b) *If the graph G is connected and the surface M^2 has minimal genus, then each of the domains into which G divides M^2 is homeomorphic to the disk.*

Proof. (a) Allowing intersections of edges, we can draw any graph on the sphere. The intersection of two edges can be eliminated by attaching a handle to the sphere and letting one edge remain on the sphere and the other go along the handle (see Figure 17).

(b) It is sufficient to consider the case where G has no multiple edges. Let $U_1, \ldots, U_m$ be the domains into which G divides M^2. By assumption, the graph G is connected and has no multiple edges; therefore, the boundary of each domain U_i is homeomorphic to the circle. A domain U_i is contractible if and only if, attaching the disk D^2 to U_i (along the boundary), we obtain the sphere S^2. Suppose that one of the domains U_i is not contractible. If we remove U_i from M^2 and attach D^2 in its place, then the graph G will turn out to lie on the surface $\widetilde{M}^2$ of genus strictly less than that of M^2. This contradicts the minimality of the genus of M^2. □

Remark 1. If a graph can be embedded in an orientable surface M^2, then it can be embedded in the nonorientable surface $M^2 \# P^2$.

Remark 2. If a graph G divides a surface M^2 into contractible domains, then the genus of M^2 is not necessarily minimal. For example, the wedge of two circles admits such an embedding in both the sphere and the torus.

The minimal genus of an orientable surface in which a graph G can be embedded is called the *genus* of this graph. The genus of a graph G is denoted by $g(G)$.

Theorem 4.10. *Let G be a connected graph without loops and multiple edges containing v vertices and e edges. Then*

$$g(G) \geq \frac{e}{6} - \frac{v}{2} + 1;$$

moreover, if G contains no cycles of length 3, then

$$g(G) \geq \frac{e}{4} - \frac{v}{2} + 1.$$

Proof. Suppose that G is embedded in an orientable surface M^2. We can assume that all domains into which G divides M^2 are contractible. In this case, we have $2 - 2g(M^2) = v - e + f$, where f the number of domains into which G divides M^2. If the boundary of each domain consists of at least n edges, then $nf \leq 2e$; therefore,

$$g(G) \geq g(M^2) \geq \frac{1}{2}\left(2 - v + e\left(1 - \frac{2}{n}\right)\right) = \frac{n-2}{2n}e - \frac{v}{2} + 1.$$

By assumption, the graph G has no loops and multiple edges. This means that $n \geq 3$, i.e., $(n-2)/2n \geq 1/6$. If, in addition, G has no cycles of length 3, then $n \geq 4$, i.e., $(n-2)/2n \geq 1/4$. □

Example. $g(K_n) \geq \frac{(n-3)(n-4)}{12}$.

Proof. The graph K_n has $\frac{n(n-1)}{2}$ edges; therefore,

$$g(K_n) \geq \frac{n(n-1)}{12} - \frac{n}{2} + 1 = \frac{n^2 - 7n + 12}{12} = \frac{(n-3)(n-4)}{12}. \qquad \square$$

Example. $g(K_{m,n}) \geq \frac{(m-2)(n-2)}{4}$.

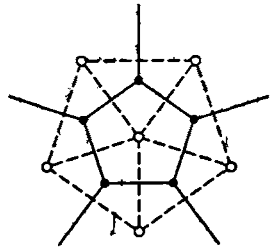

Figure 18. The dual graph

Proof. The graph $K_{m,n}$ has $m+n$ vertices and mn edges. Moreover, $K_{m,n}$ has no cycles of length 3. Hence

$$g(K_{m,n}) \geq \frac{mn}{4} - \frac{m+n}{2} + 1 = \frac{(m-2)(n-2)}{4}. \qquad \square$$

The above bounds for the genera of the graphs K_n and $K_{m,n}$ cannot be improved; that is, each graph K_n ($K_{m,n}$) can be embedded in an orientable surface of genus equal to the least integer that is larger than or equal to $\frac{(n-3)(n-4)}{12}$ (respectively, $\frac{(m-2)(n-2)}{4}$). Examples of such embeddings are fairly complicated, especially for the graphs K_n. The first examples for K_n were constructed by Ringel, Youngs, and Mayer in [**83**, **107**, **108**]. Examples of embeddings of the graphs $K_{m,n}$ were constructed in [**106**]. A modern exposition of these constructions is given in [**45**].

3.2. Map Coloring. We say that a map on a surface M^2 is n-colorable if the vertex of any (loopless) graph embedded in M^2 is n-colorable, i.e., can be colored with n colors in such a way that any two vertices joined by an edge have different colors. Visually, the problem of coloring the dual graph, i.e., the graph whose vertices correspond to domains on the surface M^2 and edges join vertices corresponding to domains with common edges (see Figure 18), is more obvious. For the dual graph, the coloring must be such that neighboring domains have different colors.

Theorem 4.11 (Heawood [**52**]). *Any map on a closed orientable surface of genus $g > 0$ is $\left[\frac{7+\sqrt{1+48g}}{2}\right]$-colorable.*

Proof. If e is an edge of a graph G, then any proper coloring of the vertices of G is also a proper coloring of the graph $G - e$. Therefore, adding an edge cannot decrease the number of colors required for coloring the vertices of the graph. Thus, drawing additional edges, we can assume that the graph G divides the surface M_g^2 into contractible triangular domains. In this case, we have $2e(G) = 3f(G)$, and the equality $v(G) - e(G) + f(G) = 2 - 2g$

implies $e(G) = 3v(G) + 6g - 6$. Clearly, the sum of degrees of the vertices of G is equal to $2e(G)$; therefore, the degree of one of the vertices is at most

$$\frac{2e(G)}{v(G)} = 6 + \frac{12(g-1)}{v(G)}. \tag{1}$$

Suppose that n is a number such that any graph on a surface of genus g has a vertex of degree at most $n-1$ (an example of such a number is $n = 7 + 12(g-1)$). It is easy to prove by induction on the number of vertices that any graph on a surface of genus g is n-colorable. Indeed, if the graph G from which a vertex v of degree at most $n-1$ is deleted is n-colorable, then the graph G itself is n-colorable (there remains at least one color for the vertex v).

Let $n(g)$ be the minimum number of colors sufficient for coloring an arbitrary map on a surface of genus g (the number $n(g)$ is finite because any map on a surface of genus g is $(7 + 12(g-1))$-colorable). Consider a graph G for which this minimum number $n(g)$ is realized and formula (1) holds. Clearly, $v(G) \geq n(g)$; therefore, if $g \geq 1$, then

$$n(g) \leq 7 + \frac{12(g-1)}{v(G)} \leq 7 + \frac{12(g-1)}{n(g)}.$$

(Note that this inequality does not hold for the sphere.) Solving the inequality $n(g)^2 - 7n(g) \leq 12g - 12$ and taking into account the positivity of $n(g)$, we obtain the required result. □

The inequality $n(g)^2 - 7n(g) \leq 12g - 12$ can be rewritten as

$$g \geq \frac{(n(g)-3)(n(g)-4)}{12}.$$

This inequality is closely related to the inequality

$$g(K_n) \geq \frac{(n-3)(n-4)}{12}.$$

Indeed, if the graph K_n is embedded in a surface of genus g, then $n(g) \geq n$, because any (proper) coloring of K_n requires n colors.

The examples of embeddings of K_n in orientable surfaces constructed by Ringel and other authors show that

$$n(g) \geq \left[\frac{7 + \sqrt{1 + 48g}}{2}\right].$$

Combining these inequalities with the inequalities of Heawood, we obtain

$$n(g) = \left[\frac{7 + \sqrt{1 + 48g}}{2}\right].$$

Recall that Heawood's argument does not apply in the case of $g = 0$.

3.3. Deleted Products of Graphs. The *deleted product* of a simplicial complex K is the subspace of $|K \times K|$ consisting of all products $\Delta^i_\alpha \times \Delta^j_\beta$, where Δ^i_α and Δ^j_β are disjoint simplices in K. The deleted product of K has the natural structure of a CW-complex. It has $n^2 - n$ vertices, where n is the number of vertices of K.

A graph G without loops and multiple edges can be viewed as a one-dimensional simplicial complex with the same vertex set. Thus, for graphs, deleted products are defined also. If a graph has a pair of disjoint edges, then its deleted product is a two-dimensional CW-complex. In what follows, we assume that the graphs G under considerations have pairs of disjoint edges.

Theorem 4.12. (a) *The deleted product of a graph G is a closed two-dimensional pseudomanifold (not necessarily connected) if and only if the graph G from which any pair of vertices v_i and v_j joined by an edge is deleted, is a set of cycles, i.e., any vertex $v_k \notin \{v_i, v_j\}$ is incident to precisely two edges not going to v_i or v_j.*

(b) *If the graph G is 2-connected (i.e., it remains connected after the deletion of any vertex), then its deleted product is connected.*

Proof. (a) Suppose that a vertex $v_k \notin \{v_i, v_j\}$ is incident to edges $v_k v_{p_1}, \dots, v_k v_{p_s}$, where $v_{p_1}, \dots, v_{p_s} \notin \{v_i, v_j\}$. Then, in the deleted product of G, the edge $v_k \times v_i v_j$ is incident to the faces $v_k v_{p_\alpha} \times v_i v_j$, where $\alpha = 1, \dots, s$. Therefore, the deleted product is a closed pseudomanifold if and only if $s = 2$ for all triples of vertices $\{v_i, v_j, v_k\}$.

(b) Vertices $v_i \times v_j$ and $v_p \times v_q$ can be joined by an edge as follows. If $p \neq j$, we join $v_i \times v_j$ with $v_p \times v_j$, and then, $v_p \times v_j$ with $v_p \times v_q$. To join vertices $v_i \times v_j$ and $v_j \times v_i$, we choose a vertex $v_k \notin \{v_i, v_j\}$ and join first $v_i \times v_j$ with $v_k \times v_j$, then $v_k \times v_j$ with $v_k \times v_i$, and finally $v_k \times v_i$ with $v_j \times v_i$. □

The conditions of Theorem 4.12 hold for the graphs K_5 and $K_{3,3}$. Therefore, the deleted products of these graphs are connected closed surfaces.

Problem 44 ([119]). Prove that the deleted product of the graph $K_{3,3}$ is a sphere with four handles, and the deleted product of the graph K_5 is a sphere with six handles.

Problem 45. Prove that the deleted product of a graph cannot be a sphere with an odd number of Möbius bands.

4. Fibrations and Homotopy Groups

The notion of a locally trivial bundle with fiber F is a generalization of the notion of a covering. For a covering, the fiber is discrete, and for a locally trivial bundle, it can be an arbitrary topological space. To discuss properties of bundles, we need the notion of a homotopy group, which is a generalization of that of a fundamental group.

4.1. Covering Homotopies. A *locally trivial fiber bundle* is a quadruple (E, B, F, p), where E, B, and F are topological spaces and $p\colon E \to B$ is a map with the following properties:

- each point $x \in B$ has a neighborhood U for which $p^{-1}(U) \approx U \times F$;
- the homeomorphism $U \times F \to p^{-1}(U)$ is compatible with p, i.e., the diagram

$$\begin{tikzcd} U \times F \arrow[r] \arrow[dr] & p^{-1}(U) \arrow[d, "p"] \\ & U \end{tikzcd}$$

  is commutative (here $U \times F \to U$ is the projection onto the first factor).

The spaces E, B, and F are called, respectively, the *total space*, *base*, and *fiber* of the bundle, and p is its *projection* map. The map $p\colon E \to B$ is also called a *fibration*.

Example. The covering $p\colon \widetilde{X} \to X$ is a locally trivial fibration with fiber $F = p^{-1}(x)$, where $x \in X$.

Example. The natural projection $p\colon B \times F \to B$ is a locally trivial fibration. This fibration is called *trivial*.

Fibrations $p_1\colon E_1 \to B$ and $p_2\colon E_2 \to B$ are *equivalent* if there exists a homeomorphism $h\colon E_1 \to E_2$ such that $p_1 = p_2 h$. A fibration equivalent to a trivial fibration is called *trivial* as well.

Theorem 4.13 (Feldbau). *Any locally trivial fibration over the cube I^k is trivial.*

Proof. First, let us show that if the cube $I^k = I^{k-1} \times I$ is partitioned into two semicubes $I^k_- = I^{k-1} \times \left[0, \frac{1}{2}\right]$ and $I^k_+ = I^{k-1} \times \left[\frac{1}{2}, 1\right]$ and a bundle is trivial over each of them, then this bundle is trivial over the entire cube I^k. In other words, we must prove that any homeomorphisms $h_\pm\colon p^{-1}(I^k_\pm) \to F \times I^k_\pm$ compatible with the projection determine a homeomorphism $h\colon p^{-1}(I^k) \to F \times I^k$ which is also compatible with the projection. Compatibility with the projection means that if $y \in p^{-1}(x)$, then $h(y) = (f, x)$ for $f \in F$, i.e., the homeomorphism h is a family of homeomorphisms $\varphi_x\colon p^{-1}(x) \to F$, where $x \in I$. On the set $p^{-1}(I^k_-)$, we can define h to be equal to h_-; i.e., we can assume that $\varphi_x = \varphi_{x,-}$ for

$x \in I^k_-$. For each point $a \in I^k_- \cap I^k_+ = I^{k-1} \times \{\frac{1}{2}\}$, we have two homeomorphisms $p^{-1}(x) \to F$, namely, $\varphi_{a,+}$ and $\varphi_{a,-}$. Consider the homeomorphism $\psi_a = \varphi_{a,-}(\varphi_{a,+})^{-1} \colon F \to F$. We use it to define homeomorphisms φ_x for $x \in I^k_+$ as follows. Let $a(x)$ be the orthogonal projection of $x \in I^k_+$ on the partition $I^{k-1} \times \{\frac{1}{2}\}$. We set $\varphi_x = \psi_{a(x)}\varphi_{x,+}$. This definition is consistent with the definition of $\varphi_{a,+}$ and $\varphi_{a,-}$, because, at the point $x = a \in I^{k-1} \times \{\frac{1}{2}\}$, we have $\psi_a\varphi_{a,+} = \varphi_{a,-}(\varphi_{a,+})^{-1}\varphi_{a,+} = \varphi_{a,-}$. Thus, the family of homeomorphisms φ_x continuously depends on x and, hence, determines a homeomorphism $h \colon p^{-1}(I^k) \to F \times I^k$ compatible with the projection.

Now, the required assertion is easy to prove by contradiction. Indeed, suppose that there is a nontrivial locally trivial bundle over the cube I^k. Cut the cube I^k into two semicubes. According to what we have proved, the bundle is nontrivial over one of the semicubes. We cut this semicube, and so on. As a result, we obtain a sequence of parallelepipeds such that their diameters tend to zero and the bundle is nontrivial over each of them. This sequence of parallelepipeds converges to some point x_0. By definition, the point x_0 has a neighborhood over which the bundle is trivial. One of the parallelepipeds is entirely contained in this neighborhood, so the bundle must be trivial over this parallelepiped. This contradiction completes the proof. □

Let $p \colon E \to B$ be a locally trivial fibration, and let $f \colon X \to B$ be a map. We say that a map $\tilde{f} \colon X \to E$ *covers* f, or is a *lifting* of f, if $p\tilde{f} = f$.

Fibrations have a property similar to the existence of a lifting of a path for coverings. The main difference is that for coverings, the lifting of a path with a given starting point is unique, while for fibrations, only the existence theorem is valid.

Theorem 4.14 (covering homotopy). *Suppose that $p \colon E \to B$ is a locally trivial fibration, X is a CW-complex, $X' \subset X$ is its subcomplex, and the following maps are given:*

- *a map $\tilde{h} \colon X \to E$;*
- *a homotopy $H \colon X \times I \to B$ of the map $h = p\tilde{h}$;*
- *a homotopy $\widetilde{H}' \colon X' \times I \to E$ covering the restriction of the homotopy H to $X' \times I$ and extending the restriction of $\tilde{h} \colon X \times \{0\} \to E$ to $X' \times \{0\}$.*

Then there exists a homotopy $\widetilde{H} \colon X \times I \to E$ which covers the homotopy H and is a simultaneous extension for the homotopy $\widetilde{H}'$ and the map $\tilde{h} \colon X \times \{0\} \to E$.

Proof. First, suppose that the fibration is trivial, i.e., $E = B \times F$ and $p(b, f) = b$. In this case, the map $\widetilde{H}$ is defined as the product of two maps,

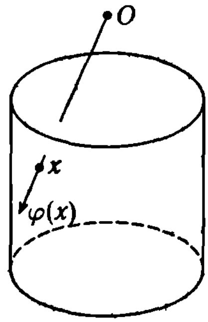

Figure 19. The projection of the cylinder

to B and to F. The map to B coincides with H; it remains to define the map to F. This is done with the help of the following lemma, which is also useful in many other situations.

Lemma (Borsuk). *Suppose that X is a CW-complex, $X' \subset X$ is its subcomplex, and $f\colon X \to Y$ is a map. Then any homotopy $F'\colon X' \times I \to Y$ of the map $f' = f \restriction X'$ can be extended to a homotopy of the map f.*

Proof. We extend the homotopy by induction. Suppose that $x_0 \in X^0$, i.e., x_0 is a vertex. If $x_0 \in X'$, then the map $\{x_0\} \times I \to Y$ is defined, and if $x_0 \notin X'$, then we send $\{x_0\} \times I$ to the point $f(x_0)$. Now, suppose that the homotopy is extended over the skeleton X^n, where $n \geq 0$. Then, for each $(n+1)$-cell, we have a map defined on $S^n \times I$ and on $D^{n+1} \times \{0\}$, and we must extend it over $D^{n+1} \times I$. To this end, we embed the cylinder $D^{n+1} \times I$ in $\mathbb{R}^{n+2}$ and choose a point O on the axis of the cylinder above its top (see Figure 19). Let $x \mapsto \varphi(x)$ be the projection of the cylinder from O onto the union of the lateral surface and the lower base. The map is defined at the point $\varphi(x)$; we send x to the image of $\varphi(x)$. □

Now, suppose that $X = D^n$ and $p\colon E \to B$ is an arbitrary (locally trivial) fibration. By assumption, a map $H\colon D^n \times I \to B$ is given. Using this map, we can construct the *induced fibration* $p_1\colon E_1 \to Y = D^n \times I$, where

$$E_1 = \{(e, y) \in E \times Y : p(e) = H(y)\}$$

and $p_1(x, y) = p(x)$. It is easy to verify that the induced fibration is locally trivial. Moreover, if $\widetilde{H}'$ is a homotopy on a subcomplex $X' \subset D^n$ which covers H, then it determines a homotopy $\widetilde{H}_1'\colon X' \times I \to E_1$ by the formula $\widetilde{H}_1'(y) = (\widetilde{H}'(y), y)$; we have $p\widetilde{H}'(y) = H(y)$ because $\widetilde{H}'$ covers H. In a similar way, from the map $\tilde{h}\colon X \to E$, a map $\tilde{h}_1\colon X \to E_1$ is constructed.

The base of the fibration p_1 is homeomorphic to D^{n+1}. According to the theorem of Feldbau, the fibration over D^{n+1} is trivial. Therefore, as in the case considered above, the maps $\widetilde{H}_1'$ and $\tilde{h}_1$ from Y to $E_1 = Y \times F$ are determined componentwise by two maps, to Y and to F; moreover, the map to Y is the identity, i.e., $\widetilde{H}_1'$ and $\tilde{h}_1$ are determined by maps to F. Thus, we have the map $X \to F$ corresponding to $\tilde{h}_1$ and a homotopy of its restriction to X'. Extending this homotopy to the entire space X, we obtain a map $\widetilde{H}_1\colon D^n \times I \to E_1$. The required covering homotopy $\widetilde{H}$ is the composition of $\widetilde{H}_1$ and the natural projection $E_1 \to E$.

Finally, consider the last case, in which the fibration $p\colon E \to B$ and the pair (X, X') are arbitrary. We construct the homotopy by induction. On the 0-skeleton, the homotopy is given at some points; at the other points, we define it to be constant. In passing from the $(n-1)$-skeleton to the n-skeleton, we must extend a homotopy given on ∂D^n to D^n. We have already learned how to do this. □

Problem 46. (a) Prove that if $Y \subset X$ is a contractible subcomplex, then $X/Y \sim X$.

(b) Let Y be a subcomplex of a CW-complex X. Prove that $X/Y \sim X \cup CY$.

Problem 47. Prove that an n-connected CW-complex is homotopy equivalent to a CW-complex having precisely one vertex and no k-cells for $1 \le k \le n$.

Problem 48. Prove that if a CW-complex X with one vertex has no k-cells for $1 \le k \le n$, then it is n-connected.

Problem 49. Given connected CW-complexes A and B with base points a_0 and b_0, prove that $A * B \sim \Sigma(A \wedge B)$, where $A \wedge B = A \times B / A \vee B$ and $A \vee B = (\{a_0\} \times B) \cup (A \times \{b_0\})$.

Problem 50. Given an n-connected CW-complex X and an m-connected CW-complex Y (both of them finite-dimensional), prove that

(a) ΣX is an $(n+1)$-connected complex;

(b) $X \wedge Y$ is an $(n+m+1)$-connected complex;

(c) $X * Y$ is an $(n+m+2)$-connected complex.

Problem 51. (a) Prove that $\Sigma(S^1 \times S^1) \sim S^2 \vee S^2 \vee S^3$.

(b) Prove that if X and Y are CW-complexes, then $\Sigma(X \times Y) \sim \Sigma X \vee \Sigma Y \vee \Sigma(X \wedge Y)$.

Problem 52. Prove that a locally trivial fibration $p\colon S^n \to B$ whose base B consists of more than one point is not homotopic to a constant map.

It follows from the covering homotopy theorem that for any locally trivial fibration $p\colon E \to B$, a path γ joining points a and b in the base B

induces a map of fibers $p^{-1}(a) \to p^{-1}(b)$, which is defined up to homotopy. First, we construct the map itself. We set $F_a = p^{-1}(a)$. To apply Theorem 4.14, we take $X = F_a$, $X' = \varnothing$, $\tilde{h} = \mathrm{id}_X$, and $H(x,t) = \gamma(t)$. According to this theorem, there exists a homotopy $\widetilde{H}\colon F_a \times I \to E$ that covers the homotopy H and is an extension of the map $\tilde{h}\colon F_a \times \{0\} \hookrightarrow E$. For this homotopy, we have $p\widetilde{H}(x,t) = \gamma(t)$. Therefore, $p\widetilde{H}(x,1) = \gamma(1) = b$, i.e., $p\widetilde{H}(x,1) \in F_b$. The required map is defined by $x \mapsto \widetilde{H}(x,1)$. This map depends on the choice of $\widetilde{H}$. Let us show that for any two homotopies $\widetilde{H}_0$ and $\widetilde{H}_1$ constructed for homotopic paths γ_0 and γ_1, the maps $x \mapsto \widetilde{H}_0(x,1)$ and $x \mapsto \widetilde{H}_1(x,1)$ are homotopic. We again apply Theorem 4.14. This time, we have two parameters, the path parameter t and the homotopy parameter τ, because we have the family of paths $\gamma_\tau(t)$. We take $X = F_a \times I_\tau$, $X' = F_a \times \{0\} \cup F_a \times \{1\}$, $\tilde{h}(y,\tau) = y$ for all $\tau \in I_\tau$ and $y \in F_a$, $H(y,\tau,t) = \gamma_\tau(t)$, $\widetilde{H}'(y,0,t) = H_0(y,t)$, and $\widetilde{H}'(y,1,t) = H_1(y,t)$. By Theorem 4.14, there exists a homotopy $\widetilde{H}(y,\tau,t)$ that covers the homotopy H and is a simultaneous extension for the homotopy $\widetilde{H}'$ and the map $\tilde{h}\colon X \times \{0\} \to E$. The required homotopy $G\colon F_a \times I_\tau \to F_b$ is defined by $G(y,\tau) = \widetilde{H}(y,\tau,1)$.

4.2. Homotopy Groups. The homotopy groups $\pi_n(X,x_0)$ are a generalization of the fundamental group $\pi_1(X,x_0)$. We first define the sets $\pi_n(X,x_0)$ for $n \geq 0$, and then endow each set $\pi_n(X,x_0)$ (for $n \geq 1$) with a group structure. We mark a point s_0 in the sphere S^n and say that two maps $(S^n,s_0) \to (X,x_0)$ are equivalent if they are homotopic (i.e., joined by a homotopy $h_t\colon S^n \to X$ such that $h_t(s_0) = x_0$ for all $t \in [0,1]$). The set $\pi_n(X,x_0)$ consists of equivalence classes. In particular, the elements of $\pi_0(X,x_0)$ are the path-connected components of the space X. The maps $(S^n,s_0) \to (X,x_0)$ are called *n-spheroids*; sometimes, it is convenient to represent such maps as $(I^n,\partial I^n) \to (X,x_0)$ or $(D^n,\partial D^n) \to (X,x_0)$ (recall that $I^n/\partial I^n \approx D^n/\partial D^n \approx S^n$).

To make the set $\pi_n(X,x_0)$ into a group, we must construct a map $fg\colon (S^n,s_0) \to (X,x_0)$ for any two maps $f,g\colon (S^n,s_0) \to (X,x_0)$. The construction is shown in the top part of Figure 20; the bottom part of Figure 20 illustrates the same construction for maps $(I^n,\partial I^n) \to (X,x_0)$.

For $n \geq 2$, the order in which two maps f and g are multiplied does not matter, because the maps fg and gf are homotopic. The homotopy is easy to construct by looking at Figure 21.

For any map $f\colon (I^n,\partial I^n) \to (X,x_0)$, there exists a map $\hat{f}\colon (I^n,\partial I^n) \to (X,x_0)$ such that the product $f\hat{f}$ is homotopic to a constant. The map $\hat{f}$ can be defined, e.g., by representing the cube I^n in the form $I^n = I^{n-1} \times [-1,1]$ and setting $\hat{f}(x,s) = f(x,-s)$. For this $\hat{f}$, we have $f\hat{f}(x,s) = f\hat{f}(x,-s)$

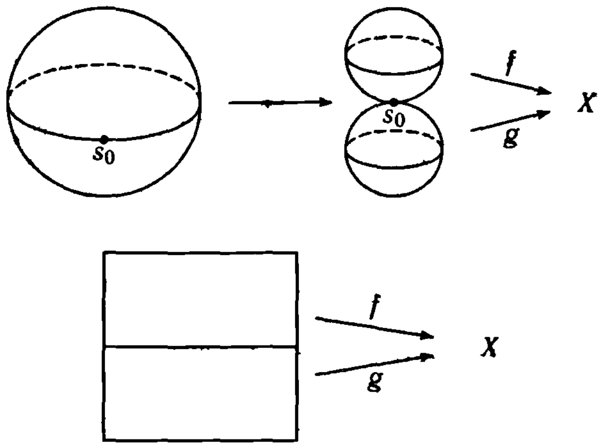

Figure 20. The product of two spheroids

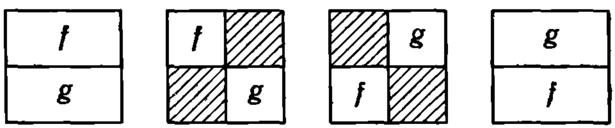

Figure 21. The commutativity of the homotopy group

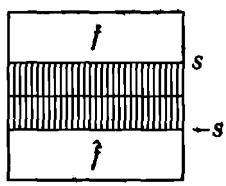

Figure 22. The inverse element

(see Figure 22). Consider the family of maps

$$g_t(x,s) = \begin{cases} f\hat{f}(x,s) & \text{if } |s| \geq t, \\ f\hat{f}(x,t) & \text{if } |s| \leq t. \end{cases}$$

Clearly, $g_0 = f\hat{f}$ and g_1 is a constant map.

Exercise 30. (a) Prove that if $f \sim f_1$ and $g \sim g_1$, then $fg \sim f_1 g_1$.

(b) Prove that $f(gh) \sim (fg)h$.

In considering homotopy groups $\pi_n(X, x_0)$ for $n \geq 1$, the space X is usually assumed to be path-connected. In this case, the groups $\pi_n(X, x_0)$ and $\pi_n(X, x_1)$ are isomorphic, but the isomorphism is not canonical: it depends on the path from x_0 to x_1. For a given path α between x_0 and x_1, the

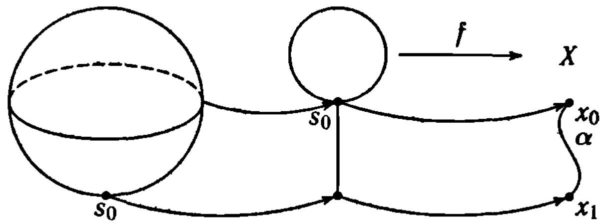

Figure 23. Changing the base point

isomorphism $\pi_n(X, x_0) \to \pi_n(X, x_1)$ is constructed as follows. Given a map $f\colon (S^n, s_0) \to (X, x_0)$, consider the map $S^n \to S^n \vee I$ which takes the equator to s_0 and each section of the southern hemisphere by a plane parallel to the equator to a point of the interval I; the southern pole is the base point s_0, and it is mapped to the free endpoint of I (see Figure 23). The composition of this map and the map $S^n \vee I \to X$ coinciding with f on S^n and with α on I is an element of the group $\pi_n(X, x_1)$. This element depends only on the homotopy classes of the map f and the path α (we assume here that homotopic paths have the same start and end points). It is easy to verify that the paths α and α^{-1} induce mutually inverse maps. In particular, if α is a loop at x_0, then α induces an automorphism of the group $\pi_n(X, x_0)$. This automorphism depends only on the element of $\pi_1(X, x_0)$ represented by the loop α. If each element of $\pi_1(X, x_0)$ induces the identity automorphism of $\pi_n(X, x_0)$, then the space X is said to be *n-simple*.

If a space X is n-simple, then the groups $\pi_n(X, x)$, where $x \in X$, are canonically isomorphic to each other for this n, and we can denote them by $\pi_n(X)$.

Problem 53. Let X be a CW-complex, and let X^n be its n-skeleton. Prove that the embedding $i\colon X^n \to X$ induces an isomorphism $i_*\colon \pi_k(X^n, x_0) \to \pi_k(X, x_0)$ for $k < n$ and an epimorphism for $k = n$.

To solve the following two problems, the relation $\pi_n(S^n) = \mathbb{Z}$ ($n \in \mathbb{N}$) is needed (see p. 235).

Problem 54. Prove that for $n \geq 2$, the group $\pi_n(S^n \vee S^1, x_0)$ is a free Abelian group of (infinite) countable rank.

Problem 55. Prove that the space $S^n \vee S^1$ is not n-simple.

4.3. Exact Sequences for Fibrations. Let $p\colon E \to B$ be a locally trivial fibration with path-connected base B. Choose a base point $b_0 \in B$. The map p induces a homomorphism $p_*\colon \pi_n(E, e_0) \to \pi_n(B, b_0)$, where $e_0 \in p^{-1}(b_0)$. Let $i\colon F \hookrightarrow E$ be the composition of the homeomorphism $F \approx p^{-1}(b_0)$ and the embedding $p^{-1}(b_0) \subset E$. The map i induces a homomorphism

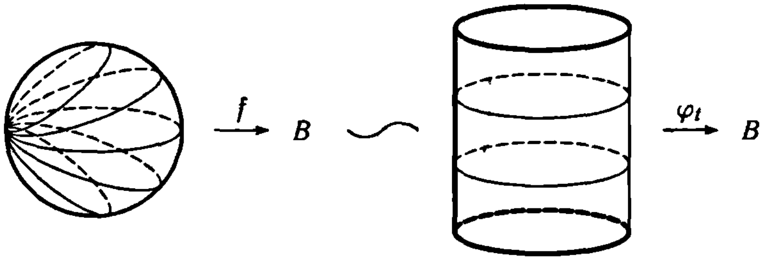

Figure 24. Representation of a map as a homotopy

$i_*\colon \pi_n(F,e_0) \to \pi_n(E,e_0)$; here and in what follows, we identify the fiber F with $p^{-1}(b_0)$.

There is also a third homomorphism $\partial_*\colon \pi_n(B,b_0) \to \pi_{n-1}(F,e_0)$, which is defined as follows. A map $f\colon (S^n,s_0) \to (B,b_0)$ can be represented as a homotopy $\varphi_t\colon (S^{n-1},s_0) \to (B,b_0)$ between two copies φ_0, φ_1 of the constant map, $\varphi_0,\varphi_1\colon S^{n-1} \to b_0$ (see Figure 24). According to the covering homotopy theorem, there exists a homotopy $\tilde{\varphi}_t\colon S^{n-1} \to E$ for which $\tilde{\varphi}_0(S^{n-1}) = e_0$, $\tilde{\varphi}_t(s_0) = e_0$, and $p\tilde{\varphi}_t = \varphi_t$ (see Figure 25). Clearly, $\tilde{\varphi}_1(S^{n-1}) \subset p^{-1}(b_0) = F$, because $\varphi_1(S^{n-1}) = b_0$. For $\partial_* f$ we take the homotopy class of the map $\tilde{\varphi}_1\colon (S^{n-1},s_0) \to (F,e_0)$. We must only verify that this map is well defined, i.e., homotopic maps f and f' correspond to homotopic maps $\tilde{\varphi}_1$ and $\tilde{\varphi}'_1$. This can easily be done by applying the covering homotopy theorem once more.

Recall that a sequence of group homomorphisms

$$\cdots \longrightarrow G_i \xrightarrow{\varphi_i} G_{i-1} \xrightarrow{\varphi_{i-1}} G_{i-2} \longrightarrow \cdots$$

is said to be *exact* if $\operatorname{Ker}\varphi_{i-1} = \operatorname{Im}\varphi_i$, where $\operatorname{Ker}\varphi_{i-1} = \{g \in G_{i-1} : \varphi_{i-1}(g) = 0\}$ (this is the *kernel* of the homomorphism φ_{i-1}) and $\operatorname{Im}\varphi_i = \{\varphi_i(g) : g \in G_i\}$ (this is the *image* of the homomorphism φ_i).

Theorem 4.15. *The sequence of homomorphisms*

$$\cdots \longrightarrow \pi_n(F,e_0) \xrightarrow{i_*} \pi_n(E,e_0) \xrightarrow{p_*} \pi_n(B,b_0) \xrightarrow{\partial_*} \pi_{n-1}(F,e_0) \longrightarrow \cdots$$

is exact.

Proof. We must prove six inclusions of the form $\operatorname{Im} i_* \subset \operatorname{Ker} p_*$, $\operatorname{Ker} p_* \subset \operatorname{Im} i_*$, etc. We prove them separately. When there is no risk of confusion, we do not mention base points and identify the elements of homotopy groups with the spheroids representing them.

(1) $\operatorname{Im} i_* \subset \operatorname{Ker} p_*$. Any spheroid belonging to $\operatorname{Im} i_*$ has a representative $f\colon S^n \to E$ whose image is contained in F. The map pf is constant; therefore, $f \in \operatorname{Ker} p_*$.

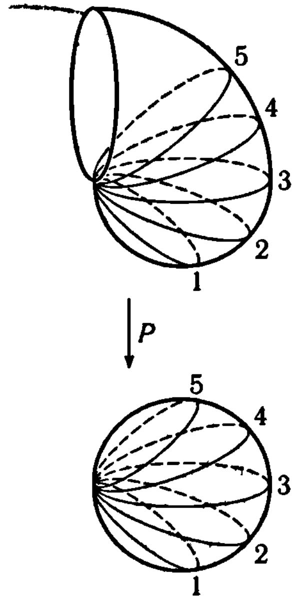

Figure 25. Lifting of the homotopy

(2) $\operatorname{Ker} p_* \subset \operatorname{Im} i_*$. Any spheroid belonging to $\operatorname{Ker} p_*$ has a representative $f: S^n \to E$ for which the spheroid $pf: S^n \to B$ is contractible. For a homotopy $H: S^n \times I \to B$ between pf and the constant map, there exists a covering homotopy $\widetilde{H}: S^n \times I \to E$ between f and some map f_1. The map pf_1 is constant, i.e., the image of f_1 is contained in F. This means that $f_1 \in \operatorname{Im} i_*$.

(3) $\operatorname{Im} p_* \subset \operatorname{Ker} \partial_*$. Suppose that $f = p\tilde{f}$, where $\tilde{f}: S^n \to E$ is a spheroid. The spheroid $\tilde{f}$ can be represented as a homotopy $\tilde{\varphi}_t$. The map $\tilde{\varphi}_1$ is constant, and hence $\partial_* f = 0$.

(4) $\operatorname{Ker} \partial_* \subset \operatorname{Im} p_*$. We represent a spheroid $f: S^n \to B$ as a homotopy $\varphi_t: S^{n-1} \to B$ and consider a covering homotopy $\tilde{\varphi}_t$. Suppose that $f \in \operatorname{Ker} \partial_*$, i.e., $\tilde{\varphi}_1: S^{n-1} \to F$ is homotopic to a constant map. Let $\tilde{\alpha}_t$ be a homotopy in F between $\tilde{\varphi}_1$ and the constant map. Consider the homotopy

$$\tilde{\psi}_t = \begin{cases} \tilde{\varphi}_{2t} & \text{if } t \in \left[0, \frac{1}{2}\right], \\ \tilde{\alpha}_{2t-1} & \text{if } t \in \left[\frac{1}{2}, 1\right]. \end{cases}$$

Corresponding to this homotopy is a spheroid $\tilde{g}: S^{n-1} \to F$ for which $g = p\tilde{g}$ is homotopic to f. Therefore, $f \in \operatorname{Im} p_*$.

(5) $\operatorname{Im} \partial_* \subset \operatorname{Ker} i_*$. Suppose that a spheroid $f: S^{n+1} \to B$ is represented as a homotopy $\varphi_t: S^n \to B$ and $\tilde{\varphi}_t$ is a lifting of this homotopy. The map $\tilde{\varphi}_0$

is constant; therefore, the homotopy $\tilde{\varphi}_t$ can be treated as a map $D^{n+1} \to B$. Hence $\tilde{\varphi}_1 \colon S^n \to F$ is homotopic in E to the constant map.

(6) $\operatorname{Ker} i_* \subset \operatorname{Im} \partial_*$. Let $f \colon S^n \to F$ be a spheroid contractible in E. The projection of a homotopy of the spheroid f to the constant map can be regarded as a spheroid $g \colon S^{n+1} \to B$, and $\partial_* g = f$. □

Remark 4.1. If the total space E of the fibration is path-connected, then the sets $\pi_0(E, e_0)$ and $\pi_0(B, b_0)$ are one-point. In this case, there is a one-to-one correspondence between $\pi_0(F, e_0)$ and the coset space $\pi_1(B, b_0)/\operatorname{Im} p_*$ (the subgroup $\operatorname{Im} p_* \subset \pi_1(B, b_0)$ is not necessarily normal, so the coset space may not be a group).

Example. For $n \geq 2$, $\pi_n(S^1) = 0$.

Proof. Consider a covering $\mathbb{R} \to S^1$. Any covering is a fibration with discrete fiber F, and $\pi_{n-1}(F) = 0$ for $n \geq 2$. Using the exact sequence, we obtain $\pi_n(S^1) = \pi_n(\mathbb{R}^1) = 0$ for $n \geq 2$. □

Example. $\pi_2(S^2) = \mathbb{Z}$, and $\pi_n(S^2) = \pi_n(S^3)$ for $n \geq 3$.

Proof. In the space $\mathbb{C}^2$ with coordinates z and w, consider the sphere S^3 given by the equation $|z|^2 + |w|^2 = 1$. On S^3, the group $S^1 = \{e^{i\alpha}\}$ acts (both coordinates z and w are multiplied by $e^{i\alpha}$). The quotient of S^3 by this action is homeomorphic to the projective space $\mathbb{C}P^1$ with homogeneous coordinates $(z : w)$. We show that the projection $p \colon S^3 \to S^3/S^1 \approx \mathbb{C}P^1$ is a locally trivial fibration (it is called the *Hopf fibration*).

We cover $\mathbb{C}P^1$ by the open sets U_1 and U_2 obtained from $\mathbb{C}P^1$ by removing the points $(1 : 0)$ and $(0 : 1)$. Over each of these sets, the map p is a trivial fibration with fiber S^1. Indeed, every point of the sphere S^3 can be represented as $(ae^{i\varphi}, be^{i\psi})$, where a and b are nonnegative numbers satisfying the equality $a^2 + b^2 = 1$. Homeomorphisms $h_i \colon p^{-1}(U_i) \to U_i \times S^1$ compatible with the projection can be defined as

$$h_1(ae^{i\varphi}, be^{i\psi}) = \left(\frac{a}{b}e^{i(\varphi-\psi)}, e^{i\psi}\right) \quad \text{and} \quad h_2(ae^{i\varphi}, be^{i\psi}) = \left(\frac{b}{a}e^{i(\psi-\varphi)}, e^{i\varphi}\right).$$

The exact sequence for the Hopf fibration is

$$\cdots \longrightarrow \pi_2(S^3) \xrightarrow{p_*} \pi_2(S^2) \xrightarrow{\partial_*} \pi_1(S^1) \xrightarrow{i_*} \pi_1(S^3) \longrightarrow \cdots.$$

We already know that $\pi_k(S^n) = 0$ for $k < n$ (see Theorem 3.19 on p. 105). Therefore, $\pi_2(S^2) = \pi_1(S^1) = \mathbb{Z}$.

Consider another segment of the exact sequence for the Hopf fibration:

$$\cdots \longrightarrow \pi_n(S^1) \xrightarrow{i_*} \pi_n(S^3) \xrightarrow{p_*} \pi_n(S^2) \xrightarrow{\partial_*} \pi_{n-1}(S^1) \longrightarrow \cdots.$$

If $n \geq 3$, then $\pi_n(S^1) = \pi_{n-1}(S^1) = 0$; therefore, $p_* \colon \pi_n(S^3) \to \pi_n(S^2)$ is an isomorphism. □

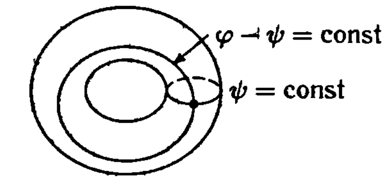

Figure 26. A meridian and a parallel

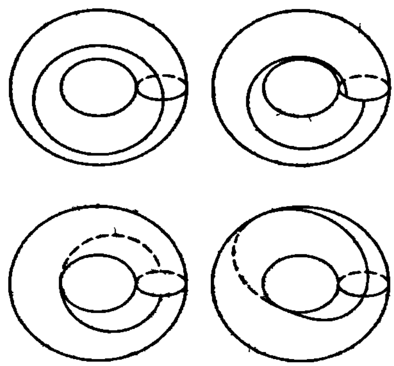

Figure 27. The transformation of the solid torus

Now, we study in more detail the geometric structure of the Hopf fibration, in particular, the arrangement of its fibers in S^3. We represent $\mathbb{C}P^1$ as the union of the two closed sets $D_1^2 \subset U_1$ and $D_2^2 \subset U_2$ determined by the inequalities $a \geq b$ and $a \leq b$ and consider the homeomorphisms $h_i: p^{-1}(D_i^2) \to D_i^2 \times S^1$ defined above. The total space S^3 of the fibration is obtained by attaching the two solid tori $D_1^2 \times S^1$ and $D_2^2 \times S^1$ to each other via a homeomorphism between their boundaries. The meridians of the solid torus $D_1^2 \times S^1$ are determined by the equations $\psi = \text{const}$ and the parallels by the equations $\varphi - \psi = \text{const}$ (see Figure 26); we assume that the meridians and parallels lie on the boundary, i.e., $a/b = 1$. Each fiber is determined by two equations of the form $\varphi - \psi = \text{const}$ and $a/b = \text{const}$.

We can transform the solid torus $D_1^2 \times S^1$ in such a way that the meridians (which are given by $\psi = \text{const}$) remain invariant, while the parallels are transformed into the lines $\varphi = \text{const}$. To this end, we cut the solid torus by a meridian plane and turn the upper part of the torus above the cut through the angle 2π leaving the lower part intact (see Figure 27). Choosing a suitable direction of rotation, we obtain a self-homeomorphism of the solid torus which takes the parallels to the lines $\varphi = \text{const}$ (see Figure 28); the fibers become linked as shown in Figure 29.

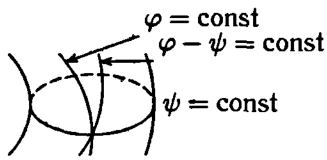

Figure 28. The image of a parallel under the transformation

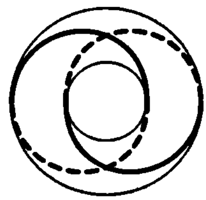

Figure 29. The Hopf link

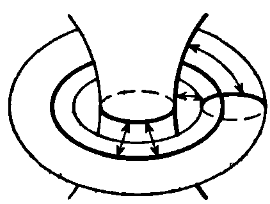

Figure 30. Gluing together solid tori

For the second solid torus $D_2^2 \times S^1$, we construct a similar homeomorphism. Then, we attach the boundaries of the transformed tori $D_1^2 \times S^1$ and $D_2^2 \times S^1$ to each other by identifying points with the same coordinates φ and ψ; thereby, we identify the meridians of one solid torus with the parallels of the other (see Figure 30).

Remark 4.2. The arrangement of the fibers of the Hopf fibration can also be described using the fact that the fibers are the sections of the sphere S^3 by the complex lines $\alpha z = \beta w$ with $\alpha, \beta \in \mathbb{C}$.

Problem 56. Suppose that S^3 is the group of quaternions of unit length and $S^1 \subset S^3$ is its subgroup consisting of the quaternions of the form $x_1 + ix_2$. The group S^1 acts on S^3 by right multiplication. Prove that the map $p\colon S^3 \to S^3/S^1$ is the Hopf fibration.

Problem 57. In the space of real matrices $\left(\begin{smallmatrix} x_1 & x_2 \\ x_3 & x_4 \end{smallmatrix}\right)$, consider the sphere given by the equation $x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$. Prove that the singular matrices divide this sphere into two solid tori.

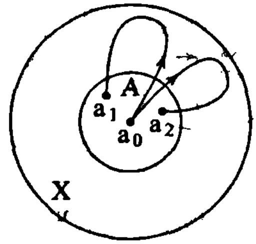

Figure 31. Elements of the set $\pi_1(X, A, a_0)$

Problem 58. Let $p\colon S^3 \to \mathbb{C}P^1$ be the Hopf fibration. Prove that $D^4 \cup_p \mathbb{C}P^1 = \mathbb{C}P^2$. (It is assumed that $S^3 = \partial D^4$.)

Problem 59. Prove that there exists no retraction $r\colon \mathbb{C}P^2 \to \mathbb{C}P^1$. (It is assumed that $\mathbb{C}P^1$ is naturally embedded in $\mathbb{C}P^2$.)

Problem 60. Prove that the complement of the Hopf link in S^3 is homotopy equivalent to the torus T^2.

The Hopf fibration $S^3 \to S^2$ with fiber S^1 has multidimensional generalizations. One of these generalizations is as follows. Let us represent S^{2n+1} as the unit sphere in $\mathbb{R}^{2n+2} \cong \mathbb{C}^{n+1}$ and identify the points $(\lambda z_1, \dots, \lambda z_{n+1}) \in S^{2n+1}$ for all $\lambda \in \mathbb{C}$ with $|\lambda| = 1$. As a result, we obtain a map $S^{2n+1} \to \mathbb{C}P^n$, which is a fibration with fiber S^1. Taking the quaternions $\mathbb{H}$ instead of the complex numbers, we obtain a fibration $S^{4n+3} \to \mathbb{H}P^n$ with fiber S^3. Thus, for $n = 1$, we obtain a fibration $S^7 \to S^4$ with fiber S^3. All these fibrations are called *Hopf fibrations* as well.

4.4. Relative Homotopy Groups. For $n \geq 1$ and a pair of spaces $X \supset A$ with base point $a_0 \in A$, an n-dimensional *relative spheroid* can be defined as a map $f\colon (D^n, \partial D^n, s_0) \to (X, A, a_0)$; it is assumed that f is a map of triples, i.e., $f(\partial D^n) \subset A$ and $f(s_0) = a_0$. Relative spheroids f_0 and f_1 are said to be *homotopic* if they are connected by a homotopy f_t such that $f_t(\partial D^n) \subset A$ and $f_t(s_0) = a_0$.

The set $\pi_n(X, A, a_0)$, where $n \geq 1$, consists of the equivalence classes of relative n-spheroids. This set does not generally admit a group structure. The point is that its elements are represented by paths starting at a_0 and ending at $a \in A$, and there is no natural way to compose a path starting at a_0 from two such paths (see Figure 31). Although $\pi_1(X, A, a_0)$ is not a group, it has a distinguished element, the class of the constant map. We refer to this element as the zero element.

For $n \geq 2$, the set $\pi_n(X, A, a_0)$ can be endowed with the structure of a group. To define it, we use the description of relative spheroids as maps of

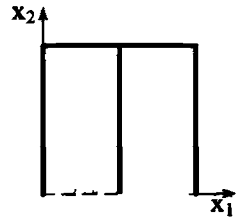

Figure 32. The composition of relative spheroids

the cube $I^n = \{(x_1, \dots, x_n) : 0 \le x_i \le 1\}$. However, we identify D^n with the quotient space $I^n/(\partial I^n \setminus I^{n-1})$, where I^{n-1} is determined by the equation $x_n = 0$, rather than with the cube I^n itself. Thus, a relative spheroid is a map $f: I^n \to X$ for which $f(\partial I^n) \subset A$ and $f(\partial I^n \setminus I^{n-1}) = a_0$.

We define the composition of relative spheroids $f, g: I^n \to X$ by

$$fg(x_1, x_2, \dots, x_n) = \begin{cases} f(2x_1, x_2, \dots, x_n) & \text{if } 0 \le x_1 \le \frac{1}{2}, \\ g(2x_1 - 1, x_2, \dots, x_n) & \text{if } \frac{1}{2} \le x_1 \le 1. \end{cases}$$

In Figure 32, the bold line represents the set of points which the map fg takes to the base point a_0 (for $n = 2$).

For $n = 2$, the construction used in the proof of the commutativity of the group $\pi_n(X)$ does not apply because a relative spheroid $f: I^2 \to X$ does not necessarily take $I^1 \subset I^2$ to the base point (in Figure 32, the set I^1 is shown by the dashed line). But for $n \ge 3$, any relative spheroid $f: I^n \to X$ maps ∂I^{n-1} to the base point, and the construction does apply.

The following assertion is often used in dealing with relative homotopy groups.

Theorem 4.16. *A relative spheroid $f: (D^n, \partial D^n, s_0) \to (X, A, a_0)$ represents the zero element of the group $\pi_n(X, A, a_0)$ if and only if there exists a homotopy h_t between the map $h_0 = f$ and some map h_1 for which $h_1(D^n) \subset A$ such that all the h_t coincide with f on S^{n-1}.*

Proof. Suppose that there exists a homotopy h_t with the properties specified in the statement of the theorem. Consider the homotopy $g_t(s) = h_1((1-t)s + ts_0)$. This homotopy joins h_1 to the constant map $D^n \to a_0$. The condition $g_t(S^{n-1}) \subset A$ holds because $h_1(D^n) \subset A$.

Now assume that there exists a homotopy $f_t: (D^n, \partial D^n, s_0) \to (X, A, a_0)$ between $f_0 = f$ and the constant map $D^n \to a_0$. It is required to construct a homotopy h_t fixed on S^{n-1}. We set

$$h_t(s) = \begin{cases} f_t\left(\frac{s}{1-t/2}\right) & \text{if } 0 \le \|s\| < 1 - t/2, \\ f_{2-2\|s\|}\left(\frac{s}{\|s\|}\right) & \text{if } 1 - t/2 \le \|s\| \le 1, \end{cases}$$

Clearly, $h_0 = f_0 = f$, $h_1(D^n) \subset A$, and $h_t(s) = f_0(s)$ for $s \in S^{n-1}$. $\square$

The embedding $i\colon A \to X$ induces a homomorphism $i_*\colon \pi_n(A, a_0) \to \pi_n(X, a_0)$. Any absolute spheroid $(D^n, s_0) \leftrightarrow (X, a_0)$ can be regarded as a relative spheroid $(D^n, S^{n-1}, s_0) \to (X, A, a_0)$, because S^{n-1} is mapped to $a_0 \in A$. Homotopic absolute spheroids are homotopic as relative spheroids; thus, we have a homomorphism $p_*\colon \pi_n(X, a_0) \to \pi_n(X, A, a_0)$. Finally, there is a homomorphism $\partial_*\colon \pi_n(X, A, a_0) \to \pi_{n-1}(A, a_0)$, which takes each relative spheroid $(D^n, S^{n-1}, s_0) \to (X, A, a_0)$ to its restriction to (S^{n-1}, s_0); from the definition of a homotopy between relative spheroids one can easily see that homotopic relative spheroids are mapped to homotopic absolute spheroids.

Theorem 4.17 (exact sequence for a pair). *The sequence of homomorphisms*

$$\cdots \to \pi_n(A, a_0) \xrightarrow{i_*} \pi_n(X, a_0) \xrightarrow{p_*} \pi_n(X, A, a_0) \xrightarrow{\partial_*} \pi_{n-1}(A, a_0) \to \cdots$$

is exact.

Proof. (1) $\operatorname{Im} i_* \subset \operatorname{Ker} p_*$. According to Theorem 4.16, any relative spheroid f with $f(D^n) \subset A$ represents the zero element of the group $\pi_n(X, A, a_0)$.

(2) $\operatorname{Ker} p_* \subset \operatorname{Im} i_*$. Suppose that $f\colon I^n \hookrightarrow X$ takes I^{n-1} to A and $\partial I^n \setminus I^{n-1}$ to a_0 and there exists a homotopy $F\colon I^n \times I \to X$ in the class of relative spheroids which joins f to the constant map to a_0. The section of the cube $I^{n+1} = I^n \times I$ by the hyperplane $tx_n + (1-t)x_{n+1} = 0$, where $t \in [0,1]$, is homeomorphic to I^n. The restriction of F to this section is a homotopy of absolute spheroids. For $t = 0$, the section is the face $x_{n+1} = 0$, and the restriction to this face coincides with f. For $t = 1$, the section is the face $x_n = 0$, which is mapped to A by assumption.

(3) $\operatorname{Im} p_* \subset \operatorname{Ker} \partial_*$. Any absolute spheroid $f\colon D^n \to X$ maps ∂D^n to a_0.

(4) $\operatorname{Ker} \partial_* \subset \operatorname{Im} p_*$. Consider a relative spheroid $f\colon I^n \to X$ whose restriction to I^{n-1} is homotopic to the constant map to a_0 (in the class of maps $I^{n-1} \to A$). Let $g_t\colon I^{n-1} \to A$ be a homotopy between $f \restriction I^{n-1}$ and the constant map. Consider the homotopy $f_t\colon \partial I^n \to X$ that coincides with g_t on I^{n-1} and takes $\partial I^n \setminus I^{n-1}$ to a_0. By the Borsuk lemma (see p. 164), this homotopy can be extended to a homotopy of f. As a result, we obtain a homotopy in the class of relative spheroids that joins f to a relative spheroid taking ∂I^n to a_0.

(5) $\operatorname{Im} \partial_* \subset \operatorname{Ker} i_*$. If an absolute spheroid $f\colon I^{n-1} \to A$ is the restriction of a relative spheroid $g\colon I^n \to X$ to $I^{n-1} = I^{n-1} \times \{0\} \subset I^n$, then g can be regarded as a homotopy $g_t\colon I^{n-1} \to X$ in X between the spheroid f and the constant map.

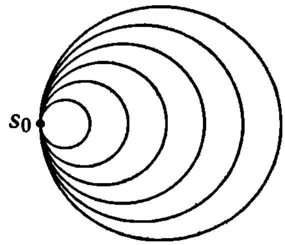

Figure 33. The homotopy $\widetilde{H}$ as a relative spheroid

(6) $\operatorname{Ker} i_* \subset \operatorname{Im} \partial_*$. A homotopy $g_t \colon I^{n-1} \to X$ between spheroids $g_0 \colon I^{n-1} \to A$ and $g_1 \colon I^{n-1} \to a_0$ can be regarded as a relative spheroid $g \colon I^n \to X$ whose restriction to I^{n-1} coincides with g_0. □

Problem 61. Prove that $\pi_{n-1}(X) \cong \pi_n(CX, X)$ for $n \geq 2$.

The following theorem shows that the exact sequence for a fibration is a special case of the exact sequence for a pair.

Theorem 4.18. *Suppose that* $p \colon E \to B$ *is a locally trivial fibration,* $e_0 \in E$, $b_0 = p(e_0)$, *and* $F = p^{-1}(e_0)$. *Then the maps* $p_* \colon \pi_n(E, F, e_0) \to \pi_n(B, b_0)$ *are isomorphisms for all* $n \geq 1$.

Proof. Suppose that $\tilde{h} \colon D^n \to E$ is a relative spheroid such that the absolute spheroid $h = p\tilde{h}$ represents the zero element of the group $\pi_n(B, b_0)$. Then there exists a homotopy $H \colon D^n \times I \to B$ between h and the constant map (in the class of maps taking S^{n-1} to b_0). By the covering homotopy theorem, there exists a homotopy $\widetilde{H} \colon D^n \times I \to E$ which covers the homotopy H and is simultaneously an extension of $\tilde{h}$ and $\widetilde{H}' \colon \{s_0\} \times I \to e_0 \in E$. The homotopy $\widetilde{H}$ joins the relative spheroid $\tilde{h}$ to a map whose image is contained in F. Such a map is homotopic to the constant map in the class of relative spheroids. Therefore, p_* is a monomorphism.

Let $h \colon S^n \to B$ be a spheroid. It can be regarded as a homotopy $H \colon S^{n-1} \times I \to B$ between the constant maps $S^{n-1} \to b_0 \in B$ (look again at Figure 24 on p. 169). Let $\widetilde{H} \colon S^{n-1} \times I \to E$ be a homotopy which covers H and is an extension of the constant map $S^{n-1} \times \{0\} \to e_0 \in E$ and of the homotopy $\widetilde{H} \colon \{s_0\} \times I \to e_0 \in E$. This homotopy can be regarded as a relative spheroid $\tilde{h} \colon D^n \to E$ (see Figure 33). Moreover, $h = p\tilde{h}$, and hence p_* is an epimorphism. □

4.5. The Whitehead Theorem. The Whitehead theorem asserts that if a map $X \to Y$ of connected CW-complexes induces isomorphisms of all homotopy groups, then it is a homotopy equivalence. We start with the special case in which the space Y is one-point. The general Whitehead theorem is easily derived from the relative version of this special case.

Theorem 4.19. *Suppose that X is a CW-complex, x_0 is a vertex of X, and $\pi_n(X, x_0) = 0$ for all $n \geq 0$. Then X is a contractible space.*

Proof. Let $f_0 = \mathrm{id}_X$, and let f_0' be the restriction of f_0 to X^0 (the 0-skeleton of X). The equality $\pi_0(X, x_0) = 0$ means that each vertex $x_\alpha \in X^0$ is joined by a path γ_α with x_0. The formula $F_0'(x_\alpha) \coloneqq \gamma_\alpha(t)$ defines a homotopy between f_0' and the constant map $X^0 \to x_0$. According to the Borsuk lemma, this homotopy can be extended to a homotopy $F_0(x, t)$ on the entire space X. This homotopy joins the identity map $F_0(x, 0) = f_0(x)$ to the map $F_0(x, 1) = f_1(x)$ taking X^0 to x_0.

Let f_1' be the restriction of f_1 to X^1. The equality $\pi_1(X, x_0) = 0$ implies the existence of a homotopy F_1' between f_1' and the constant map $X^1 \to x_0$. Extending this homotopy to X, we obtain a homotopy F_1 between f_1 and a map f_2 taking X^1 to x_0.

Similarly, applying the equality $\pi_n(X, x_0) = 0$, we can construct a homotopy F_n between maps f_n and f_{n+1} such that f_{n+1} takes X^n to x_0 for every n.

This completes the proof of the theorem in the case of $\dim X < \infty$. If $\dim X = \infty$, then a homotopy F between id_X and the constant map $X \to x_0$ can be constructed as follows. We set $F(x, t_n) \eqqcolon f_n(x)$ for $t_n = \frac{2^n - 1}{2^n}$ and define F to coincide with F_n between t_n and t_{n+1}. Clearly, the restrictions of F to $X^n \times I$ are continuous for all n; therefore, the entire map F is continuous. □

A space $A \subset X$ is called a *deformation retract* of X if there exists a homotopy $f_t\colon X \to X$, where $t \in [0, 1]$, such that $f_t \restriction A = \mathrm{id}_A$ for all t, $f_0 = \mathrm{id}_X$, and $f_1(X) \subset A$. For the map $r = f_1\colon X \to A$, we have $ri = \mathrm{id}_A$, and ir is homotopic to id_X (here $i\colon A \to X$ is the natural embedding of A into X). Clearly, any deformation retract A of X is homotopy equivalent to X. Indeed, for the maps $r\colon X \to A$ and $i\colon A \to X$, we have $ir \simeq \mathrm{id}_X$ and $ri = \mathrm{id}_A$.

Problem 62. Let $m, n \geq 1$. Prove that the space $S^m \vee S^n \coloneqq (S^m \times \{x_0\}) \cup (\{y_0\} \times S^n)$, where $x_0 \in S^n$ and $y_0 \in S^m$, is a deformation retract of the space obtained from $S^m \times S^n$ by removing a point not belonging to $S^m \vee S^n$.

Theorem 4.19 has the following relative version.

Theorem 4.20. *Suppose that X is a path-connected CW-complex, A is its subcomplex, a_0 is a vertex of A, and $\pi_n(X, A, a_0) = 0$ for all $n \geq 1$. Then A is a deformation retract of X.*

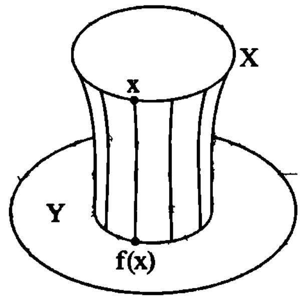

Figure 34. The cylinder of a map

Proof. Suppose that $n \geq 1$ and $g_n \colon X \to X$ is a map such that $g_n(X^{n-1}) \subset A$ and $g_n \restriction A = \mathrm{id}_A$. According to Theorem 4.16, the equality $\pi_n(X, A, a_0) = 0$ implies the existence of a homotopy which is fixed on A and joins the map $g_n \restriction X^n \cup A$ to a map $g'_{n+1} \colon X^n \cup A \to A$. Extending this homotopy to the entire space X, we obtain g_{n+1}.

The map g_1 is constructed as follows. We join a vertex $x_\alpha \in X^0 \setminus A^0$ to a_0 by a path and set $g'_1(x_\alpha) = a_0$. Clearly, the map $g'_1 \colon X^0 \cup A \to X$ is homotopic to $\mathrm{id}_{X^0 \cup A}$. Extending the homotopy to X, we obtain the required map g_1. □

Now, we can prove the Whitehead theorem itself.

Theorem 4.21 (Whitehead). *Suppose that X and Y are connected CW-complexes, $x_0 \in X$ and $y_0 \in Y$ are their vertices, and a map $f \colon (X, x_0) \to (Y, y_0)$ induces isomorphisms $f_* \colon \pi_n(X, x_0) \to \pi_n(Y, y_0)$ for all $n \geq 1$. Then f is a homotopy equivalence.*

Proof. Let C_f be the space obtained by attaching $X \times I$ to Y via the map $X \times \{0\} = X \xrightarrow{f} Y$ (see Figure 34). The space C_f is called the *cylinder* of the map f.

The identity map $C_f \to C_f$ is homotopic to $g \colon C_f \to Y \subset C_f$, and the homotopy is the identity on Y. Therefore, the spaces C_f and Y are homotopy equivalent.

The map f can be represented as the composition $X \xrightarrow{i} C_f \xrightarrow{g} Y$, where i is the natural embedding $X = X \times \{1\} \to C_f$. The homomorphisms $i_* g_*$ and g_* are isomorphisms; therefore, i_* is an isomorphism also.

Let us write the exact sequence for the pair (C_f, X):

$$\cdots \to \pi_n(X) \xrightarrow{i_*} \pi_n(C_f) \xrightarrow{p_*} \pi_n(C_f, X) \xrightarrow{\partial_*} \pi_{n-1}(X) \xrightarrow{i_*} \pi_{n-1}(C_f) \to \cdots .$$

We have $\pi_n(C_f, X) = 0$. Indeed, $\operatorname{Ker} p_* = \operatorname{Im} i_* = \pi_n(C_f)$, whence $p_* = 0$. Moreover, $\operatorname{Im} \partial_* = \operatorname{Ker} i_* = 0$; therefore, $\partial_* = 0$, i.e., $\pi_n(C_f, X) = \operatorname{Ker} \partial_* = \operatorname{Im} p_* = 0$.

According to Theorem 4.20, the space X is a deformation retract of C_f, i.e., the identity map $C_f \to C_f$ is homotopic to $r\colon C_f \to X \subset C_f$ and the homotopy is fixed on X. For $f' = rj\colon Y \to X$, we have $f'f = rjgi \sim r(\mathrm{id}_{C_f})i = \mathrm{id}_X$ and $ff' = girj \sim g(\mathrm{id}_{C_f})j = \mathrm{id}_Y$. □

Manifolds

1. Definition and Basic Properties

A second countable topological space X is called a *topological manifold* if it is Hausdorff and each point $x \in X$ has an open neighborhood homeomorphic to an open subset of $\mathbb{R}^n$, the number n being the same for all points $x \in X$. (The example on p. 117 shows that there exist non-Hausdorff spaces in which every point has an open neighborhood homeomorphic to an open subset of Euclidean space.)

According to the Brouwer theorem on the invariance of dimension (Theorem 2.9 on p. 61), the dimension of a locally Euclidean space is determined uniquely. Indeed, suppose that $U \ni x$ and $V \ni x$ are open sets in X homeomorphic to open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, where $n \neq m$. Then the open set $U \cap V$ is homeomorphic to both an open subset of $\mathbb{R}^n$ and an open subset of $\mathbb{R}^m$, which contradicts the invariance of dimension.

Let M^n be a topological manifold of dimension n. A pair (U, φ), where $U \subset M^n$ is a connected open set and $\varphi \colon U \to \varphi(U)$ is a homeomorphism onto an open subset of $\mathbb{R}^n$, is called a *chart*, or a *local coordinate system*. If $\varphi(x) = 0$, then we say that (U, φ) is a *local coordinate system with origin* x.

A *smooth structure* on a topological manifold M is a family $\mathcal{A}$ of local coordinate systems $\{(U_\alpha, \varphi_\alpha) : \alpha \in A\}$ with the following properties:

(i) the sets U_α cover M;

(ii) if $U_\alpha \cap U_\beta \neq \varnothing$, then $\varphi_\alpha \varphi_\beta^{-1}$ is a smooth map;

(iii) the family $\mathcal{A}$ is maximal in the sense that if (U, φ) is a local coordinate system and the maps $\varphi_\alpha \varphi^{-1}$ and $\varphi \varphi_\alpha^{-1}$ are smooth for all α such that $U_\alpha \cap U \neq \varnothing$, then $(U, \varphi) \in \mathcal{A}$.

To define a smooth structure on M, it is sufficient to specify a family $\mathcal{A}_0$ of local coordinate systems having properties (i) and (ii). Indeed, property (iii) can be regarded as the definition of the set of local coordinate systems required for completing $\mathcal{A}_0$. A set of charts covering M and having property (ii) is referred to as an *atlas*.

A topological manifold M endowed with a smooth structure is called a *smooth manifold*, or simply a *manifold*. (In mathematics, nonsmooth manifolds are much less frequent than smooth ones.)

Example. The topological space $\mathbb{R}P^n$ admits a structure of a manifold.

Proof. The space $\mathbb{R}P^n$ can be covered by open sets $U_i = \{(x_1 : \dots : x_{n+1}) \in \mathbb{R}P^n\colon x_i \neq 0\}$, where $i = 1, \dots, n+1$. Consider the homeomorphisms $\varphi_i\colon U_i \to \mathbb{R}^n$ given by

$$\varphi_i(x_1 : \dots : x_{n+1}) = \left(\frac{x_1}{x_i}, \dots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \dots, \frac{x_{n+1}}{x_i}\right).$$

We must prove that the maps $\varphi_j\varphi_i^{-1}$, which are defined on $\varphi_i(U_i \cap U_j)$, are smooth. Without loss of generality, we can assume that $i = 1$ and $j = 2$. Then the set $U_i \cap U_j$ is determined by the inequality $x_1x_2 \neq 0$. Let

$$\varphi_1(x_1 : \dots : x_{n+1}) = \left(\frac{x_2}{x_1}, \frac{x_3}{x_1}, \dots, \frac{x_{n+1}}{x_1}\right) = (y_1, y_2, \dots, y_n).$$

Then

$$\begin{aligned}\varphi_2\varphi_1^{-1}(y_1, y_2, \dots, y_n) &= \varphi_2(x_1 : \dots : x_{n+1}) \\ &= \left(\frac{x_1}{x_2}, \frac{x_3}{x_2}, \dots, \frac{x_{n+1}}{x_2}\right) = \left(\frac{1}{y_1}, \frac{y_2}{y_1}, \dots, \frac{y_n}{y_1}\right).\end{aligned}$$

By assumption, $y_1 \neq 0$; therefore, the map $\varphi_2\varphi_1^{-1}$ is smooth. □

Exercise 31. Prove that the space $\mathbb{C}P^n$ is a manifold.

1.1. Manifolds with Boundary. The definition of a *manifold with boundary* is obtained from that of a manifold by allowing the images $\varphi(U)$ of the charts $\varphi\colon U \to \varphi(U)$ to be open subsets of the topological space $\mathbb{R}^n_+ = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1 \geq 0\}$.

To define a smooth structure on a manifold with boundary, we modify property (ii) as follows:

(ii) if $U_\alpha \cap U_\beta \neq \varnothing$, then there exist open (in $\mathbb{R}$) sets $V_\alpha \supset \varphi_\alpha(U_\alpha \cap U_\beta)$ and $V_\beta \supset \varphi_\beta(U_\alpha \cap U_\beta)$ and mutually inverse smooth maps $f_{\alpha\beta}\colon V_\beta \to V_\alpha$ and $f_{\beta\alpha}\colon V_\alpha \to V_\beta$ such that the restrictions of these maps to $\varphi_\beta(U_\alpha \cap U_\beta)$ and $\varphi_\alpha(U_\alpha \cap U_\beta)$ coincide with $\varphi_\alpha\varphi_\beta^{-1}$ and $\varphi_\beta\varphi_\alpha^{-1}$, respectively.

Let M^n be a manifold with boundary. We say that $x \in M^n$ is a *boundary point* if there exists a chart $\varphi\colon U \to \varphi(U) \subset \mathbb{R}^n_+$ such that $x \in U$ and

$$\varphi(x) \in \partial\mathbb{R}^n_+ = \{(x_1,\dots,x_n) \in \mathbb{R}^n : x_1 = 0\};$$

by a chart we mean a smooth chart, i.e., a chart from a smooth structure. The *boundary* of a manifold M^n is the set of all boundary points of M^n. We say that $x \in M^n$ is an *interior point* if there exists a chart $\varphi\colon U \to \varphi(U)$ such that $x \in U$ and $\varphi(U)$ is open in $\mathbb{R}^n$; we again assume the chart to be smooth. To show that an interior point of a manifold cannot be a boundary point, we need the following theorem.

Theorem 5.1 (Inverse function theorem). *Suppose that $f\colon U \to V$ is a smooth map of open subsets of $\mathbb{R}^n$ and its Jacobian matrix $\left(\frac{\partial f_i}{\partial x_j}\right)$ has nonzero determinant at some point $x \in U$. Then x has a neighborhood $\widetilde{U} \subset U$ such that the set $f(\widetilde{U})$ is open in $\mathbb{R}^n$ and the map $f \restriction \widetilde{U}\colon \widetilde{U} \to f(\widetilde{U})$ is a homeomorphism. Moreover, the map $(f \restriction \widetilde{U})^{-1}$ is smooth.*

Proof. We write f in the form $f = (f_1,\dots,f_n)$, where $f_1,\dots,f_n$ are functions on U. Suppose that B is a convex set in the domain of f and $y, z \in B$. On the interval $[0,1]$, consider the function $\varphi_i(t) \rightleftharpoons f_i(ty+(1-t)z)$. Clearly,

$$(1)\qquad f_i(y) - f_i(z) = \varphi_i(1) - \varphi_i(0) = \frac{d\varphi_i}{dt_i}(t_i) = \sum \frac{\partial f_i}{\partial x_j}(w_i)(y_i - z_i),$$

where $t_i \in (0,1)$ and $w_i = t_i y + (1-t_i)z \in B$.

Now, consider the function $J(w_1,\dots,w_n) = \det\left(\frac{\partial f_i}{\partial x_j}(w_i)\right)$ of n^2 real variables (the coordinates of $w_1,\dots,w_n \in U$). This function is continuous, and $J(x,\dots,x) \neq 0$. Hence there exists a number ε such that if a point $w_1,\dots,w_n$ belongs to the disk $D^n_{x,\varepsilon}$, then $J(w_1,\dots,w_n) \neq 0$. This inequality and formula (1) show that if $y, z \in D^n_{x,\varepsilon}$ and $y \neq z$, then $f(y) \neq f(z)$. Thus, the restriction of f to the compact set $D^n_{x,\varepsilon}$ is a one-to-one map to some subset of the Hausdorff space $\mathbb{R}^n$. Such a map is a homeomorphism.

Let $\widetilde{U} = \operatorname{int} D^n_{x,\varepsilon}$. We have already proved that $f \restriction \widetilde{U}\colon \widetilde{U} \to f(\widetilde{U})$ is a homeomorphism. Now, let us prove that the set $f(\widetilde{U})$ is open in $\mathbb{R}^n$. Take a point $u \in \widetilde{U}$. We must show that any open disk of sufficiently small radius centered at $f(u)$ is entirely contained in $f(\widetilde{U})$. On the compact set $\partial D^n_{x,\varepsilon}$, the function $\varphi(y) = \|f(y) - f(u)\|$ attains its minimum. This minimum is positive, because $f(u) \notin f(\partial D^n_{x,\varepsilon})$. Therefore, we can choose a positive number δ such that $\|f(y) - f(u)\| > 2\delta$ for all $y \in \partial D^n_{x,\varepsilon}$. Let us show that the open disk of radius δ centered at $f(u)$ is contained in $f(\widetilde{U})$. Suppose that $\|f(u) - z\| < \delta$. If $y \in D^n_{x,\varepsilon}$, then

$$\|f(y) - z\| \geq \|f(y) - f(u)\| - \|f(u) - z\| > 2\delta - \delta = \delta.$$

Therefore, on the set $D^n_{x,\varepsilon}$, the smooth function $\psi(w) = \|f(w) - z\|^2$ attains its minimum at an interior point (the values of this function on the boundary are larger than δ^2, and its value at u is smaller than δ^2). Let a be a point of minimum for ψ on the set $D^n_{x,\varepsilon}$. Then

$$0 = \frac{\partial \psi}{\partial x_j} = 2\sum \frac{\partial f_i}{\partial x_j}(a)(f_i(a) - z_i).$$

We already know that if $a \in D^n_{x,\varepsilon}$, then $J(a, \dots, a) = \det\left(\frac{\partial f_i}{\partial x_j}(a)\right) \neq 0$. Hence $z = f(a) \in f(\widetilde{U})$, as required.

It remains to prove the last assertion, namely, that the map $(f \restriction \widetilde{U})^{-1}$ is smooth. Since the determinant $\det\left(\frac{\partial f_i}{\partial x_j}(y)\right)$ is nonzero for $y \in \widetilde{U}$, it follows that f^{-1} has smooth partial derivatives on the open set $\widetilde{U}$. Therefore, f^{-1} is smooth on $\widetilde{U}$. □

Corollary. *A point of a manifold with boundary is an interior point if and only if it is not a boundary point.*

Proof. If M^n is a manifold with boundary and a point of M is both interior and boundary, then there exist smooth mutually inverse maps $f: U \to V$ and $f: V \to U$ of open subsets of $\mathbb{R}^n$ satisfying the following conditions:

(i) $0 \in U$, $0 \in V$, and $f(0) = g(0) = 0$;

(ii) U has an open subset $U' \ni 0$ such that $f(U') \subset \mathbb{R}^n_+$.

Since the map f has a smooth inverse in a neighborhood of 0, it follows that $\det\left(\frac{\partial f_i}{\partial x_j}(0)\right) \neq 0$. Therefore, the point 0 has a neighborhood $\widetilde{U} \subset U$ such that $f(\widetilde{U})$ is open in $\mathbb{R}^n$ and $f \restriction \widetilde{U}: \widetilde{U} \to f(\widetilde{U})$ is a homeomorphism. Thus, $f(\widetilde{U} \cap U')$ is an open neighborhood of 0 in $\mathbb{R}^n$. This contradicts the inclusions $f(\widetilde{U} \cap U') \subset f(U') \subset \mathbb{R}^n_+$. □

Remark 5.1. An even stronger assertion is valid: *An open neighborhood of* 0 *in* $\mathbb{R}^n_+$ *is never homeomorphic to an open subset of* $\mathbb{R}^n$. The proof of this assertion for $n = 2$ is contained in that of Theorem 4.1 on p. 140. In the general case, the proof uses the multidimensional analog of the Jordan theorem.

A compact manifold without boundary is said to be *closed.*

A subset $N \subset M^n$ is called a k-dimensional *submanifold* of M^n if for an arbitrary point $x \in N$, there exists a chart (U, φ) such that $U \cap N = \varphi^{-1}(\mathbb{R}^k \cap \varphi(U))$ (it is assumed that $\mathbb{R}^k$ is standardly embedded in $\mathbb{R}^n$, i.e., consists of the points whose last $n - k$ coordinates are zero).

Exercise 32. (a) Let M^n be a manifold without boundary. Prove that a submanifold of M^n is a manifold without boundary.

(b) Let M^n be a manifold with boundary. Prove that any of its submanifolds is either a manifold with boundary or a manifold without boundary.

Exercise 33. Prove that the boundary of a manifold is a manifold without boundary.

Problem 63. Let M^n be a connected manifold, and let N^n be a submanifold of M^n (of the same dimension). Prove that if N^n is closed, then $N^n = M^n$.

1.2. Maps of Manifolds. Let M^m and N^n be manifolds with smooth structures $\mathcal{A} = \{(U_\alpha, \varphi_\alpha)\}$ and $\mathcal{B} = \{(V_\beta, \psi_\beta)\}$. A map $f: M^m \to N^n$ is said to be *smooth* if all the maps $\psi_\beta f \varphi_\alpha^{-1}$ are smooth. Each map $\psi_\beta f \varphi_\alpha^{-1}$ is defined on the open set $\varphi_\alpha(U_\alpha \cap f^{-1}(V_\beta)) \subset \mathbb{R}^m$ and takes values in $\mathbb{R}^n$.

If $f: M^m \to N^n$ and $g: N^n \to M^m$ are smooth mutually inverse maps, then the map f is called a *diffeomorphism*, and the manifolds M^m and N^n are said to be *diffeomorphic*. It follows from the inverse function theorem that if manifolds M^m and N^n are diffeomorphic, then $m = n$.

Let $f: M^m \to N^n$ be a smooth map between manifolds without boundary, and let $x \in M^m$. Choose local coordinates near x and $f(x)$ and consider the Jacobian matrix $\left(\frac{\partial f_i}{\partial x_j}(x)\right)$. The rank of this matrix does not depend on the local coordinate systems. This rank is called the *rank* of the map f at the point x and denoted by $\operatorname{rank} f(x)$.

Let $f: M^m \to N^n$ be a smooth map. If $\operatorname{rank} f(x) = m$, then the map f is called an *immersion* at the point x, and if $\operatorname{rank} f(x) = n$, then f is called a *submersion* at x. A map f that is an immersion (submersion) at all points $x \in M^m$ is called an immersion (submersion). An immersion f that is a homeomorphism of M^m onto $f(M^m) \subset N^n$ is an *embedding*.

Theorem 5.2. (a) *Any smooth map $f: M^m \to N^n$ that is an immersion at some point a has the structure of the standard embedding $\mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^{n-m}$ in a neighborhood of this point. To reduce it locally to such an embedding, it suffices to choose a suitable coordinate system on the image, i.e., in a neighborhood of $f(a)$.*

(b) *Any smooth map $f: M^m \to N^n$ that is a submersion at some point a has the structure of the standard projection $\mathbb{R}^n \times \mathbb{R}^{m-n} \to \mathbb{R}^n$ in a neighborhood of this point. To reduce it to such a projection, it suffices to choose a suitable coordinate system on the preimage, i.e., in a neighborhood of a.*

Proof. Let (U_a, φ) and $(V_{f(a)}, \psi)$ be local coordinate systems with origins $a \in M^m$ and $f(a) \in N^n$, respectively; we assume that $f(U_a) \subset V_{f(a)}$. We write f in local coordinates, i.e., consider the map $\tilde{f} = \psi f \varphi^{-1}$. For this map, the Jacobian matrix $J = \left(\frac{\partial \tilde{f}_i}{\partial x_j}(x)\right)$ is defined. If f is an immersion, then the rank of J at the origin equals m, and if f is a submersion, then the rank is n. Therefore, we can choose $i_1, \ldots, i_m \in \{1, \ldots, n\}$ (if f is an immersion) or $j_1, \dots, j_n \in \{1, \dots, m\}$ (if f is a submersion) such that at the origin,

$$\det \begin{pmatrix} \frac{\partial \tilde f_{i_1}}{\partial x_1} & \cdots & \frac{\partial \tilde f_{i_1}}{\partial x_m} \\ \cdots & \cdots & \cdots \\ \frac{\partial \tilde f_{i_m}}{\partial x_1} & \cdots & \frac{\partial \tilde f_{i_m}}{\partial x_m} \end{pmatrix} \neq 0 \quad \text{or} \quad \det \begin{pmatrix} \frac{\partial \tilde f_1}{\partial x_{j_1}} & \cdots & \frac{\partial \tilde f_1}{\partial x_{j_n}} \\ \cdots & \cdots & \cdots \\ \frac{\partial \tilde f_n}{\partial x_{j_1}} & \cdots & \frac{\partial \tilde f_n}{\partial x_{j_n}} \end{pmatrix} \neq 0,$$

respectively. Without loss of generality, we can assume that $i_k = k$ (respectively, $j_k = k$).

The Jacobian matrix J can be augmented to a square matrix of the form $\begin{pmatrix} J_1 & J_2 \\ 0 & I \end{pmatrix}$ or $\begin{pmatrix} J_1 & 0 \\ J_2 & I \end{pmatrix}$, where I is the identity matrix of order $|n-m|$. The augmented matrix is the Jacobian matrix for some map $\widetilde F$, namely, for the map

$$\widetilde F(x_1, \dots, x_m, y_1, \dots, y_{n-m}) \\ = (\tilde f_1(x), \dots, \tilde f_m(x), \tilde f_{m+1}(x) + y_1, \dots, \tilde f_n(x) + y_{n-m})$$

if f is an immersion, and for the map

$$\widetilde F(x_1, \dots, x_m, y_1, \dots, y_{m-n}) = (\tilde f(x,y), y_1, \dots, y_{m-n})$$

if f is a submersion; we write $\widetilde F(x,y) = \tilde f(x) + (0,y)$ and $\widetilde F(x,y) = (\tilde f(x,y), y)$ respectively.

At the origin, the Jacobian determinant of the map $\widetilde F$ is nonzero; therefore, according to the inverse function theorem, $\widetilde F$ has a smooth inverse $\widetilde F^{-1}$ in some neighborhood of the origin.

In the case of immersion, we replace the map ψ by $\widetilde F^{-1}\psi$ (in other words, we change the local coordinate system on the image of f). In the new local coordinates, the map f has the form

$$x \mapsto \psi f \varphi^{-1} \tilde f(x) = \widetilde F(x,0) \mapsto \widetilde F^{-1}(x,0).$$

In the submersion case, we replace the map φ by $\widetilde F\varphi$ (in other words, we change the local coordinate system on the preimage of f). In the new local coordinates, f has the form

$$(\tilde f(x,y), y) \mapsto \widetilde F^{-1}(x,y) \mapsto \psi f \varphi^{-1} \tilde f(x,y),$$

i.e., $(x', y) \mapsto x'$, where $x' = \tilde f(x,y)$. It remains to note that for y fixed, the map $x \mapsto \tilde f(x,y)$ is an epimorphism in a small neighborhood of the origin, because the Jacobian determinant of this map is nonzero. □

Using Theorem 5.2, we can easily prove the following assertion.

Theorem 5.3. *Suppose that $f\colon M^m \to N^n$ is a smooth map, X^k is a submanifold of N^n, and f is a submersion at each point of $f^{-1}(X^k)$. Then $f^{-1}(X^k)$ is a submanifold of dimension $k+m-n$ in M^m.*

Proof. Let $a \in f^{-1}(X^k)$. In a neighborhood of $f(a)$, we can choose local coordinates $x = (u, v)$ so that $u \in \mathbb{R}^k$, $v \in \mathbb{R}^{n-k}$, and the equation $v = 0$ determines X^k in this neighborhood.

The map f is a submersion at a; therefore, by Theorem 5.2, we can choose local coordinates (x, y) in a neighborhood of a so that $x \in \mathbb{R}^n$, $y \in \mathbb{R}^{m-n}$, and the map f has the form $(x, y) \mapsto x$ (i.e., $(u, v, y) \mapsto (u, v)$) in the chosen local coordinates near a and $f(a)$. Clearly, the equation of the set $f^{-1}(X^k)$ in these coordinates is $v = 0$. Therefore, $f^{-1}(X^k)$ is a submanifold of dimension $k + m - n$. $\square$

Theorem 5.3 can be used to prove that certain subsets of manifolds are submanifolds and, thereby, manifolds.

Example. The sphere S^n is a manifold.

Proof. Consider the map $f\colon \mathbb{R}^{n+1} \to \mathbb{R}^1$ defined by $f(x_1, \dots, x_{n+1}) = x_1^2 + \dots + x_{n+1}^2$. The Jacobian matrix $J = (2x_1, \dots, 2x_{n+1})$ has rank 1 at each point of $f^{-1}(1)$; therefore, the set $f^{-1}(1) = S^n$ is a submanifold in $\mathbb{R}^{n+1}$. $\square$

1.3. Smooth Partitions of Unity. Let $\{U_\alpha, \alpha \in \mathcal{A}\}$ be an open cover of a manifold M^n. A partition of unity $\{\varphi_\beta, \beta \in \mathcal{B}\}$ subordinate to this cover is said to be *smooth* if all the functions φ_β are smooth.

Theorem 5.4. (a) *For an open cover $\{U_\alpha, \alpha \in \mathcal{A}\}$ of a manifold M^n, there exists a smooth partition of unity $\{\varphi_\beta, \beta \in \mathcal{B}\}$ subordinate to it.*

(b) *If the index set $\mathcal{A}$ is no more than countable, then it can be assumed that $\mathcal{B} = \mathcal{A}$ and* $\operatorname{supp} \varphi_\alpha \subset U_\alpha$ *for each $\alpha \in \mathcal{A}$.*

Proof. (a) First, let us construct open sets $X_k \subset M^n$ $(k = 1, 2, \dots)$ such that their closures $\overline{X}_k$ are compact, $\overline{X}_k \subset X_{k+1}$, and $M^n = \bigcup_{k=1}^\infty X_k$. To this end, consider a countable base of the space M^n and take all open sets with compact closures in this base. We denote the chosen sets by $W_1, W_2, \dots$. They cover the manifold M^n. Indeed, any point $x \in M^n$ has a neighborhood $U(x)$ with compact closure. The set $U(x)$ can be represented as a union of elements of the base; clearly, all these elements have compact closures. Therefore, $x \in W_i$ for some i.

We put $X_1 = W_1$. The compact set $\overline{X}_1$ is covered by the open sets W_i; hence $\overline{X}_1 \subset W_{i_1} \cup \dots \cup W_{i_p}$, where $i_1 < i_2 < \dots < i_p$. We put $X_2 = W_1 \cup W_2 \cup W_{i_1} \cup \dots \cup W_{i_p}$. The sets $X_3, X_4, \dots$ are defined similarly.

Now, let us construct open sets $V_{\beta,1} \subset V_{\beta,3}$. We put $D_r^n = \{x \in \mathbb{R}^n : \|x\| < r\}$. For each point $z \in \overline{X}_k \setminus X_{k-1}$, we choose an open set $V_{z,3}$ such that $V_{z,3} \subset U_\alpha$ for some α, $V_{z,3} \subset X_{k+1}$, and $V_{z,3} \cap X_{k-2} = \varnothing$; in addition, we require that there exist a chart $\psi_x\colon V_{x,3} \to D_3^n$. The open sets $V_{x,1} = \psi_x^{-1}(D_1^n)$ cover the compact set $\overline{X}_k \subset X_{k+1}$; hence $\overline{X}_k$ is covered by finitely many sets of the form $V_{x,1}$. Let $\{V_{\beta,1}\}$ be the family of all such sets for all k; by $V_{\beta,3}$ we denote the sets of the form $V_{x,3}$ corresponding to the elements of this family. Note that the index set $\{\beta\}$ is at most countable; in addition, $\{V_{\beta,1}\}$ is a cover of M^n and $\{V_{\beta,3}\}$ is a locally finite refinement of $\{U_\alpha\}$.[1]

It is easy to verify that the function that takes the value $e^{-1/t}$ at $t > 0$ and 0 at $t \le 0$ is smooth. Therefore, the function

$$\gamma(t) = \begin{cases} e^{-1/(1-t)} & \text{if } t < 1, \\ 0 & \text{if } t \ge 1 \end{cases}$$

is smooth also. Thus, $\tilde{\gamma}(x) = \gamma(\|x\|^2/4)$ is a smooth function on $\mathbb{R}^n$; it takes positive values at all points of the open disk D_2^n and vanishes outside D_2^n.

Consider the functions

$$g_\beta(x) = \begin{cases} \tilde{\gamma}(\psi_\beta(x)) & \text{if } x \in V_{\beta,3}, \\ 0 & \text{if } x \notin V_{\beta,3} \end{cases}$$

and $h(x) = \sum_\beta g_\beta(x)$. The function h is smooth, because the cover $\{V_{\beta,3}\}$ is locally finite. The sets $V_{\beta,1}$ cover the entire manifold M^n, and $g_\beta(x) > 0$ for $x \in V_{\beta,1}$. Therefore, $h(x) > 0$ for any point $x \in M^n$. The functions $\varphi_\beta = g_\beta/h$ form the required partition of unity, because $\operatorname{supp}\varphi_\beta \subset V_{\beta,3} \subset U_\alpha$.

(b) Consider the open cover $U_1, U_2, \ldots$. We have constructed the partition of unity $\varphi_1, \varphi_2, \ldots$ such that for each i, $\operatorname{supp}\varphi_i \subset U_j$ for some $j = j(i)$. Let us define $\tilde{\varphi}_i$ as the sum of all functions φ_k such that $\operatorname{supp}\varphi_k \subset U_i$ and $\operatorname{supp}\varphi_k \not\subset U_j$ for $j < i$. Then each function φ_k is contained as a term in precisely one $\tilde{\varphi}_i$, and $\operatorname{supp}\tilde{\varphi}_i \subset U_i$. □

1.4. Sard's Theorem. To state Sard's theorem, we need the notion of a set of measure zero. We say that a set $X \subset \mathbb{R}^n$ $(n \ge 1)$ is of *measure zero* if, for any $\varepsilon > 0$, X can be covered by countably many cubes such that the sum of their volumes is less than ε. The cubes may be open or closed; instead of cubes, disks or parallelepipeds can be taken. It is assumed that in $\mathbb{R}^0$, only the empty set has measure zero.

Let M^n be a manifold. We say that a set $X \subset M^n$ is of *measure zero* if the manifold M^n can be covered by a countable family of charts $\varphi_i\colon U_i \to \mathbb{R}^n$, where each $\varphi_i(X \cap U_i) \subset \mathbb{R}^n$ has measure zero. The consistency of the two definitions is implied by the following two lemmas.

Lemma 1. *Any countable union of sets of measure zero in $\mathbb{R}^n$ is a set of measure zero.*

[1] Each point $x \in X_{k-1}$ belongs only to those sets $V_{\beta,3}$ that are contained in X_{k+1}. The number of such sets is finite by construction.

Proof. Suppose that $X = \bigcup_{i=1}^{\infty} X_i \subset \mathbb{R}^n$ and each set X_i has measure zero. Covering every X_i by a countable set of cubes such that the sum of their volumes is less than $\varepsilon/2^i$, we obtain a cover of X by a countable set of cubes for which the sum of volumes is less than $\varepsilon(1/2 + 1/4 + 1/8 + \cdots) = \varepsilon$. □

Lemma 2. *If $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is a smooth map and $X \subset \mathbb{R}^n$ is a set of measure zero, then $f(X) \subset \mathbb{R}^n$ is a set of measure zero.*

Proof. The space $\mathbb{R}^n$ can be represented as a countable union of cubes; therefore, it is sufficient to consider the case where X is contained in a cube I^n. Recall that

$$\|f(u) - f(v)\| \leq \max_{x\in[u,v]} \left|\frac{\partial f_i}{x_j}(x)\right| \|u - v\|$$

(this follows from equality (1) on p. 183). Let $K = \max_{x\in I^n} \left|\frac{\partial f_i}{x_j}(x)\right|$. Then $\|f(u) - f(v)\| \leq K\|u - v\|$ whenever $u, v \in I^n$. This means that f takes any disk of radius r contained in I^n to a disk of radius Kr. Thus, if a set X is covered by disks such that the sum of their volumes is less than ε, then $f(X)$ can be covered by disks such that the sum of their volumes is less than $K^n\varepsilon$. □

Exercise 34. (a) Suppose that $m < n$ and $M^m \subset N^n$ is a submanifold. Prove that the set $M^m \subset N^n$ has measure zero.

(b) Suppose that $m < n$ and $f\colon M^m \to N^n$ is a smooth map. Prove that the set $f(M^m) \subset N^n$ has measure zero.

The proof of Sard's theorem uses the following assertion.

Theorem 5.5 (Fubini). *If $C \subset \mathbb{R}^n$ is a compact set and any section of C by a hyperplane of the form $x_n = a$ has measure zero, then C has measure zero.*

Proof. If $n = 1$, then $C = \varnothing$ by definition. Suppose that $n \geq 2$. Since C is compact, we have $C \subset \mathbb{R}^{n-1} \times [a, b]$. For $t \in [a, b]$, we define sets $C_t \subset \mathbb{R}^{n-1}$ as follows:

$$C_t \times \{t\} = C \cap (\mathbb{R}^{n-1} \times \{t\}) = C \cap \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_n = t\}.$$

By assumption, each set $C_t \subset \mathbb{R}^{n-1}$ has measure zero. Take $\varepsilon > 0$ and consider a cover of C_t by open $(n-1)$-cubes $I_{t,1}^{n-1}, I_{t,2}^{n-1}, \dots$ such that the sum of their volumes is less than ε. The set $J_t^{n-1} = \bigcup_{j=1}^{\infty} I_{t,j}^{n-1}$ is open; therefore, its complement $C \setminus (J_t^{n-1} \times [a, b])$ is closed and hence compact. Thus, we can choose $\delta = \delta(t) > 0$ such that the set C_τ is covered by $J_t^{n-1} \times I_t$ for every $\tau \in (t - \delta, t + \delta) = I_t$. Indeed, the function $|x_n - t|$ attains its minimum δ on $C \setminus (J_t^{n-1} \times I_t)$; this minimum is positive, because $|x_n - t|$ can vanish only at those points of C that belong to C_t. The number δ has the required property, because if $\tau \in I_t$ and $x \in C_\tau$, then $|x_n - \tau| < \delta$, i.e., $x \notin C \setminus (J_t^{n-1} \times [a, b])$ and

$$x \in (J_t^{n-1} \times [a, b]) \cap C_\tau \subset J_t^{n-1} \times I_t.$$

The open sets I_t form a cover of the interval $[a, b]$. We take a finite subcover $I_{t_1}, \dots, I_{t_k}$ of this cover and replace each open set $I_{t_i} = (\alpha, \beta)$ by the interval $I_i = [\alpha, \beta] \cap [a, b]$. The cubes $I_{t_i,j}^{n-1} \times I_i$ cover C, and the sum of their volumes does not exceed εL, where L is the sum of the lengths of the intervals I_i.

Lemma 3. *Any finite cover of an interval $I = [a, b]$ by intervals $I_i = [a_i, b_i] \subset I$ has a subcover such that the sum of lengths of the intervals from this subcover is less than $2(b - a)$.*

Proof. We can assume that the cover $I_1, \dots, I_n$ is minimal (that is, it ceases to be a cover of I after any interval I_i is removed). Let us order the intervals $I_1, \dots, I_n$ in such a way that $a_1 \le a_2 \le \dots \le a_n$. Since the cover is minimal, we have $b_1 \le b_2 \le \dots \le b_n$. These inequalities and the minimality of the cover imply $b_k < a_{k+2}$ (otherwise, the intervals $[a_k, b_k]$ and $[a_{k+2}, b_{k+2}]$ would cover $[a_{k+1}, b_{k+1}]$). Therefore, the sum of the lengths of $I_1, I_3, \dots$ and the sum of the lengths of $I_2, I_4, \dots$ are both less than $b - a$. □

According to Lemma 3, the cover $I_1, \dots, I_k$ of the interval $[a, b]$ can be chosen so that $L < 2(b - a)$. This completes the proof of the Fubini theorem. □

Let $f \colon M^m \to N^n$ be a smooth map. A point $x \in M^m$ is said to be *critical* if $\operatorname{rank} f(x) < \min(m, n)$. Otherwise this point is said to be *regular*. A point $y \in N^n$ is called a *critical value* if $y = f(x)$ for some critical point x.

Theorem 5.6 (Sard [117]). *The set of critical values of any smooth map of manifolds has measure zero.*

Proof. It is sufficient to prove the theorem for a smooth map $f \colon U \to \mathbb{R}^n$, where $U \subset \mathbb{R}^m$ is an open set. We prove it by induction on m. For $m = 0$, the assertion is obvious: if $n = 0$, then the set of critical values is empty, and if $n \ge 1$, then this set is a one-point subset of $\mathbb{R}^n$ and, hence, has measure zero. Suppose that Sard's theorem is valid for all smooth maps $V \to \mathbb{R}^n$, where $V \subset \mathbb{R}^{m-1}$.

Let C be the set of critical points of a map $f \colon U \to \mathbb{R}^n$, and let C_i be the set of all points $x \in U$ at which all partial derivatives of f of order at most i vanish. Clearly, $C \supset C_1 \supset C_2 \supset \cdots$.

Step 1. The set $f(C \setminus C_1)$ has measure zero.

If $n = 1$, then $C_1 = C$. Suppose that $n \ge 2$. Let $c \in C \setminus C_1$. Then $\frac{\partial f_i}{\partial x_j}(c) \ne 0$ for some i and j. We can assume that $i = j = 1$. Consider the map $h \colon U \to \mathbb{R}^m$ defined by

$$h(x_1, \dots, x_m) = (f_1(x), x_2, \dots, x_m).$$

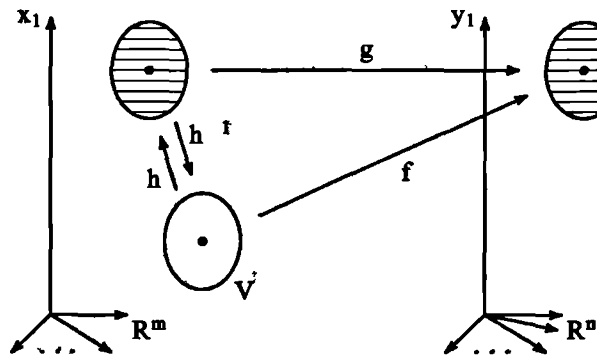

Figure 1. Construction of the map g

Clearly, the Jacobian matrix of h at c has the form $\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(c) & * \\ 0 & I \end{pmatrix}$, where I is the identity matrix. Thus, we can apply the inverse function theorem to the map h at the point c. According to this theorem, c has a neighborhood V such that the restriction of h to V is a homeomorphism (see Figure 1). For the map $g = fh^{-1}$, we have

$$g(x_1, \ldots, x_m) = (x_1, g_2(x), \ldots, g_n(x)).$$

In particular, g takes the set $(\{t\} \times \mathbb{R}^{m-1}) \cap h(V)$ to $\{t\} \times \mathbb{R}^{n-1}$. Let g^t be the restriction of g to this set. It is easy to show that a point of $(\{t\} \times \mathbb{R}^{m-1}) \cap h(V)$ is critical for the map g^t if and only if it is critical for g. Indeed,

$$\left(\frac{\partial g_i}{\partial x_j}\right) = \begin{pmatrix} 1 & 0 \\ * & \left(\frac{\partial g_i^t}{\partial x_j}\right) \end{pmatrix}.$$

By the induction hypothesis, the set of critical points of the map g^t has measure zero in $\{t\} \times \mathbb{R}^{n-1}$. This suggests that applying the Fubini theorem, it is possible to show that the set $h(V \cap C)$ has measure zero (and, hence, the image of $V \cap C$ under f has measure zero). The Fubini theorem does not apply directly because the set $f(V \cap C)$ is not compact. But $V \cap C$ can be represented as a countable union of compact sets; therefore, $f(V \cap C)$ can also be represented as a countable union of compact sets. To each of these sets, we can apply the Fubini theorem and obtain the desired result.

Thus, each point $c \in C \setminus C_1$ has a neighborhood V such that the set $f((C \setminus C_1) \cap V) \subset f(C \cap V)$ has measure zero. Since $C \setminus C_1$ can be covered by countably many such neighborhoods (instead of V we can take an open disk of rational radius centered at a point with rational coordinates such that it contains c and is contained in V), it follows that the set $f(C \setminus C_1)$ is of measure zero.

Step 2. The set $f(C_k \setminus C_{k-1})$ has measure zero for any $k \geq 1$.

The proof of this assertion is similar to that of Step 1. Let $c \in C_k \backslash C_{k-1}$. Then $\frac{\partial^{k+1} f_i}{\partial x_{j_1} \cdots \partial x_{j_{k+1}}}(c) \neq 0$ for some $i, j_1, \ldots, j_{k+1}$. We can assume that $i = j_1 = 1$. Let $w(x) = \frac{\partial^k f_1}{\partial x_{j_2} \cdots \partial x_{j_{k+1}}}(x)$. Then $w(c) = 0$, because $c \in C_k$, and

$$\frac{\partial w}{\partial x_1}(c) = \frac{\partial^{k+1} f_i}{\partial x_{j_1} \cdots \partial x_{j_{k+1}}}(c) \neq 0.$$

Consider the map $h\colon U \to \mathbb{R}^m$ defined by the formula

$$h(x_1, \ldots, x_m) = (w(x), x_2, \ldots, x_m).$$

The rest of the proof is a repetition of that of Step 1. The only difference is that it uses the fact that

$$h\left((C_k \setminus C_{k+1}) \cap V\right) \cap (\{t\} \times \mathbb{R}^{m-1})$$

is a critical point for the map g and, hence, for the map g^t.

Step 3. The set $f(C_k)$ has measure zero for sufficiently large k (e.g., for $k > (m/n) - 1$).

It suffices to consider the case where $U = (0,1)^m$ and the smooth map f is defined in some neighborhood of the cube $[0,1]^m$, because any open set U can be covered by countably many open cubes with this property.

Let $c \in C_k \cap U$. Then the Taylor expansion of $f(c+h)$ has no terms of order lower than $k+1$. Hence there exists a constant K such that $\|f(c+h) - f(c)\| \leq K\|h\|^{k+1}$ whenever $c + h \in U$.

We divide the open cube $U = (0,1)^m$ into l^m cubes with side length $1/l$ and consider the small cubes intersecting C_k. The image of any such cube under f is contained in a disk of radius Kd^{k+1}, where $d = \sqrt{n}/l$ is the maximum distance between points in the cube. Thus, the set $f(U \cap C_k)$ can be covered by disks such that the sum of their volumes is at most $K' l^m (d^{k+1})^n = K'' l^{m-(k+1)n}$. If $m < (k+1)n$, i.e., $k > (m/n) - 1$, then $\lim_{l \to \infty} l^{m-(k+1)n} = 0$. □

1.5. An Important Example: Grassmann Manifolds. Consider the set $G(n,k)$ whose elements are the k-dimensional subspaces of $\mathbb{R}^n$. A topology on this set is defined as follows. Let $v_1, \ldots, v_k \in \mathbb{R}^n$ be linearly independent vectors. The k-tuples $(v_1, \ldots, v_k)$ form an open subset X in $\mathbb{R}^n \times \cdots \times \mathbb{R}^n = \mathbb{R}^{nk}$. On the set X, we take the natural topology. The set $G(n,k)$ is a quotient of X; it is obtained by identifying k-tuples of vectors which generate the same subspaces. We endow $G(n,k)$ with the quotient topology. In other words, a set $U \subset G(n,k)$ is open if and only if the set of all bases of k-dimensional subspaces contained in U is open in X.

The set $G(n,k)$ can also be described as the quotient of X under the natural action of the group $\mathrm{GL}_k(\mathbb{R})$, which is defined as follows. We represent each k-tuple $(v_1,\dots,v_k)$ as the matrix $\begin{pmatrix} v_{11} & \dots & v_{1n} \\ \dots & \dots & \dots \\ v_{k1} & \dots & v_{kn}\end{pmatrix}$ whose rows are the coordinates of the vectors v_i. The matrix $A \in \mathrm{GL}_k(\mathbb{R})$ takes this k-tuple to the k-tuple $(w_1,\dots,w_k)$ of vectors whose coordinates form the rows of the matrix $W = AV$.

It is easy to verify that the same topology on $G(n,k)$ is obtained by considering only orthonormal sets of vectors rather than all linearly independent sets. Indeed, the map which takes every k-tuple of linearly independent vectors to the k-tuple obtained by applying the Gram Schmidt orthogonalization procedure is continuous.

This approach to the definition of the topology on $G(n,k)$ has the advantage that the space $G(n,k)$ is obtained as the quotient of a compact space under the action of the compact group $\mathrm{O}(k)$. This implies, in particular, that the space $G(n,k)$ is compact and Hausdorff.

Another proof of the Hausdorffness of $G(n,k)$ is as follows. Take a point $x \in \mathbb{R}^n$ and consider the function d_x on $G(n,k)$ equal to the distance from x to a subspace $\Pi \in G(n,k)$. This function is continuous. Clearly, if k-subspaces Π_1 and Π_2 are different, then the point x can be chosen so that $x \in \Pi_1$ and $x \notin \Pi_2$. In this case, we have $d_x(\Pi_1) = 0$ and $d_x(\Pi_2) \neq 0$.

Theorem 5.7. *The topological space $G(n,k)$ is a manifold (without boundary) of dimension $k(n-k)$.*

Proof. Take linearly independent vectors $v_1,\dots,v_k$ in a subspace $\Pi \in G(n,k)$ and consider the rectangular matrix $V = V(\Pi)$ whose rows are the coordinates of these vectors. For each multi-index $I = \{i_1,\dots,i_k\}$, where $1 \le i_1 < \dots < i_k \le n$, let V_I be the square matrix formed by the columns of V with numbers $i_1,\dots,i_k$. Since the vectors $v_1,\dots,v_k$ are linearly independent, it follows that there exists a multi-index I for which $\det V_I \neq 0$. Consider the rectangular matrix $(V_I)^{-1}V$. Its columns with numbers $i_1,\dots,i_k$ form the identity matrix of order k.

If $w_1,\dots,w_k$ are some other linearly independent vectors in Π, then $W = AV$, where $A \in \mathrm{GL}_k(\mathbb{R})$ and $W_I = AV_I$. Therefore,

$$(W_I)^{-1}W = (V_I)^{-1}A^{-1}AV = (V_I)^{-1}V.$$

This means that the rectangular matrix $(V_I)^{-1}V$ depends only on the subspace Π and the multi-index I; we denote this matrix by Π^I. For any multi-index J, the square matrix Π^I_J is defined. As mentioned before, Π^I_I is the identity matrix. The elements of the other columns of Π^I can be arbitrary.

For every multi-index I, let $U_I \subset G(n,k)$ be the set of those subspaces $\Pi \in G(n,k)$ for which $\det V_I \neq 0$. The sets U_I cover $G(n,k)$. Clearly, they are open. Indeed, if $\det V_I \neq 0$, then $\det V'_I \neq 0$ for any matrix V' whose elements are sufficiently close to those of V.

For each subspace $\Pi \subset U_I$, we construct the matrix Π^I and take the $n-k$ of its columns whose numbers are not included in the multi-index I. As a result, we obtain a homeomorphism $\varphi_I : U_I \to \mathbb{R}^{k(n-k)}$. It remains to verify that if I and J are two multi-indices, then the map $\varphi_J \varphi_I^{-1}$, which is defined on the open set $\varphi_I(U_I \cap U_J)$, is smooth.

The map $\varphi_J \varphi_I^{-1}$ can be described as follows. To each point $x \in \mathbb{R}^{k(n-k)}$ we assign the matrix Π^I whose $n-k$ columns are filled with the coordinates of x and the remaining columns (with numbers contained in the multi-index I) form the identity matrix. Then, we take a matrix V for which $(V_I)^{-1}V = \Pi^I$. The final result is $\Pi^J = (V_J)^{-1}V = (V_J)^{-1}V_I\Pi^I$. To remove the ambiguity involved in the choice of V, we assume that V_I is the identity matrix. In this case, to $X = \Pi^I$ we assign the matrix $(X_J)^{-1}X$ and take its columns with numbers not included in the multi-index J. By assumption, the matrix X_J is nonsingular on the entire domain; therefore, the resulting map is smooth. □

Exercise 35. Prove that the correspondence between subspaces and their orthogonal complements induces a diffeomorphism $G(n,k) \longleftrightarrow G(n,n-k)$.

Exercise 36. Prove that $G(n,1) \approx \mathbb{R}P^{n-1}$.

Choose a basis in the space $\mathbb{R}^n$. To each k-dimensional subspace Π we assign $\binom{n}{k}$ numbers x_I, called the *Plücker coordinates*, as follows. Take linearly independent vectors $v_1, \dots, v_k$ in Π and consider the matrix V whose rows are the coordinates of these vectors. For each multi-index I, we set $x_I = \det V_I$. There are $\binom{n}{k}$ multi-indices, and hence we obtain $\binom{n}{k}$ numbers.

The Plücker coordinates are determined uniquely up to proportionality. Indeed, when we take a different basis in Π, the matrix V is replaced by AV, where $A \in \mathrm{GL}_k(\mathbb{R})$, and each Plücker coordinate is multiplied by $\det A$.

Theorem 5.8 (Plücker embedding). *The Plücker coordinates determine an embedding* $i\colon G(n,k) \to \mathbb{R}P^{\binom{n}{k}-1}$.

Proof. First, we show that i is a homeomorphism of $G(n,k)$ onto its image $i(G(n,k))$. For this purpose, it suffices to verify that the map i is injective, because any one-to-one continuous map from a compact space to a Hausdorff space is a homeomorphism.

Suppose that $v_1, \dots, v_k \in \mathbb{R}^n$ are variable linearly independent vectors and $c_1, \dots, c_{n-k} \in \mathbb{R}^n$ are constant linearly independent vectors. Consider the square matrix $\binom{V}{C}$ whose rows are the coordinates of these vectors. By the Laplace theorem, we have

$$\det\begin{pmatrix} V \\ C \end{pmatrix} = \sum a_I \det \hat{V}_I,$$

where a_I is a constant depending on the matrix C. Clearly, $\det\binom{V}{C} = 0$ if and only if the intersection of the subspaces spanned by $v_1, \dots, v_k$ and $c_1, \dots, c_{n-k}$ is nontrivial. Thus, the Plücker coordinates of a k-dimensional subspace Π satisfy the equation $\sum a_I x_I = 0$ if and only if the intersection of Π with the subspace spanned by $c_1, \dots, c_{n-k}$ is nontrivial. It remains to note that for two different k-dimensional subspaces in $\mathbb{R}^n$, there exists an $(n-k)$-dimensional subspace such that its intersection with one of the subspaces is trivial and the intersection with the other is nontrivial.

Now, we show that i is an immersion. Take a chart U_I. On this chart, we introduce coordinates as explained above. To avoid cumbersome notation, we consider only the simple example of U_I consisting of matrices Π^I of the form $\left(\begin{smallmatrix} 1 & 0 & 0 & x_1 & x_4 \\ 0 & 1 & 0 & x_2 & x_5 \\ 0 & 0 & 1 & x_3 & x_6 \end{smallmatrix}\right)$. The image of a matrix Π^I under i is the set of all determinants of its submatrices of order $k = 3$. Considering only the submatrices in which $k-1$ columns coincide with those of the identity matrix, we see that in the chart U_I, the map i has the form

$$\begin{pmatrix} 1 & 0 & 0 & x_1 & x_4 \\ 0 & 1 & 0 & x_2 & x_5 \\ 0 & 0 & 1 & x_3 & x_6 \end{pmatrix} \mapsto (1, \pm x_1, \pm x_2, \pm x_3, \pm x_4, \pm x_5, \pm x_6, \dots).$$

One of the coordinates of the image is 1. This means that the image of the chart U_I is contained in a standard chart of the projective space, i.e., locally, i is a map of Euclidean spaces. Obviously, the rank of this map is equal to the dimension of the manifold $G(n,k)$. $\square$

Remark 5.2. The image of $G(n,k)$ under the Plücker embedding into $\mathbb{R}P^{\binom{n}{k}-1}$ can be determined explicitly by a system of equations called the *Plücker relations*; see [**100**, Section 30] for details.

By analogy with the manifold of k-subspaces in $\mathbb{R}^n$, we can consider the manifold of complex k-subspaces in $\mathbb{C}^n$. To distinguish between these manifolds, we call them the *real* and *complex* Grassmann manifolds, respectively. Moreover, in the real case, we can consider the *oriented Grassmann manifold* $G_+(n,k)$, whose points are oriented k-dimensional subspaces. In this case, sets of vectors are considered equivalent if they not only span the same k-subspace but also determine the same orientation on this subspace.

Exercise 37. Prove that there exists a double covering $G_+(n,k) \to G(n,k)$.

Exercise 38. Prove that the Plücker coordinates determine an embedding $G_+(n,k) \to S^{\binom{n}{k}-1}$.

Problem 64. (a) Prove that the manifold $G_+(n,k)$ is always orientable.

(b) Prove that the real Grassmann manifold $G(n,k)$ is orientable if and only if n is even.

Problem 65. Prove that $G_+(4,2) \approx S^2 \times S^2$.

Problem 66. Prove that the quadric $z_1^2+\dots+z_n^2 = 0$ in $\mathbb{C}P^{n-1}$ is diffeomorphic to $G_+(n,2)$. Moreover, under the diffeomorphism, complex conjugation corresponds to reversing the orientation of the plane.

Now, we describe the cell structure of the Grassmann manifolds. For each k-dimensional subspace $\Pi \subset \mathbb{R}^n$, consider the sequence of $a_i = \dim(\Pi \cap \mathbb{R}^i)$, where $i = 0,1,\dots,n$. It is assumed that $\mathbb{R}^i$ consists of vectors of the form $(x_1,\dots,x_i,0,\dots,0)$. Clearly, $a_0 = 0$, a_{i+1} is equal to a_i or a_i+1, and $a_n = k$. Therefore, to the sequence a_i we can assign the *Schubert symbol* $\sigma = (\sigma_1,\dots,\sigma_k)$, where $1 \le \sigma_1 < \dots < \sigma_k \le n$, $\dim(\Pi \cap \mathbb{R}^{\sigma_j}) = j$, and $\dim(\Pi \cap \mathbb{R}^{\sigma_j-1}) = j-1$.

Lemma. *A subspace Π has Schubert symbol σ if and only if it contains vectors $v_1,\dots,v_k$ such that the matrix whose rows are formed by their coordinates has the form*

$$\begin{pmatrix}
* & \dots & * & 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots \\
* & \dots & * & 0 & * & \dots & * & 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots \\
* & \dots & * & 0 & * & \dots & * & 0 & 0 & \dots & * & 1 & 0 & \dots & 0 & 0 & 0 & \dots \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
* & \dots & * & 0 & * & \dots & * & 0 & 0 & \dots & * & 0 & * & \dots & * & 1 & 0 & \dots
\end{pmatrix};$$

the elements $\sigma_1,\dots,\sigma_k$ of the Schubert symbol are the numbers of the columns consisting of zeros and ones; the elements $*$ are arbitrary. Moreover, the vectors $v_1,\dots,v_k$ are determined uniquely.

Proof. The one-dimensional space $\Pi \cap \mathbb{R}^{\sigma_1}$ is generated by the vector $v_1 = (x_1,\dots,x_{\sigma_1},0,\dots,0)$. The condition $\dim(\Pi \cap \mathbb{R}^{\sigma_1-1}) = 0$ implies $x_{\sigma_1} \ne 0$. Therefore, we can assume that $x_{\sigma_1} = 1$; in this case, the vector v_1 is determined uniquely.

In the two-dimensional space $\Pi \cap \mathbb{R}^{\sigma_2}$, the vector v_1 can be extended to a basis by a vector $v_2 = (y_1,\dots,y_{\sigma_2},0,\dots,0)$. The condition $\dim(\Pi \cap \mathbb{R}^{\sigma_2-1}) = 1$ implies $y_{\sigma_2} \ne 0$. Therefore, we can assume that $y_{\sigma_2} = 1$. In this case, the vector v_2 is determined up to the addition of λv_1. The σ_1th coordinate of the vector $v_2 + \lambda v_1$ is $y_{\sigma_1} + \lambda$. We can kill it by choosing suitable λ. Now, the vector v_2 is determined uniquely. The rest of the proof is similar. □

The set of all subspaces with a given Schubert symbol σ is called an *open Schubert cell* and denoted by $e(\sigma)$. Any open Schubert cell $e(\sigma) \subset G(n,k)$

is homeomorphic to the open disk of dimension $d(\sigma) = (\sigma_1 - 1) + (\sigma_2 - 2) + \dots + (\sigma_k - k)$. The open Schubert cells are pairwise disjoint and cover the entire Grassmann manifold.

The distinguishing feature of any cell $e(\sigma)$ is that the subspaces which it contains have bases $v_1, \dots, v_k$ such that $v_i = (v_{i1}, \dots, v_{i\sigma_i}, 0, \dots, 0)$, where $v_{i\sigma_i} > 0$. Hence the distinguishing feature of the closure $\overline{e(\sigma)}$ of such a cell is that the subspaces which it contains have bases $v_1, \dots, v_k$ such that $v_i = (v_{i1}, \dots, v_{i\sigma_i}, 0, \dots, 0)$, where $v_{i\sigma_i} \geq 0$. Note that the bases $v_1, \dots, v_k$ can be assumed to be orthonormal.

Our immediate goal is to construct a continuous map $\chi_\sigma \colon D^{d(\sigma)} \to \overline{e(\sigma)} \subset G(n, k)$ with the following properties:

- the restriction of χ_σ to $\operatorname{int} D^{d(\sigma)}$ is a homeomorphism onto an open Schubert cell $e(\sigma)$;
- the set $\chi_\sigma(\partial D^{d(\sigma)})$ is contained in the union of open Schubert cells $e(\tau)$ for which $d(\tau) < d(\sigma)$.

We construct such a map by induction on k. If $k = 1$, then the Schubert symbol σ consists of one element σ_1, and $d(\sigma) = \sigma_1 - 1$. In this case, the set $\overline{e(\sigma)}$ consists of all one-dimensional subspaces generated by nonzero vectors of the form $(v_{11}, \dots, v_{1\sigma_1}, 0, \dots, 0)$, where $v_{1\sigma_1} \geq 0$, and we define the map $\chi_\sigma \colon D^{d(\sigma)} \to \overline{e(\sigma)}$ as follows. We identify the disk $D^{d(\sigma)}$ with the hemisphere $x_1^2 + \dots + x_{\sigma_1}^2 = 1$, where $x_{\sigma_1} \geq 0$, and define the image of each point $(x_1, \dots, x_{\sigma_1})$ to be the one-dimensional subspace spanned by the vector $(x_1, \dots, x_{\sigma_1}, 0, \dots, 0)$. Clearly, the restriction of the map χ_σ thus defined to $\operatorname{int} D^{d(\sigma)}$ is a homeomorphism onto $e(\sigma)$ and the set $\chi_\sigma(\partial D^{d(\sigma)})$ consists of the subspaces spanned by nonzero vectors of the form $(x_1, \dots, x_{\sigma_1 - 1}, 0, \dots, 0)$; multiplying such a vector by a nonzero number, we can reduce it to the form $(x_1, \dots, x_{\tau_1}, 0, \dots, 0)$, where $x_{\tau_1} = 1$ and $\tau_1 < \sigma_1$.

To construct the map χ_σ for $k \geq 2$, we need an auxiliary proper orthogonal transformation of $\mathbb{R}^n$ with positive determinant which takes a given unit vector u to another unit vector v and leaves all vectors orthogonal to u and v fixed. Such a transformation $R(u, v)$ exists if $u \neq -v$, and it is unique. Indeed, it is easy to verify that the transformation defined by

$$R(u, v)x = x - \frac{(u + v, x)}{1 + (u, v)}(u + v) + 2(u, x)v$$

has the required properties (on the plane spanned by u and v, it acts as a rotation because it takes u to v and v to the vector symmetric to u about v). Thus, the point $R(u, v)x$ continuously depends on u, v, and x. Clearly, if $u, v \in \Pi$, then the projections of the vectors x and $R(u, v)x$ on $\Pi^\perp$ coincide.

Suppose that the required map $\chi_\sigma \colon D^{d(\sigma)} \to \overline{e(\sigma)}$ is constructed for any

Schubert symbol $\sigma = (\sigma_1, \dots, \sigma_k)$ of length k. Consider a Schubert symbol $\sigma' = (\sigma_1, \dots, \sigma_k, \sigma_{k+1})$ of length $k+1$ (as usual, we assume that $1 \le \sigma_1 < \sigma_2 < \cdots < \sigma_k < \sigma_{k+1} \le n$). Instead of the map $\chi_{\sigma'}\colon D^{d(\sigma')} \to \overline{e(\sigma')}$, we construct a map (which is even a homeomorphism) $\varphi\colon \overline{e(\sigma)} \times D^{d(\sigma')-d(\sigma)} \to \overline{e(\sigma')}$. The map $\chi_{\sigma'}$ will be defined as the composition

$$D^{d(\sigma')} \approx D^{d(\sigma)} \times D^{d(\sigma')-d(\sigma)} \xrightarrow{\chi_\sigma \times \mathrm{id}} \overline{e(\sigma)} \times D^{d(\sigma')-d(\sigma)} \xrightarrow{\varphi} \overline{e(\sigma')};$$

here $d(\sigma') - d(\sigma) = \sigma_{k+1} - k - 1$.

In each subspace $\Pi \in \overline{e(\sigma)}$, we choose an orthonormal basis $v_1, \dots, v_k$ such that $v_i = (v_{i1}, \dots, v_{i\sigma_i}, 0, \dots, 0)$, where $v_{i\sigma_i} \geq 0$. Let $e_1, \dots, e_n$ be the canonical basis of $\mathbb{R}^n$. We set

$$R = R(e_{\sigma_k}, v_k) \circ \cdots \circ R(e_{\sigma_2}, v_2) \circ R(e_{\sigma_1}, v_1).$$

It is easy to verify that $R e_{\sigma_i} = v_i$ for all $i = 1, \dots, k$. Indeed, the transformations $R(e_{\sigma_1}, v_1), \dots, R(e_{\sigma_{i-1}}, v_{i-1})$ leave the vector e_{σ_i} fixed, because this vector is orthogonal to $v_1, \dots, v_{i-1}$ and $e_{\sigma_1}, \dots, e_{\sigma_{i-1}}$. The transformation $R(e_{\sigma_i}, v_i)$ takes e_{σ_i} to v_i. Finally, the transformations $R(e_{\sigma_{i+1}}, v_{i+1}), \dots, R(e_{\sigma_k}, v_k)$ leave the vector v_i fixed because it is orthogonal to $v_{i+1}, \dots, v_k$ and $e_{\sigma_{i+1}}, \dots, e_{\sigma_k}$.

We identify the disk $D^{\sigma_{k+1}-k-1}$ with the set of unit vectors $w = (w_1, \dots, w_{\sigma_{k+1}}, 0, \dots, 0)$ such that $w_{\sigma_{k+1}} \ge 0$ and $w_{\sigma_i} = 0$ for $i = 1, \dots, k$. We define the map $\varphi\colon \overline{e(\sigma)} \times D^{d(\sigma')-d(\sigma)} \to \overline{e(\sigma')}$ by

$$\varphi(v_1, \dots, v_k, w) = (v_1, \dots, v_k, Rw).$$

We must verify that the space spanned by the vectors $v_1, \dots, v_k, Rw$ indeed belongs to $\overline{e(\sigma')}$. The vectors w and Rw have equal projections on the orthogonal complement of $\mathbb{R}^{\sigma_k}$; therefore, $Rw = (*, \dots, *, w_{\sigma_k+1}, \dots, w_{\sigma_{k+1}}, 0, \dots, 0)$. The linear independence of $v_1, \dots, v_k, Rw$ follows from

$$(v_i, Rw) = (Re_{\sigma_i}, Rw) = (e_{\sigma_i}, w) = w_{\sigma_i} = 0$$

and $(Rw, Rw) = (w, w) = 1$.

The map φ is surjective: its inverse φ^{-1} is defined by the formula

$$\varphi^{-1}(v_1, \dots, v_k, v_{k+1}) = (v_1, \dots, v_k, R^{-1}v_{k+1}),$$

where the orthonormal basis $v_1, \dots, v_k, v_{k+1}$ is constructed in precisely the same way as for the k-dimensional subspace with a given Schubert symbol and R is the orthogonal transformation defined above. It is easy to verify that the vector $w = R^{-1}v_{k+1}$ has the required properties; namely,

$w = (w_1, \dots, w_{\sigma_{k+1}}, 0, \dots, 0)$, where $w_{\sigma_{k+1}} = v_{k+1,\sigma_{k+1}} \ge 0$;
$w_{\sigma_i} = (e_{\sigma_i}, w) = (Re_{\sigma_i}, Rw) = (v_i, v_{k+1}) = 0$ for $i = 1, \dots, k$; and
$(w, w) = (R^{-1}v_{k+1}, R^{-1}v_{k+1}) = (v_{k+1}, v_{k+1}) = 1$.

The map ϕ^{-1} is continuous and by the induction hypothesis, χ_σ is a homeomorphism between $\operatorname{int} D^{d(\sigma)}$ and $d(\sigma)$; therefore, $\chi_{\sigma'}$ is a homeomorphism between $\operatorname{int} D^{d(\sigma')}$ and $d(\sigma')$ because

$$\operatorname{int}(\overline{e(\sigma)} \times D^{d(\sigma')-d(\sigma)}) = e(\sigma) \times \operatorname{int} D^{d(\sigma')-d(\sigma)}.$$

2. Tangent Spaces

It is easy to define a tangent vector at a point $x \in M^n$ in a local coordinate system, but the passage to other coordinate systems involves difficulties. For this reason, there are several definitions of tangent vectors, which are used in different situations.

One of the most natural definitions is as follows. A *tangent vector* at a point $x \in M^n$ is an object to which a certain vector $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ is associated in every local coordinate system (U, φ) with origin x; in a different local coordinate system (V, ψ) with origin x, associated to the same tangent vector is the vector $w = (w_1, \dots, w_n)$ related to v by

$$w_i = \sum \frac{\partial(\psi\varphi^{-1})_i}{\partial x_j}(0) v_j. \tag{1}$$

In other words, w is the image of v under the action of the Jacobian matrix of the transition function $\psi\varphi^{-1}$. This definition is consistent because the Jacobian matrix of a composition of two maps is the product of the Jacobian matrices of these maps.

The major drawback of this definition is that it depends on the coordinate system. An invariant definition can be given in several ways.

Tangent vectors as classes of equivalent curves. To a vector $v \in \mathbb{R}^n$ we can assign the family of all smooth curves $\gamma\colon (-1,1) \to \mathbb{R}^n$ for which $\gamma(0) = 0$ and $\frac{d\gamma}{dt}(0) = v$. If (U, φ) is a local coordinate system with origin $x \in M$, then the curve $\gamma(t)$ determines the curve $\widetilde{\gamma} = \varphi^{-1}\gamma$ on the manifold M^n; moreover, $\widetilde{\gamma}(0) = x$. Thus, a tangent vector at a point $x \in M$ can be defined as an equivalence class of smooth curves $\widetilde{\gamma}\colon (-1,1) \to M^n$ for which $\widetilde{\gamma}(0) = x$. Curves $\widetilde{\gamma}_1$ and $\widetilde{\gamma}_2$ are considered equivalent if

$$\left.\frac{d(\varphi\widetilde{\gamma}_1(t))}{dt}\right|_{t=0} = \left.\frac{d(\varphi\widetilde{\gamma}_2(t))}{dt}\right|_{t=0}$$

in some coordinate system (U, φ) with origin x. If (V, ψ) is another coordinate system with origin x, then

$$\left.\frac{d(\psi\widetilde{\gamma}(t))_i}{dt}\right|_{t=0} = \left.\frac{d(\psi\varphi^{-1}\varphi\widetilde{\gamma}(t))_i}{dt}\right|_{t=0} = \sum_j \frac{\partial(\psi\varphi^{-1})_i}{\partial x_j}(0) \left.\frac{d(\varphi\widetilde{\gamma}(t))_j}{dt}\right|_{t=0}.$$

This implies, first of all, that the equivalence of curves does not depend on the choice of local coordinates. Second, we see that the coordinates $\frac{d(\psi\tilde{\gamma}(t))_i}{dt}\Big|_{t=0}$ of a tangent vector do indeed obey law (1) under a change of local coordinates.

Tangent vectors as differentiation operators. Suppose that (U, φ) is a local coordinate system with origin $x \in M^n$, $v \in \mathbb{R}^n$, and f is a smooth function defined in a neighborhood of x. To the function f we assign the number $\sum_i \frac{\partial(f\varphi^{-1})}{\partial x_i}(0)v_i$ and call it the *derivative of f in the direction of the vector field v*. In a different coordinate system (V, ψ), v is replaced by the vector w with coordinates $w_i = \sum_j \frac{\partial(\psi\varphi^{-1})_i}{\partial x_j}(0)v_j$; therefore, in the new coordinate system, the function f is assigned the number

$$\sum_{i,j} \frac{\partial(f\psi^{-1})}{\partial x_i}(0)\frac{\partial(\psi\varphi^{-1})_i}{\partial x_j}(0)v_j = \sum_j \frac{\partial(f\varphi^{-1})}{\partial x_j}(0)v_j.$$

Thus, the number assigned to f does not depend on the choice of a coordinate system.

We have associated to each tangent vector v at the point $x \in M^n$ a linear operator $v\colon C^\infty(M^n) \to \mathbb{R}$ (instead of $C^\infty(M^n)$, we could take $C^\infty(U)$, where U is a neighborhood of x; the number $v(f)$ depends only on the behavior of the function f in an arbitrarily small neighborhood of x). This operator has the following properties:

(i) $(\lambda v + \mu w)(f) = \lambda v(f) + \mu w(f)$;

(ii) $v(fg) = f(x)v(g) + g(x)v(f)$.

The second property follows from the Leibniz rule $\frac{\partial(fg)}{\partial x_i} = g\frac{\partial f}{\partial x_i} + f\frac{\partial g}{\partial x_i}$.

Exercise 39. Show that (ii) implies $v(c) = 0$ for any constant function c.

We can take properties (i) and (ii), together with the linearity of the operator v, as a definition of the linear space of tangent vectors at the point $x \in M^n$, but we must verify that no "irrelevant" operators arise, i.e., that if $v\colon C^\infty(\mathbb{R}^n) \to \mathbb{R}$ is a linear operator such that $v(fg) = f(0)v(g) + g(0)v(f)$, then $v(f) = \sum_i \frac{\partial f}{\partial x_i}(0)v_i$ for some $v_1, \ldots, v_n \in \mathbb{R}$. For this purpose, we need the following auxiliary assertion.

Lemma. *Suppose that $f \in C^\infty(U)$, where $U \subset \mathbb{R}^n$ is a convex neighborhood of the origin, and $f(0) = 0$. Then there exist functions $g_1, \ldots, g_n \in C^\infty(U)$ such that $f(x) = \sum x_i g_i(x)$ and $g_i(0) = \frac{\partial f}{\partial x_i}(0)$.*

Proof. Clearly,

$$f(x) = f(x) - f(0) = \int_0^1 \frac{df(tx)}{dt}dt = \int_0^1 \sum x_i \frac{\partial f(tx)}{\partial x_i}dt.$$

therefore, we can take $g_i(x) = \int_0^1 \frac{\partial f(tx)}{\partial x_i} dt$. $\square$

This lemma immediately implies the required assertion. Indeed, since $f(x) - f(0) = \sum x_i g_i(x)$, it follows that

$$v(f) = \sum 0 \cdot v(g_i) + \sum g_i(0) v(x_i) = \sum \frac{\partial f}{\partial x_i}(0) v_i,$$

where $v_i = v(x_i)$.

The tangent vectors at a point $x \in M^n$ form a linear space. This space is called the *tangent space* at the point x and denoted by $T_x M^n$.

Exercise 40. Suppose that $J = \{f \in C^\infty(\mathbb{R}^n) : f(0) = 0\}$, $J^2 = \{\sum f_i g_i \mid f_i, g_i \in J\}$ (the sum is finite), $(J/J^2)^*$ is the space of linear functions on J/J^2, and V is the tangent space at the point $0 \in \mathbb{R}^n$.

(a) Prove that if $v \in V$ and $f \in J^2$, then $v(f) = 0$. Thus, each tangent vector v can be treated as an element of the space $(J/J^2)^*$.

(b) For $l \in (J/J^2)^*$, let $v_l(f) = l(f(x) - f(0))$. Prove that the operator v_l has property (ii) at $x = 0$.

(c) Prove that the maps $V \to (J/J^2)^*$ and $(J/J^2)^* \to V$ constructed in (a) and (b) are mutually inverse.

2.1. The Differential. Suppose that $f\colon M^m \to N^n$ is a smooth map and $v \in T_x M^m$ is a tangent vector. Then we can define a vector $df(v) \in T_{f(x)} N^n$. For example, if v is given as a curve $\gamma(t)$, then $df(v)$ is the curve $f(\gamma(t))$. If v is given as a linear operator on smooth functions, then we define $df(v)$ as the operator $df(v)(\varphi) = v(\varphi f)$ (we can do this because $\varphi \in C^\infty(U_{f(x)})$ implies $\varphi f \in C^\infty(V_x)$).

The map $df\colon T_x M^m \to T_{f(x)} N^n$ is linear; it is called the *differential* of the map f at the point x.

Exercise 41. Prove that $d(f \circ g) = df \circ dg$.

A map f is an immersion (submersion) at a point x if and only if the differential of f at x is a monomorphism (epimorphism). This criterion for f to be an immersion (submersion) is sometimes very convenient.

Example. Consider the map f from the space $\mathbb{R}^{n^2}$ of all matrices of order n to the space $\mathbb{R}^{n(n+1)/2}$ of symmetric matrices defined by $f(X) = X^T X$. This map is a submersion at all points of the set $f^{-1}(I_n)$, where I_n is the identity matrix. (Therefore, according to Theorem 5.3, the topological space $f^{-1}(I_n) = \mathrm{O}(n)$ is a manifold.)

Proof. Suppose that $U \in \mathrm{O}(n)$, i.e., $U^T U = I_n$. Consider a smooth curve of the form $\gamma(t) + U + tA$ in the space of all matrices. The map f takes it to the curve $I_n + t(UA^T + AU^T) + o(t)$; therefore, its differential takes A to $UA^T + (UA^T)^T$. Clearly, any symmetric matrix can be represented as $X + X^T$, and any matrix X can be represented as $X = UA^T$. Thus, the differential of f is an epimorphism at the point U. □

Exercise 42. Prove that the unitary matrix space $\mathrm{U}(n)$ is a manifold.

Exercise 43. Prove that the map $f\colon \mathrm{U}(n) \to S^1$ defined by $f(U) = \det(U)$ is a submersion. In particular, $f^{-1}(1) = \mathrm{SU}(n)$ is a manifold.

2.2. Vector Fields. On the set $TM^n = \bigcup_{x\in M^n} T_xM^n$, the structure of a manifold can be defined as follows. Let (U, φ) be a local coordinate system on the manifold M^n. To each tangent vector at a point $x \in M^n$ we assign the pair $(\varphi(x), v)$, where $v = (v_1, \dots, v_n)$ is the set of coordinates of this tangent vector in the given coordinate system. As a result, we obtain a one-to-one map

$$T\varphi\colon TU = \bigcup_{x\in U} T_xM^n \to \varphi(U) \times \mathbb{R}^n \subset \mathbb{R}^{2n}.$$

The sets TU cover TM^n. Declaring all maps $T\varphi$ to be homeomorphisms, we endow TM^n with the structure of a topological space. The charts $(TU, T\varphi)$ determine the structure of a manifold on this space. The manifold TM^n is called the *tangent bundle* of the manifold M^n.

Problem 67. Prove that the manifold TS^n is homeomorphic to the subset of the complex space $\mathbb{C}^{n+1}$ given by the equation $z_1^2 + \dots + z_{n+1}^2 = 1$.

Assigning to a tangent vector $v \in T_xM^n$ the point x, we obtain a projection $p\colon TM^n \to M^n$. The map p is smooth.

Any smooth map $f\colon M^m \to N^n$ induces the smooth map $df\colon TM^m \to TN^n$ (the *differential* of f).

A *vector field* on a manifold M^n is defined as a smooth section of the projection p, i.e., a smooth map $s\colon M^n \to TM^n$ such that $ps = \mathrm{id}_{M^n}$; it takes each point $x \in M^n$ to a vector $v \in T_xM^n$. The smoothness of the map s means that in any local coordinate system (U, φ), the vector field has the form $\sum a_i \frac{\partial}{\partial x_i}$, where the a_i are smooth functions on $\varphi(U_i)$.

A point $x \in M^n$ is said to be *singular* for a vector field $s\colon M^n \to TM^n$ if $s(x) = 0$.

Example. On the sphere S^{2n+1}, there exists a vector field without singular points.

Proof. If $\gamma(t) = x + ty + \cdots$ is a smooth curve on the sphere S^m, then the vector $y \in \mathbb{R}^{m+1}$ satisfies the equality $(x, y) = 0$. Indeed, $\|\gamma(t)\| = 1$ implies $\|x\|^2 + t(x, y) + \cdots = 1$; therefore, $(x, y) = 0$. The dimension of the space formed by such vectors coincides with that of the tangent space; hence the vectors $y \in \mathbb{R}^{m+1}$ for which $(x, y) = 0$ are in one-to-one correspondence with the tangent vectors at the point $x \in S^m$.

The $(2n+1)$-sphere lies in $(2n+2)$-space. This means that the coordinates of x can be divided into pairs, $x = (u_1, v_1, \dots, u_{n+1}, v_{n+1})$. We set $y = (-v_1, u_1, \dots, -v_{n+1}, u_{n+1})$. As a result, we obtain a vector field without singular points on the sphere S^{2n+1}. □

Problem 68. Prove that there exist three vector fields on the sphere S^{4n+3} which are linearly independent at each point $x \in S^{4n+3}$.

Problem 69. (a) Prove that the maps $f, g\colon S^{2n+1} \to S^{2n+1}$ defined by $f(x) = x$ and $g(x) = -x$ are homotopic.

(b) Prove that the maps $f, g\colon S^{2n} \to S^{2n}$ defined by $f(x) = -x$ and $g(x_0, x_1, \dots, x_{2n}) = (-x_0, x_1, \dots, x_{2n})$ are homotopic.

Theorem 5.9. *On the sphere S^{2n}, there exists no vector field without singular points.*

Proof (see [88]). Suppose that $v(x)$ is a vector field on S^m without singular points, i.e., $v\colon S^m \to \mathbb{R}^{m+1}$ is a smooth map such that $v(x) \neq 0$ and $(x, v(x)) = 0$ for all $x \in S^m$. Replacing $v(x)$ by $v(x)/\|v(x)\|$, we obtain a vector field consisting of unit vectors.

Let us extend the map v over $\mathbb{R}^{m+1} \setminus \{0\}$ by setting $v(rx) = rv(x)$ for $r > 0$ and $x \in S^m$. For $t \in \mathbb{R}$, we define the map $f_t\colon \mathbb{R}^{m+1} \setminus \{0\} \to \mathbb{R}^{m+1}$ by $f_t(x) = x + tv(x)$. If $\|x\| = r$, then $\|f_t(x)\| = \sqrt{1+t^2}\, r$, i.e., $f(S_r^m) \subset S^m_{\sqrt{1+t^2}\, r}$, where S_r^m is the sphere of radius r centered at the origin.

The Jacobian matrix of the map $f_t(x)$ has the form $I + tJ(x)$, where I is the identity matrix and $J(x)$ is the Jacobian matrix of $v(x)$. Thus, for small t, the map f_t satisfies the hypothesis of the inverse function theorem, and the set $f(S_r^m)$ is open in $S^m_{\sqrt{1+t^2}\, r}$. On the other hand, $f(S_r^m)$ is compact and, hence, closed in $S^m_{\sqrt{1+t^2}\, r}$. The connectedness of $S^m_{\sqrt{1+t^2}\, r}$ implies $f(S_r^m) = S^m_{\sqrt{1+t^2}\, r}$.

Let $0 < a < b$. Consider the set $A = \{x \in \mathbb{R}^{m+1} : a \le \|x\| \le b\}$. If t is sufficiently small, then

$$f_t(A) = \left\{x \in \mathbb{R}^{m+1} : \sqrt{1+t^2}\, a \le \|x\| \le \sqrt{1+t^2}\, b\right\}.$$

Therefore, the ratio of the volumes of the sets $f_t(A)$ and A is $\left(\sqrt{1+t^2}\right)^{m+1}$.

The ratio of the volumes of $f_t(A)$ and A can be calculated differently. First, let us show that for a sufficiently small t, the map f_t is one-to-one on A. Indeed, all partial derivatives of $v(x)$ are uniformly bounded on A (because A is compact); hence there exists a constant c such that $\|v(x) - v(y)\| \le c\|x - y\|$

for any $x, y \in A$. Suppose that $x, y \in A$ and $f_t(x) = f_t(y)$. Then $x - y = t(v(x) - v(y))$, whence $\|x-y\| \le c|t| \cdot \|x-y\|$. For $|t| < c^{-1}$, we obtain $x = y$.

The Jacobian determinant for $f_t(x)$ is equal to $\det(I + tJ(x)) = 1 + t\sigma_1(x) + \cdots + t^{m+1}\sigma_{m+1}(x)$, where $\sigma_1, \dots, \sigma_{m+1}$ are smooth functions. If t is sufficiently small, then this determinant is positive and f_t is a homeomorphism between A and $f_t(A)$; therefore, the volume of $f_t(A)$ equals $a_0 + a_1 t + \cdots + a_{m+1}t^{m+1}$, where a_0 is the volume of A and

$$a_k = \int_{a \le |x| < b} \dots \int \sigma_k(x)\, dx_1 \cdots dx_{m+1}.$$

Thus, $(\sqrt{1+t^2})^{m+1}$ is a polynomial in t of degree $m+1$. This is possible only if the number $m+1$ is even. But we have assumed that $m = 2n$, i.e., $m+1$ is odd. □

Remark 5.3. The nonexistence theorem for vector fields without singular points on the sphere S^{2n} has many different proofs. The proof given above is very nonstandard. A more standard proof is based on the Poincaré–Hopf theorem (Theorem 5.28 on p. 228); it uses also Theorem 5.40 on p. 244 and Example 2 on p. 250.

Problem 70. (a) Given a smooth map $f\colon S^{2n} \to S^{2n}$, prove that there exists a point $x \in S^{2n}$ such that either $f(x) = x$ or $f(x) = -x$.

(b) Prove that any smooth map $f\colon \mathbb{R}P^{2n} \to \mathbb{R}P^{2n}$ has a fixed point.

Problem 71* ([2]). A *division algebra* is defined as a finite-dimensional real space K with bilinear multiplication $\mu\colon K \times K \to K$ without zero divisors (i.e., such that if $v \ne 0$ and $w \ne 0$, then $\mu(v, w) \ne 0$) and with two-sided identity element e (this means that $\mu(e, v) = v = \mu(v, e)$ for all $v \in K$). Prove that if $\dim K \ge 2$, then K contains a subalgebra isomorphic to $\mathbb{C}$.

2.3. Riemannian Metric. To define a *Riemannian metric* on a manifold M^n means to specify a smooth inner product (u, v) on the tangent space T_xM^n (an inner product (u, v) is smooth if the function $f\colon TM^n \to \mathbb{R}$ defined by $f(v) = (v, v)$ is smooth, or, equivalently, the function (X, Y) is smooth for any smooth vector fields X and Y on M^n).

Theorem 5.10. *Any manifold M^n admits a Riemannian metric.*

Proof. Let us cover M^n by a countable set of charts $\varphi_i\colon U_i \to \mathbb{R}^n$ and construct a smooth partition of unity $\{f_i\}$ for which $\operatorname{supp} f_i \subset U_i$.

For $x \in U_i$, we define an inner product $(\,\cdot\,,\,\cdot\,)_i$ on T_xM^n as follows. Let vectors $v, w \in T_xM^n$ have coordinates $(v_1, \dots, v_n)$ and $(w_1, \dots, w_n)$ in a local coordinate system (U_i, φ_i). Then we set $(v, w)_i = v_1w_1 + \cdots + v_nw_n$.

For a point $x \in M^n$ and a tangent vectors $v, w \in T_x M^n$, we define

$$(v, w) = \sum_{i-1}^{\infty} f_i(x)(v, w)_i$$

(if the value $(v, w)_i$ is not defined, then $x \notin U_i$ and $f_i(x) = 0$; in this case, we set $f_i(x)(v, w)_i = 0$).

For fixed x, we obtain an expression of the form $\lambda_1 A_1 + \cdots + \lambda_k A_k$, where the λ_i are positive numbers with $\sum \lambda_i = 1$ and the A_i are positive definite symmetric bilinear forms. Any such sum is positive definite. □

2.4. Differential Forms and Orientability. The *cotangent space* at a point $x \in M^n$ is defined as the space of linear functions on $T_x M^n$; it is denoted by $T_x^* M^n$. The set $T^* M^n = \bigcup_{x \in M^n} T_x^* M^n$ can be made into a manifold in the same way as TM^n. Namely, suppose that (U, φ) is a local coordinate system with origin x, $v \in T_x M^n$, and $l \in T_x^* M^n$. In this local coordinate system, the vector v has coordinates $(v_1, \dots, v_n)$, and $l(v) = l_1 v_1 + \cdots + l_n v_n$, where the numbers $l_1, \dots, l_n$ are the same for all vectors. We say that $(l_1, \dots, l_n)$ are the coordinates of the covector l in the given coordinate system. Then we proceed as in the case of TM^n.

Any smooth map $f \colon M^m \to N^n$ induces the map of tangent bundles $df \colon TM^m \to TN^n$; it moves the tangent vectors in the direction in which the map f acts. For cotangent bundles, the map δf induced in a similar way acts in the opposite direction, i.e., $\delta f \colon T^* N^n \to T^* M^m$. Indeed, the map δf is defined by $\delta f(l)(v) = l(df(v))$. This formula shows that if $l \in T^*_{f(x)} N^n \subset T^* N^n$, then $\delta f(l) \in T_x^* M^m \subset T^* M^m$,

Let $\Lambda_x^k M^n$ be the kth exterior power[2] of the space $T_x^* M^n$. On the set $\Lambda^k M^n = \bigcup_{x \in M^n} \Lambda_x^k M^n$, the natural structure of a manifold is defined. A *differential k-form* on a manifold M^n is a smooth section of the canonical projection $p \colon \Lambda^k M^n \to M^n$, i.e., a smooth map $s \colon M^n \to \Lambda^k M^n$ such that $ps = \mathrm{id}_{M^n}$.

In a local coordinate system, the form $\omega \in \Lambda^k M^n$ is written as

$$\omega = \sum_{i_1 < \cdots < i_k} a_{i_1, \dots, i_k} dx_{i_1} \wedge \cdots \wedge dx_{i_k},$$

where $dx_i(v) = v_i$ (the ith coordinate of v in this coordinate system) for all i and v.

A manifold M^n is said to be *orientable* if there exists a set of charts $\{U_\alpha, \varphi_\alpha : \alpha \in A\}$ which covers M^n and has the property that the Jacobian

[2]The definition of the exterior powers of linear spaces can be found, e.g., in [100]. See also [**135**, 8.4].

determinant of $\varphi_\alpha\varphi_\beta^{-1}$ is positive for any $\alpha, \beta \in A$. Such a set of charts is called an *orientation atlas*.

Exercise 44. Prove that the product of any two orientable manifolds is orientable.

Problem 72. Let f be a self-diffeomorphism of the manifold $M^n = \{x \in \mathbb{R}^n : 0 < \|x\| < 1\}$ defined by $f(tu) = (1-t)u$, where $0 < t < 1$, for all unit vectors u. Prove that f is orientation-reversing.

Theorem 5.11. *A manifold M^n is orientable if and only if there exists a nowhere vanishing n-form Ω on M^n.*

Proof. First, suppose that the manifold M^n is orientable. Let $\{U_\alpha, \varphi_\alpha\}$ be an orientation atlas on M^n. The space of n-forms on an n-manifold is one-dimensional; the basis form ω on $\mathbb{R}^n$ is $\omega = dx_1 \wedge \cdots \wedge dx_n$. Consider the form $\omega_\alpha = (\delta\varphi_\alpha)\omega$ on U_α.

If $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is a smooth map, then $(\delta f)\omega = J(x)\omega$, where $J(x)$ is the Jacobian determinant of f. Therefore, $\delta(\varphi_\beta\varphi_\alpha^{-1})\omega = \lambda\omega$, where λ is a positive function on $\varphi_\alpha(U_\alpha \cap U_\beta)$. Thus, $\delta(\varphi_\alpha^{-1})\delta(\varphi_\beta)\omega = \lambda\omega$, and hence $\omega_\beta = \delta(\varphi_\beta)\omega = \delta(\varphi_\alpha)(\lambda\omega) = \lambda_\alpha\omega_\alpha$, where $\lambda_\alpha(x) = \lambda(\varphi_\alpha(x))$ is a positive function on $U_\alpha \cap U_\beta$.

We can assume that the cover $\{U_\alpha\}$ is no more than countable (see Problem 1). Let f_α be a partition of unity subordinate to the cover $\{U_\alpha\}$ and such that $\operatorname{supp} f_\alpha \subset U_\alpha$ for all α. We set $\Omega = \sum f_\alpha\omega_\alpha$. Clearly, the form Ω does not vanish on M^n.

Now let us assume that Ω is a nowhere vanishing n-form on a manifold M^n. Take an arbitrary atlas $\{U_\alpha, \varphi_\alpha\}$ and let $\omega = dx_1 \wedge \cdots \wedge dx_n$ be the basis form on $\mathbb{R}^n$. The form $\delta(\varphi_\alpha)\omega$, which is defined on U_α, is proportional to Ω; therefore, $\delta(\varphi_\alpha)\omega = \mu_\alpha\Omega$, where μ_α is a function on U_α. Since the function μ_α does not vanish, it follows that $\mu_\alpha > 0$ or $\mu_\alpha < 0$ on the entire set U_α (we assume that U_α is connected). If $\mu_\alpha < 0$, then we replace φ_α by the composition ψ_α of φ_α and the map $\mathbb{R}^n \to \mathbb{R}^n$ defined by $(x_1, x_2, \ldots) \mapsto (x_2, x_1, \ldots)$. Clearly,

$$\begin{aligned}\delta(\psi_\alpha)\omega &= \delta(\psi_\alpha)dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n \\ &= \delta(\varphi_\alpha)dx_2 \wedge dx_1 \wedge \cdots \wedge dx_n = -\mu_\alpha\Omega = \nu_\alpha\Omega,\end{aligned}$$

where $\nu_\alpha = -\mu_\alpha > 0$.

As a result, we obtain an orientation atlas, because

$$\delta(\varphi_\beta\varphi_\alpha^{-1})\omega = \delta(\varphi_\alpha^{-1})\delta(\varphi_\beta)\omega = \delta(\varphi_\alpha^{-1})(\mu_\beta\Omega) = \mu_\alpha^{-1}\mu_\beta\omega. \qquad \square$$

On each space $\Lambda_x^n M^n$, we can define an inner product in the same way as in constructing a Riemannian metric on the manifold M^n. Let $\widetilde{M}^n$ be the set of unit vectors in $\Lambda^n M^n$. The space $\Lambda^n_x M^n$ is one-dimensional; therefore, at each point x there are precisely two unit vectors. Thus, the natural projection $p: \widetilde{M}^n \to M^n$ is a double covering. This covering is called the *orientation covering* of the manifold M^n, and $\widetilde{M}^n$ is called the *orientation covering manifold*. This term is explained by the following theorem.

Theorem 5.12. (a) *A manifold $\widetilde{M}^n$ is connected if and only if the manifold M^n is nonorientable.*

(b) *The manifold $\widetilde{M}^n$ is always orientable.*

Proof. (a) If the manifold $\widetilde{M}^n$ is disconnected, then each of its connected components is a nowhere vanishing n-form Ω. Conversely, if an n-form Ω nowhere vanishes, then $\Omega(x)/\|\Omega(x)\|$ is a connected component of the manifold $\widetilde{M}^n$.

(b) By definition, each point $x \in \widetilde{M}^n$ is an n-form $\Omega(x)$ in the tangent space at the point $p(x) \in M^n$. The covering $p: \widetilde{M}^n \to M^n$ induces an isomorphism of cotangent spaces; therefore, $\Omega(x)$ can be regarded as a form on $\widetilde{M}^n$. □

Exercise 45. Prove that if $\pi_1(M^n) = 0$, then the manifold M^n is orientable.

The points of $\widetilde{M}^n$ can be treated as pairs of the form

$$(\text{point } x \in M^n, \text{ orientation of the space } T_x M^n).$$

Lifting a path $\gamma \subset M^n$ to the covering manifold $\widetilde{M}^n$ corresponds to transferring orientation along the path γ. The transfer of orientation along a path can be interpreted geometrically as follows. Let us cover the path γ by finitely many charts $\varphi_i: U_i \to \mathbb{R}^n$ in such a way that each set $\gamma \cap U_i$ is connected. Using a chart φ_i, we can define compatible orientations on all spaces $T_x M^n$ with $x \in U_i$. Thus, if $[a, b] \subset \gamma \cap U_i$, then an orientation given at a can be transferred to b.

Exercise 46. Give a definition of an orientation covering based on the geometric definition of orientation transfer along a path.

3. Embeddings and Immersions

We have defined embeddings and immersions only for manifolds without boundary; thus, we assume that all manifolds considered in this section are without boundary. There is a fairly simple procedure for constructing embeddings of compact n-manifolds in $\mathbb{R}^{2n+1}$. We describe it in Section 3.1. Then, in Section 3.2, we apply these embeddings to prove that any closed manifold is triangulable.

For noncompact manifolds, the construction is quite different. It is based on the fact that if $n \geq 2m$, then any smooth map $M^m \mapsto \mathbb{R}^n$ can be approximated with arbitrary accuracy by an immersion. We discuss immersions of manifolds in Section 3.3; in Section 3.4, we prove that any n-manifold can be embedded in $\mathbb{R}^{2n+1}$ as a closed submanifold.

All these embedding and immersion theorems were proved by Whitney [**145**]. A more delicate argument, also due to Whitney, shows that any n-manifold with $n \geq 2$ immerses in $\mathbb{R}^{2n-1}$, and any compact n-manifold embeds in $\mathbb{R}^{2n}$. A modern exposition of the proofs of these assertions is contained in [**1**].

3.1. Embeddings of Compact Manifolds. In this section, we prove that any compact manifold M^n can be embedded in $\mathbb{R}^{2n+1}$. The construction includes two steps. First, we prove that M^n can be embedded in $\mathbb{R}^N$ for sufficiently large N. Then, we prove that if M^n embeds in $\mathbb{R}^N$ with $N > 2n+1$, then M^n embeds in $\mathbb{R}^{N-1}$.

Theorem 5.13. *A compact manifold M^n can be embedded in $\mathbb{R}^N$ for sufficiently large N.*

Proof. On the compact manifold M^n, there exists a finite family of charts $\varphi_i \colon U_i \to \mathbb{R}^n$, where $i = 1, \ldots, k$, with the following properties:

(i) each set $\varphi_i(U_i)$ coincides with the open disk of radius 2 centered at the origin;

(ii) the preimages of the open unit disk under the maps φ_i cover M^n; we denote them by V_i.

Let us construct a smooth function $\lambda \colon \mathbb{R}^n \to \mathbb{R}$ such that

$$\lambda(y) = \begin{cases} 1 & \text{if } \|y\| \leq 1, \\ 0 & \text{if } \|y\| \geq 2, \end{cases}$$

and $0 < \lambda(y) < 1$ if $2 > \|y\| > 1$. For this purpose, we take the function

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

Then, we set $\beta(t) = \alpha(x-1)\alpha(2-x)$; the function β is positive on the interval $(1, 2)$. Finally, we define

$$\gamma(\tau) = \int_\tau^2 \beta(t)dt \Big/ \int_1^2 \beta(t)dt$$

and $\lambda(y) = \gamma(\|y\|)$.

Let $\lambda_i(x) = \lambda(\varphi_i(x))$. The map $\lambda_i(x)\varphi_i(x)$ is defined on the entire manifold M^n (if $x \notin U_i$, then $\lambda_i(x) = 0$). It is easy to verify that the map

$f\colon M^n \to \mathbb{R}^{(n+1)k}$ defined by

$$x \mapsto \big(\lambda_1(x), \lambda_1(x)\varphi_1(x), \dots, \lambda_k(x), \lambda_k(x)\varphi_k(x)\big)$$

is one-to-one. Indeed, take $x_1 \in \bar{V}_i$. If $x_2 \in \bar{V}_i$, then $\lambda_i(x_1) = \lambda_i(x_2) = 1$; therefore, the equality $\lambda_i(x_1)\varphi_i(x_1) = \lambda_i(x_2)\varphi_i(x_2)$ is equivalent to $\varphi_i(x_1) = \varphi_i(x_2)$, i.e., $x_1 = x_2$. If $x_2 \notin \bar{V}_i$, then $\lambda_i(x_1) = 1$, while $\lambda_i(x_2) < 1$.

The restriction to U_i of the map $x \mapsto \lambda_i(x)\varphi_i(x)$ is an immersion. Indeed, if $x \in U_i$, then $\lambda_i(x) = 1$, and the map $x \mapsto \varphi_i(x)$ is a local diffeomorphism. Hence $f\colon M^n \to \mathbb{R}^{(n+1)k}$ is an immersion. It remains to note that any one-to-one map from the compact space M^n to the Hausdorff space $\mathbb{R}^{(n+1)k}$ is a homeomorphism of M^n onto its image. □

Theorem 5.14. (a) *Let $f\colon M^n \to \mathbb{R}^N$ be an immersion. If $N > 2n$, then there exists a hyperplane $\mathbb{R}^{N-1} \subset \mathbb{R}^N$ such that the composition of f and the projection onto this hyperplane is an immersion.*

(b) *Suppose that M^n is a compact manifold and $f\colon M^n \to \mathbb{R}^N$ is an embedding. If $N > 2n+1$, then there exists a hyperplane $\mathbb{R}^{N-1} \subset \mathbb{R}^N$ such that the composition of f and the projection on this hyperplane is an embedding.*

Proof. (a) The kernel of the projection of $\mathbb{R}^N$ onto the hyperplane $\mathbb{R}^{N-1}_v$ orthogonal to a vector v consists of the vectors proportional to v. Therefore, the composition of f and the projection onto $\mathbb{R}^{N-1}_v$ is an immersion at a point $x \in M^n$ if and only if the vector v does not belong to the image of the map

$$T_xM^n \xrightarrow{df} T_{f(x)}\mathbb{R}^N \cong \mathbb{R}^N.$$

To eliminate the zero vector, we consider S^{N-1} instead of $\mathbb{R}^N$. The images of collinear vectors under the map df are collinear, and the image of any nonzero vector is nonzero. Therefore, we can endow M^n with a Riemannian metric and construct a map $g\colon T_1M^n \to S^{N-1}$, where T_1M^n is the set of unit tangent vectors.

It is easy to verify that T_1M^n is a manifold of dimension $2n-1$. Indeed, the smooth function $TM^n \to \mathbb{R}$ evaluating the squared length of a tangent vector is a submersion at the point $1 \in \mathbb{R}$, and T_1M^n is the preimage of $1 \in \mathbb{R}$.

The map g constructed above is smooth; therefore, if $2n-1 < N-1$, then its image has measure zero. In particular, there exists a vector $v \in S^{N-1}$ that does not belong to the image of g. The composition of f and the projection onto the hyperplane $\mathbb{R}^{N-1}_v$ is an immersion.

(b) We have already proved that if $N > 2n$, then the composition of the map f and the projection onto the hyperplane $\mathbb{R}^{N-1}_v$ is an immersion for almost all $v \in S^{N-1}$. Let us show that if $N > 2n+1$, then the composition of f and the projection onto $\mathbb{R}_v^{N-1}$ is one-to-one for almost all $v \in S^{N-1}$. Consider the map $g\colon (M^n \times M^n) \setminus \Delta \to S^{N-1}$ defined by

$$g(x, y) = \frac{f(x) - f(y)}{\|f(x) - f(y)\|}.$$

Here $\Delta = \{(x, y) \in M^n \times M^n : x = y\}$ is the *diagonal* of $M^n \times M^n$. The map g is well defined because if $x \neq y$, then, by assumption, $f(x) \neq f(y)$. The manifold $(M^n \times M^n) \setminus \Delta$ has dimension $2n$; therefore, if $N > 2n + 1$, then the image of g has measure zero. Clearly, if a vector $v \in S^{N-1}$ does not belong to the image of g, then the composition of f and the projection onto $\mathbb{R}_v^{N-1}$ is one-to-one.

It remains to note that any continuous one-to-one map of the compact space M^n to the Hausdorff space $\mathbb{R}_v^{N-1}$ is a homeomorphism of M^n onto its image. □

3.2. Triangulations of Closed Manifolds. Let M^n be a compact manifold without boundary. In this section, we prove, following [**26**], that M^n is triangulable, i.e., there exists a homeomorphism $M^n \to |K|$, where K is a simplicial complex. To construct a triangulation, we embed M^n in $\mathbb{R}^N$. Each point $x \in M^n$ in $\mathbb{R}^N$ is associated with two affine subspaces containing it, namely, the tangent subspace T_xM^n and the normal subspace N_xM^n (the orthogonal complement of T_xM^n). We say that the sphere of radius r centered at y is tangent to M^n at a point x if $y \in N_xM^n$ and $\|y - x\| = r$.

Lemma. *For any closed manifold $M^n \subset \mathbb{R}^N$, there exists a number $r > 0$ such that any sphere of radius less than r tangent to M^n intersects M^n only in the point of tangency.*

Proof. First, consider the case $M^1 \subset \mathbb{R}^2$. Suppose that the graph $y = f(x)$ of a smooth function f intersects the circle $x^2 + (y - r)^2 = r^2$ at a point $(x_0, r - \sqrt{r^2 - x_0^2})$ and is tangent to the circle at the origin, i.e., $f'(0) = 0$. Suppose also that $\max_{t \in [0,x]} f''(t) = C$. Then $f'(\tau) = \int_0^\tau f''(t)dt \leq Cx$ (for $\tau \in [0, x]$) and

$$r - \sqrt{r^2 - x_0^2} = f(x_0) = \int_0^{x_0} f'(\tau)\, d\tau \leq \frac{Cx_0^2}{2}.$$

Performing a simple algebraic calculation, we see that if $x_0 \in (0, r]$, then

$$\frac{r - \sqrt{r^2 - x_0^2}}{x_0^2} \geq \frac{1}{2r};$$

therefore, $C \geq \frac{1}{r}$. This means that if the radius r is small, then the interval $[0, r]$ contains a point at which the second derivative of f is large; to be more specific, its value at this point is at least $1/r$.

Now, consider the general case $M^n \subset \mathbb{R}^N$. The compact manifold M^n can be covered by finitely many open sets U_i such that for each $x \in U_i$, the orthogonal projection $p_{i,x}\colon U_i \to T_{x_i}M^n$ is a diffeomorphism of U_i onto $U'_{i,x} = p_{i,x}(U_i)$ and, moreover, U_i is the graph of a smooth map $\varphi_{i,x}\colon U'_{i,x} \to N_{x_i}M^n$. If the sphere of radius r tangent to M^n at $x \in U_i$ intersects U_i at a point different from x, then the inequality $C \geq 1/r$ proved above implies a certain inequality for the second partial derivatives of $\varphi_{i,x}$. Using it, we can find a lower bound for the radii of the tangent spheres intersecting U_i. If a tangent sphere has radius smaller than the minimum of these bounds over all i and is tangent to M^n at a point $x \in U_i$, then it does not intersect U_i (but it may intersect U_j for $j \neq i$).

Suppose that there exists a sequence of spheres which are tangent to M^n at points $x_1, x_2, \ldots$, intersect M^n at other points $y_1, y_2, \ldots$, and have radii $r_k \to 0$. We can assume that $x_1, x_2, \ldots \in U_i$ and $x_k \to x \in U_i$ (otherwise, we take a subsequence). Moreover, we can assume that all the r_k are smaller than the minimum bound mentioned above; then none of the points y_k belongs to U_i. On the other hand, $\|x_k - y_k\| \leq 3r_k \to 0$ in $\mathbb{R}^N$. Therefore, $y_k \to x \in U_i$, contradicting the fact that all limit points of the sequence y_k must belong to the closed set $M^n \setminus U_i$. □

For each $\rho > 0$, we can choose points $a_1, \ldots, a_m \in M^n$ such that the open sets

$$\sigma^n(a_k, \rho) = M^n \cap \{y \in \mathbb{R}^n : \|a_k - y\| < \rho\}$$

cover M^n. Take the number r from the lemma and choose a number $\rho < r/2$ so small that each set $\sigma^n(a_k, \rho)$ is homeomorphic to $\operatorname{int} D^n$ and any straight line containing two points of the set $\sigma^n(a_k, \rho)$ makes an angle no larger than $\pi/4$ with the subspace $T_{a_k}M^n$. Then the orthogonal projection $\sigma^n(a_k, \rho) \to T_{a_k}M^n$ is a homeomorphism onto its image.

The sets $c_k^n = \{x \in M^n : \|x - a_k\| \leq \|x - a_i\|, i = 1, \ldots, m\}$ cover M^n. Moreover, $c_k^n \subset \sigma^n(a_k, \rho)$, because if $x \in M^n$ and $\|x - a_k\| > \rho$, then $\|x - a_i\| < \rho$ for some i. Each set c_k^n is the intersection of the manifold M^n with the convex subset of $\mathbb{R}^N$ determined by the inequalities $\|x - a_k\| \leq \|x - a_i\|$, where $i = 1, \ldots, m$. Consider the hyperplane L_{ki} given by $\|x - a_k\| = \|x - a_i\|$. If L_{ki} intersects $N_{a_k}M^n$ in some point y, then the sphere of radius $\|y - a_k\|$ centered at y is tangent to M^n at a_k and intersects M^n in a_i; therefore,

$$\|y - a_k\| > r > 2\rho. \tag{1}$$

Let τ be a ray emanating from a_k in the space $T_{a_k}M^n$. This ray and the subspace $N_{a_k}M^n$ generate a half-space H_τ of dimension $N - n + 1$. The half-space H_τ intersects $\sigma^n(a_k, \rho) \subset M^n$ in some curve γ; the projection of γ on $T_{a_k}M^n$ lies on the ray τ. The curve γ intersects at least one of

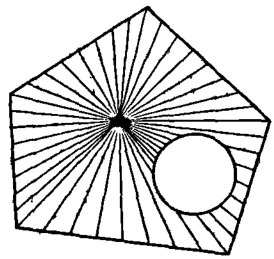

Figure 2. A "bad" set c_k^n

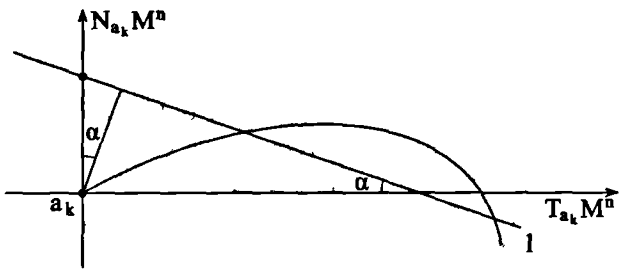

Figure 3. The section by a half-plane contained in H_r

the hyperplanes L_{ki}. Let us show that γ intersects this L_{ki} in precisely one point and is not tangent to L_{ki}. (If the curve γ intersected L_{ki} in two points, then the set c_k^n might be of the form shown in Figure 2.) Suppose that γ is tangent to L_{ki} or intersects it in two points. Let l be the tangent or the line passing through the two intersection points. Both intersection points belong to $\gamma \subset \sigma^n(a_k, \rho)$; therefore, by assumption, the line l makes an angle of $\alpha \leq \pi/4$ with $T_{a_k}M^n$ (for the tangent line, this assertion is proved by passage to the limit). Since the line l intersects $\sigma^n(a_k, \rho)$, the distance from a_k to l is at most ρ. Taking into account the fact that $\alpha \leq \pi/4$, we obtain

$$\|a_k - y\| \leq \rho/\cos\alpha \leq \rho/\cos(\pi/4) = \sqrt{2}\rho$$

(see Figure 3), which contradicts (1).

Thus, each curve γ intersects L_{ki} in at most one point. This enables us to construct a homeomorphism between c_k^n and the convex polyhedron in $T_{a_k}M^n$ determined by the same hyperplanes L_{ki} as c_k^n (see Figure 4). (Note that no hyperplane L_{ki} intersecting c_k^n is parallel to $T_{a_k}M^n$, because such a hyperplane would intersect $N_{a_k}M^n$ in a point y for which $\|a_k - y\| \leq \rho$. And, of course, L_{ki} cannot intersect c_k^n and $T_{a_k}M^n$ on different sides of the point a_k.) The homeomorphism transfers the combinatorial structure of the convex polyhedron to c_k^n; the i-faces of c_k^n have an invariant description based on

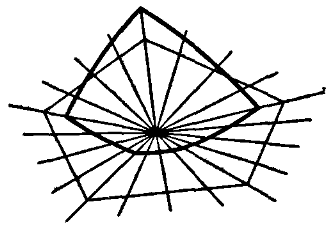

Figure 4. A set c_k^n and a convex polyhedron

the intersections of the hyperplanes L_{ki}. This allows us to triangulate all the c_k^n face by face: first we triangulate the 2-faces, then the 3-faces, and so on.

A compact manifold M^n with boundary ∂M^n can be triangulated by the same method. Consider the closed manifold $\widetilde{M}^n$ obtained from two copies of M^n by identifying the respective points on their boundaries. We embed $\widetilde{M}^n$ in $\mathbb{R}^N$, apply the lemma to $\widetilde{M}^n$ and ∂M^n, and choose a number $r > 0$ such that the assertion of the lemma holds for both manifolds. Then, we choose numbers $\rho_1 < r/2$ and $\rho_2 < r/2$ which can be used as ρ in triangulating $\widetilde{M}^n$ and ∂M^n, respectively. For $\rho = \min\{\rho_1, \rho_2\}$, we first choose points $a_1, \dots, a_m \in \partial M^n$ and then supplement them by $a_{m+1}, \dots, a_{m+k} \in M^n \subset \widetilde{M}^n$.

3.3. Immersions. In this section, we prove the following assertion: *Any manifold M^n (not necessarily compact) can be immersed in $\mathbb{R}^{2n}$; moreover, if $2m \le n$, then any smooth map $f\colon M^m \to \mathbb{R}^n$ can be approximated with arbitrary accuracy by an immersion.*

First, we calculate the dimension of the set of matrices of a given rank. Let $M_{n,m}$ be the set of all matrices $\begin{pmatrix} a_{11} & \dots & a_{1m} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nm} \end{pmatrix}$ with real coefficients; such matrices are in one-to-one correspondence with the linear maps $\mathbb{R}^m \to \mathbb{R}^n$. The set $M_{n,m}$ is naturally identified with $\mathbb{R}^{mn}$. Let $M_{n,m,k}$ be the subset of $M_{n,m} = \mathbb{R}^{mn}$ consisting of all matrices of rank k.

Theorem 5.15. *If $k \le \min(m,n)$, then $M_{n,m,k}$ is a manifold of dimension $k(m+n-k)$.*

Proof. Take any element of $M_{n,m,k}$. Without lost of generality, we can assume that this element has the form $\begin{pmatrix} A_0 & B_0 \\ C_0 & D_0 \end{pmatrix}$, where A_0 is a nonsingular matrix of order k. If $\varepsilon > 0$ is sufficiently small, then any matrix A of order k such that the absolute values of all elements of the matrix $A - A_0$ are smaller than ε is nonsingular. It is easy to verify that for such a matrix A, we have

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in M_{n,m,k} \iff D = CA^{-1}B.$$

Indeed,

$$\operatorname{rank}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \operatorname{rank}\left[\begin{pmatrix} I_k & 0 \\ -CA^{-1} & I_{n-k} \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix}\right]$$
$$= \operatorname{rank}\begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix}.$$

The rank of the last matrix coincides with the rank of A if and only if $D - CA^{-1}B = 0$.

Take a sufficiently small neighborhood U of $\left(\begin{smallmatrix} A_0 & B_0 \\ C_0 & D_0 \end{smallmatrix}\right) \in M_{n,m,k} \subset \mathbb{R}^{mn}$ and consider the map $\varphi\colon U \to \mathbb{R}^{mn}$ defined by

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \mapsto \begin{pmatrix} A & B \\ C & D - CA^{-1}B \end{pmatrix}.$$

This map is invertible; its inverse has the form

$$\begin{pmatrix} A & B \\ C & X \end{pmatrix} \mapsto \begin{pmatrix} A & B \\ C & X + CA^{-1}B \end{pmatrix}.$$

Moreover, $U \cap M_{n,m,k} = \varphi^{-1}\left(\mathbb{R}^{k(m+n-k)} \cap \varphi(U)\right)$, where $\mathbb{R}^{k(m+n-k)} \subset \mathbb{R}^{mn}$ is the subspace consisting of the matrices of the form $\left(\begin{smallmatrix} A & B \\ C & 0 \end{smallmatrix}\right)$. □

Now, consider the local situation where M^m is an open subset of $\mathbb{R}^m$. Recall that "almost all" means "all except those from a set of measure zero."

Theorem 5.16. *Suppose that $U \subset \mathbb{R}^m$ is an open set, $f\colon U \to \mathbb{R}^n$ is a smooth map, and $n \geq 2m$. Then, for almost every linear map $A\colon \mathbb{R}^m \to \mathbb{R}^n$, the map $g\colon U \to \mathbb{R}^n$ defined by $g(x) = f(x) + Ax$ is an immersion.*

Proof. Consider the map $F_k\colon M_{n,m,k} \times U \to M_{n,m}$ defined by $F_k(X, x) \doteq X - df(x)$. According to Theorem 5.15, the dimension of the manifold $M_{n,m,k} \times U$ equals $k(n+m-k)+m$. The function $k(n+m-k)$ monotonically increases in k for fixed m and n provided that $k < (m+n)/2$. Thus, if $k \leq m-1$, then $k < m < 3m/2 \leq (m+n)/2$ and

$$k(n+m-k)+m \leq (m-1)(n+1)+m = (2m-n)+mn-1.$$

By assumption, $2m \leq n$; therefore, $\dim(M_{n,m,k} \times U) < \dim M_{n,m}$, and the image of F_k has measure zero. This means that the set of linear maps of the form $X - df(x)$, where $X \in M_{n,m,k}$ ($k = 1, \ldots, m-1$), has measure zero; i.e., for almost all linear maps A and all $x \in U$, the matrices $A + df(x)$ have rank m. □

Now, we are ready to prove the main theorem.

Theorem 5.17. *Suppose that $n \geq 2m$ and $f\colon M^m \to \mathbb{R}^n$ is a smooth map. Then, for any $\varepsilon > 0$, there exists an immersion $g\colon M^m \to \mathbb{R}^n$ such that $\|f(x) - g(x)\| < \varepsilon$ for all $x \in M^m$.*

Proof. Take a countable family of triples of open sets $U_{i,1} \subset U_{i,2} \subset U_{i,3}$ such that the sets $U_{i,1}$ cover M^m and for each i we have $U_{i,k} = \varphi_i^{-1}(D_k^m)$, where $D_k^m = \{x \in \mathbb{R}^m : \|x\| \le k\}$ and $\varphi_i\colon U_{i,3} \to D_3^m \subset \mathbb{R}^m$ is a smooth chart; in addition, we require that the cover $\{U_{i,3}\}$ be locally finite.[3] We construct the desired map g by induction, successively replacing the map f_{i-1} constructed at the $(i-1)$st step by a map f_i such that

(i) $\|f_i(x) - f_{i-1}(x)\| < \varepsilon/2^i$ for all $x \in M^m$;

(ii) the rank of f_i on the set $\overline{U}_{i,1}$ equals m;

(iii) the map f_i coincides with f_{i-1} outside $U_{i,2}$;

(iv) the rank of f_i equals m at all points of the set $C_i = \overline{U}_{i,2} \cap \big(\bigcup_{j=1}^{i-1} \overline{U}_{j,1}\big)$.

Setting $f_0 = f$ and $g(x) = \lim_{i\to\infty} f_i(x)$, we obtain the required map; it is smooth because the cover $\{U_{i,3}\}$ is locally finite.

We proceed to constructing the maps f_i. For this purpose, we need a smooth function $\lambda\colon \mathbb{R}^m \to \mathbb{R}$ such that

$$\lambda(y) = \begin{cases} 1 & \text{if } \|y\| \le 1, \\ 0 & \text{if } \|y\| \ge 2. \end{cases}$$

(The construction of such a function was described on p. 208.) The map f_i will have the form $f_i(x) = f_{i-1}(x) + \lambda(\varphi_i(x))A\varphi_i(x)$. Obviously, all maps of this form have property (iii). We can perform the construction in the local coordinates associated with the chart $\varphi_i\colon U_{i,3} \to \mathbb{R}^m$. In other words, we can assume that f_i is a map from $D_3^m \subset \mathbb{R}^m$ to $\mathbb{R}^n$ and $f_i(y) = f_{i-1}(y) + \lambda(y)Ay$. If $x \in \overline{U}_{i,1}$, then $y = \varphi_i(x) \in \overline{D_1^m}$; therefore, $\lambda(y) = 1$. In this case, we have $f_i(y) = f_{i-1}(y) + Ay$. According to Theorem 5.16, the map f_i has rank m at all points for almost every A. This ensures (i) and (ii). It remains to achieve (iv).

For each i, the set $C_i = \overline{U}_{i,2} \cap \big(\bigcup_{j=1}^{i-1} \overline{U}_{j,1}\big)$ is compact, and f_{i-1} has rank m at all of its points. Therefore, if all elements of a matrix A are sufficiently small, then the map $f_i(y) = f_{i-1}(y) + \lambda(y)Ay$ has rank m for all $y \in C_i$ (the function $a(y) = \min_{A:\ \operatorname{rank} f_i(y) \le m} \max |a_{ij}|$ attains its minimum on the set C_i). □

3.4. Embeddings of Noncompact Manifolds. In this section, we show that any n-manifold can be embedded in $\mathbb{R}^{2n+1}$ as a closed submanifold. The proof is valid for both compact and noncompact manifolds, but we have already given a simple proof for compact manifolds.

We say that an immersion $f\colon M^m \to N^n$ is *one-to-one* if the map $M^m \to f(M^m)$ is one-to-one. If the manifold M^m is compact, then any

[3] Such open sets were constructed in the proof of Theorem 5.4.

Figure 5. A one-to-one immersion which is not an embedding

one-to-one immersion of this manifold is an embedding. But for noncompact manifolds, this is not so (see Figure 5).

Theorem 5.18. *Suppose that $n \geq 2m+1$ and $f\colon M^m \to \mathbb{R}^n$ is an immersion. Then, for any $\varepsilon > 0$, there exists a one-to-one immersion $g\colon M^m \to \mathbb{R}^n$ such that $\|f(x) - g(x)\| < \varepsilon$ for all $x \in M^m$.*

Proof. As in the proof of Theorem 5.17, we construct the desired map g by successive approximations. We define open sets $U_{i,1} \subset U_{i,2} \subset U_{i,3}$ and a smooth function $\lambda\colon \mathbb{R}^m \to \mathbb{R}$ with the same properties as in Theorem 5.17 but require in addition that the restriction of f to $U_{i,3}$ be one-to-one (this can be achieved because any immersion is locally one-to-one). This time, we seek maps f_i in the form $f_{i-1}(x)+\lambda(\varphi_i(x))v_i$, where the $v_i \in \mathbb{R}^n$ are constant vectors. We want the vectors v_i to be sufficiently small, so that $\|v_i\| < \varepsilon/2^i$.

The maps f_i and f_{i-1} differ only on the compact set $\overline{U}_{i,2}$; therefore, if f_{i-1} is an immersion and v_i is sufficiently small, then f_i is an immersion also.

The equality $f_i(x) = f_i(y)$ is equivalent to $f_{i-1}(x) + \lambda(\varphi_i(x))v_i = f_{i-1}(y) + \lambda(\varphi_i(y))v_i$, and it implies that

$$v_i = -\frac{f_{i-1}(x) - f_{i-1}(y)}{\lambda(\varphi_i(x)) - \lambda(\varphi_i(y))} \tag{1}$$

(provided $\lambda(\varphi_i(x)) \neq \lambda(\varphi_i(y))$).

Let N be the open subset of $M^m \times M^m$ consisting of all pairs (x, y) for which $\lambda(\varphi_i(x)) \neq \lambda(\varphi_i(y))$. Consider the map $N \to \mathbb{R}^n$ defined by the right-hand side of (1). The manifold N has dimension $2m < n$; therefore, the image of this map has measure zero. This means that we can choose an arbitrarily small vector v_i which satisfies (1) nowhere on N. For this vector, the equality $f_i(x) = f_i(y)$ implies $f_{i-1}(x) = f_{i-1}(y)$ and $\lambda(\varphi_i(x)) = \lambda(\varphi_i(y))$. In particular, if $x \in U_{i,1}$ and $f_i(x) = f_i(y)$, then $\lambda(\varphi_i(x)) = \lambda(\varphi_i(y)) = 1$, which means that $y \in U_{i,1}$. Since the restriction of f_{i-1} to $U_{i,1}$ is one-to-one, we have $x = y$. Summarizing, we conclude that the restriction of f_i to $\bigcup_{j=1}^{i} U_{j,1}$ is one-to-one. □

As we will see, the obstruction to a one-to-one immersion f being an embedding is that the sequence $\{f(x_n)\}$ may converge even when the sequence $\{x_n\}$ has no limit points. We denote the set of limits of all such sequences $\{f(x_n)\}$ by $L(f)$.

Exercise 47. Prove that if f is a map from a compact manifold to $\mathbb{R}^n$, then $L(f) = \varnothing$.

Theorem 5.19. *A one-to-one immersion $f\colon M^m \to \mathbb{R}^n$ is an embedding if and only if $L(f) \cap f(M^m) = \varnothing$.*

Proof. A one-to-one immersion $f\colon M^m \to \mathbb{R}^n$ is an embedding if and only if the map $f^{-1}\colon f(M^m) \to M^m$ is continuous.

First, suppose that f^{-1} is continuous. Then the relations $\lim_{k\to\infty} f(x_k) = y$ and $y \in f(M^m)$ imply $\lim_{k\to\infty} x_k = f^{-1}(y)$. Therefore, $L(f) \cap f(M^m) = \varnothing$.

Now, suppose that f^{-1} is discontinuous. Then there exists a point $y \in f(M^n)$ and a sequence $y_k \to y$ such that the sequence $\{x_k = f^{-1}(y_k)\}$ does not converge. The sequence $\{x_k\}$ cannot have limit points different from $x = f^{-1}(y)$. Indeed, if $x_{k_i} \to x' \neq x$, then $y_{k_i} \to f(x') \neq y$. Therefore, the sequence $\{x_k\}$ contains a subsequence having no limit points. □

Now, we can prove the main theorem.

Theorem 5.20. *For any manifold M^m, there exists an embedding $f\colon M^m \to \mathbb{R}^n$, where $n = 2m + 1$. Moreover, there exists an embedding such that the set $f(M^m)$ is closed in $\mathbb{R}^n$.*

Proof. First, let us show that there exists a smooth function $f_1\colon M^m \to \mathbb{R}$ for which $L(f_1) = \varnothing$.

Take the same sets $U_{i,1} \subset U_{i,2} \subset U_{i,3}$, smooth function $\lambda\colon \mathbb{R}^m \to \mathbb{R}$, and charts $\varphi_i\colon U_{i,3} \to \mathbb{R}^m$ as in the proof of Theorem 5.17. We set $f_1(x) = \sum_{i=1}^{\infty} i\lambda(\varphi_i(x))$. If $x \in U_{i,1}$, then $\lambda(\varphi_i(x)) = 1$; therefore, $f(x) \geq 1$. Since the cover $\{U_{i,3}\}$ is locally finite, it follows that the function f_1 is smooth, because $x \notin U_{i,3}$ implies $\lambda(\varphi_i(x)) = 0$.

Suppose that a sequence $\{x_k\}$ has no limit points. Then, for any positive integer N, we can choose $k(N)$ such that $x_k \notin \overline{U}_{1,1} \cap \dots \cap \overline{U}_{N,1}$ for all $k \geq k(N)$. Thus, every such x_k belongs to $U_{i,1}$ with $i > N$, and $f_1(x_k) \geq i > N$. Hence the sequence $\{f_1(x_k)\}$ has no limit points.

Consider the map $f_2\colon M^m \to \mathbb{R}^{2m+1}$ defined by $f_2(x) = (f_1(x), 0, \dots, 0)$. According to Theorem 5.17, for any $\varepsilon > 0$, there exists an immersion $f_3\colon M^m \to \mathbb{R}^{2m+1}$ such that $\|f_2 - f_3\| < \varepsilon$, and by Theorem 5.18 there exists a one-to-one immersion $f\colon M^m \to \mathbb{R}^{2m+1}$ such that $\|f_3 - f\| < \varepsilon$.

We show that $L(f) = \varnothing$ (for all ε). Suppose that a sequence $\{x_k\}$, where $x_k \in M^m$, has no limit points. Then, for any positive integer N, we can choose $k(N)$ such that $f_1(x_k) > N$ whenever $k \geq k(N)$. Therefore, the sequence $\{f(x_k)\}$ has no limit, because $\|f(x_k) - f_2(x_k)\| < 2\varepsilon$.

It remains to verify that the set $f(M^m)$ is closed. This is implied by the following lemma.

Lemma. *The set $f(M^m)$ is closed in $\mathbb{R}^n$ if and only if $L(f) \subset f(M^m)$.*

Proof. First, suppose that $f(M^m)$ is closed in $\mathbb{R}^n$ and $y \in L(f)$. Then $y = \lim_{k\to\infty} f(x_k)$, where $x_k \in M^m$, and hence $y \in f(M^m)$.

Now, suppose that $L(f) \subset f(M^m)$ and a point y belongs to the closure of $f(M^m)$. Then there exists a sequence $\{x_k\} \subset x_k \in M^m$ such that $f(x_k) \to y$. If the sequence $\{x_k\}$ has a limit point x, then it has a subsequence $x_{k_i} \to x$, and $y = \lim_{i\to\infty} f(x_{k_i}) = f(\lim_{i\to\infty} x_{k_i}) = f(x) \in f(M^m)$. If the sequence $\{x_k\}$ has no limit points, then $y \in L(f) \subset f(M^m)$. □

For the embedding f constructed above, the set $L(f)$ is empty, and hence $f(M^m)$ is closed. □

3.5. Impossibility of Certain Embeddings. In this section, we prove that no closed nonorientable manifold of dimension n can be embedded in $\mathbb{R}^{n+1}$. The proof uses fairly obvious assertions concerning transversality and general position, which we do not prove rigorously. We only give the definition of transversality.

Suppose that X and Y are smooth manifolds and $W \subset Y$ is a submanifold of Y. We say that a smooth map $f: X \to Y$ is *transversal* to the submanifold W at a point $x \in X$ if one of the following conditions holds:

(a) $f(x) \notin W$;

(b) $f(x) \in W$ and $T_{f(x)}W + (df)_x(T_xX) = T_{f(x)}Y$.

If a map f is transversal to W at all points $x \in X$, then we say that f is *transversal* to W.

Example. If $\dim X + \dim W < \dim Y$, then $f: X \to Y$ is transversal to W if and only if $f(X) \cap W = \varnothing$.

Theorem 5.21. *Suppose that M^n is a manifold without boundary (not necessarily compact) and $f: M^n \to N^{n+1}$ is an embedding into a simply connected manifold N^{n+1} such that $f(M^n)$ is closed. Then M^n is orientable.*

Proof (see [116]). Suppose that the manifold M^n is nonorientable. Let γ be a curve on M^n such that the transfer along this curve is orientation-reversing. Then the transfer along γ reverses the directions of normal vectors to M^n. If the length of a normal vector transferred along γ is constant and sufficiently small, then the curve described by the end of this vector does not intersect M^n. Using this nonclosed curve, we can easily construct a smooth closed curve $\widetilde{\gamma}$ which transversally intersects M^n in one point. We show that in fact such a curve does not exist.

The contraction of the curve $\widetilde{\gamma}$ in the space N^{n+1} determines a map $g\colon D^2 \to N^{n+1}$ whose restriction to ∂D^2 coincides with $\widetilde{\gamma}$. We can assume that g is smooth. Slightly varying f and g, we bring $f(M^n)$ and $g(D^2)$ into general position. For $n \geq 3$, the generic disk $g(D^2)$ is self-avoiding; therefore, the intersection of $f(M^n)$ and $g(D^2)$ consists of closed curves and arcs of curves whose endpoints are different points of $\widetilde{\gamma}$; these curves and arcs are self-avoiding and pairwise disjoint. If $n = 2$, then the self-intersections of the disk may be impossible to remove by slightly changing g, but they can be made double and triple; the double points sweep out some curves, and the triple points are isolated. In general position, the submanifold $f(M^n)$ does not pass through the triple points of the disk $g(D^2)$. Thus, the intersection of $f(M^n)$ and $g(D^2)$ again consists of closed curves and arcs, but this time, the curves may transversally intersect each other and have transversal self-intersections. However, the number of intersection points of these curves with the curve $\widetilde{\gamma}$ is again even. This is sufficient to obtain a contradiction, because M^n intersects $\widetilde{\gamma}$ in precisely one point. □

Corollary. *No closed nonorientable manifold of dimension n can be embedded in $\mathbb{R}^{n+1}$.*

Thus, any closed two-dimensional surface embedded in S^3 is orientable. Using this fact, we can obtain a complete description of all closed two-dimensional surfaces which can be embedded in $\mathbb{R}P^3$. Clearly, $\mathbb{R}P^2$ is one of them. To the surface $\mathbb{R}P^2$ embedded in $\mathbb{R}P^3$ we can attach any number of handles. Thus, any closed nonorientable surface with odd Euler characteristic can be embedded in $\mathbb{R}P^3$.

Theorem 5.22 (see [21]). *No closed nonorientable two-dimensional surface with even Euler characteristic can be embedded in $\mathbb{R}P^3$.*

Proof (see [30]). Suppose that M^2 is a closed nonorientable surface embedded in $\mathbb{R}P^3$. We want to prove that its Euler characteristic $\chi(M^2)$ is odd. Let $\sigma\colon S^3 \to S^3$ be the antipodal involution (it is defined by $\sigma(x) = -x$), and let $p\colon S^3 \to \mathbb{R}P^3 = S^3/\sigma$ be the natural projection. Take the equatorial sphere S^2 in S^3. We assume that $\mathbb{R}P^2 = p(S^2) \subset \mathbb{R}P^3$. Slightly moving M^2 if necessary, we can assume also that it intersects $\mathbb{R}P^2$ transversally. If $M^2 \cap \mathbb{R}P^2$ is disconnected, then by attaching several handles to M^2, we can construct a new two-dimensional surface N^2 for which $N^2 \cap \mathbb{R}P^2$ is connected. Clearly, the surface N^2 is nonorientable and $\chi(M^2) \equiv \chi(N^2) \pmod 2$. In what follows, we assume that $M^2 \cap \mathbb{R}P^2$ is connected.

Let us show that $p^{-1}(M^2) \cap S^2$ is connected as well. Since the surface $p^{-1}(M^2)$ is embedded in S^3, it is orientable (by the corollary of Theorem 5.21). The quotient of $p^{-1}(M^2)$ under the antipodal involution σ is a nonorientable two-dimensional surface; therefore, the restriction of σ to

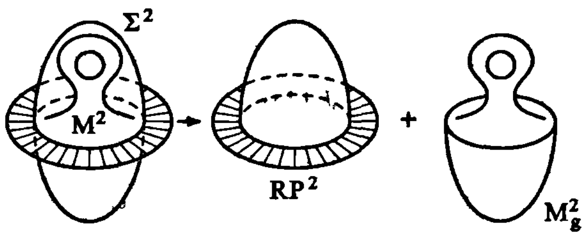

Figure 6. The surgery

$p^{-1}(M^2)$ is orientation-reversing. On the other hand, σ preserves the orientation of the sphere S^3. Hence σ cannot take connected components of $S^3 \setminus p^{-1}(M^2)$ to themselves. Each connected component of $S^2 \setminus p^{-1}(M^2)$ is contained in a connected component of $S^3 \setminus p^{-1}(M^2)$; thus, σ cannot take connected components of $S^2 \setminus p^{-1}(M^2)$ to themselves either. It follows that the number of connected components of $S^2 \setminus p^{-1}(M^2)$ is even.

By assumption, $M^2 \cap \mathbb{R}P^2$ is connected. Therefore, $p^{-1}(M^2) \cap S^2$ has one or two connected components, i.e., $S^2 \setminus p^{-1}(M^2)$ has two or three connected components. But we have proved that the number of connected components of $S^2 \setminus p^{-1}(M^2)$ is even. Hence $S^2 \setminus p^{-1}(M^2)$ has two connected components. This means that $p^{-1}(M^2) \cap S^2$ is connected, i.e., $p^{-1}(M^2) \cap S^2 \approx S^1$.

In what follows, it is convenient to assume that $\mathbb{R}P^3$ is obtained from D^3 by identifying antipodal points of the sphere $S^2 = \partial D^3$. In this case, the restriction of $p\colon D^3 \to \mathbb{R}P^3$ to $D^3 \setminus S^2$ is a homeomorphism, while the restriction of p to S^2 remains a double covering. Let $D^3 \doteq D^3_R = \{x \in \mathbb{R}^3 ;\ \|x\| \le R\}$. We can choose $\varepsilon > 0$ such that the intersection of the closure of $D^3_R \setminus D^3_{R-\varepsilon}$ with M^2 is homeomorphic to the product of $p^{-1}(M^2) \cap S^2$ and the interval $[R-\varepsilon, R]$, i.e., to the cylinder $S^1 \times I$. In $\mathbb{R}P^3$, this cylinder transforms into a Möbius band.

Let Σ^2 be the sphere $\partial D^3_{R-\varepsilon}$; it intersects M^2 in a circle. We cut Σ^2 and M^2 along this circle and glue together the four pieces thus obtained into two closed surfaces as shown in Figure 6. Namely, we attach one half of the sphere Σ^2 to the Möbius band and obtain $\mathbb{R}P^2$. The other half of Σ^2 is attached to the remaining part of M^2. As a result, we obtain an orientable surface, because it is embedded in D^3 (we again apply the corollary of Theorem 5.21). Suppose that this surface has g handles. Then $\chi(M^2) + \chi(\Sigma^2) = \chi(\mathbb{R}P^2) + 2 - 2g$, and hence $\chi(M^2) \equiv \chi(\mathbb{R}P^2) \equiv 1 \pmod 2$. □

4. The Degree of a Map

4.1. The Degree of a Smooth Map. Let $f\colon M^n \to N^n$ be a smooth map of manifolds of the same dimension n. We assume that the manifolds M^n

and N^n are closed and oriented. According to Sard's theorem, the map f has a regular value $y \in N^n$. Let $x \in f^{-1}(y)$. Since $df(x)\colon T_xM^n \to T_yN^n$ is an isomorphism, we can choose local coordinates near x and y which determine orientations compatible with those of the manifolds M^n and N^n. Let $\operatorname{sign} J_f(x)$ denote the sign of the Jacobian of f at the point x. The *degree* of the map f with respect to the point y is

$$\deg(f, y) = \sum_{x \in f^{-1}(y)} \operatorname{sign} J_f(x).$$

This sum is defined because the set $f^{-1}(y)$ is finite. Indeed, suppose that $f^{-1}(y)$ contains infinitely many different points. Since the manifold M^n is compact, there exists a sequence of pairwise different points $x_i \in f^{-1}(y)$, where $i \in \mathbb{N}$, which converges to x_0. We have $f(x_0) = y$; by the inverse function theorem, the point x_0 has a neighborhood U in which f is a homeomorphism. In particular, $(U \setminus \{x_0\}) \cap f^{-1}(y) = \varnothing$. This contradiction shows that $f^{-1}(y)$ is finite.

We assume that if the set $f^{-1}(y)$ is empty, then $\deg(f, y) = 0$.

Example. Let $S^1 = \{z \in \mathbb{C} : |z| = 1\}$. Consider the map $f\colon S^1 \to S^1$ defined by $f(z) = z^n$, where $n \in \mathbb{Z}$. If $n \neq 0$, then $\deg(f, w) = n$ for any point $w \in S^1$. (If $n = 0$, then the irregular point $w = 1$ is excluded from consideration.)

Let $f, g\colon M^m \to N^n$ be smooth maps. We say that f and g are *smoothly homotopic* if there exists a smooth map (homotopy) $F\colon M^m \times I \to N^n$ such that $F(x, 0) = f(x)$ and $F(x, 1) = g(x)$ for all $x \in M^m$.

Theorem 5.23. *Suppose that $f, g\colon M^n \to N^n$ are smoothly homotopic maps of closed oriented manifolds and $y \in N^n$ is a regular value for both maps. Then* $\deg(f, y) = \deg(g, y)$.

Proof. Let $f^{-1}(y) = \{x_1, \dots, x_k\}$. The points $x_1, \dots, x_k$ have pairwise disjoint neighborhoods $U_1, \dots, U_k$ which f diffeomorphically maps to neighborhoods $V_1, \dots, V_k$ of y. Consider the set $V = \bigcap V_i \setminus f(M \setminus \bigcup U_i)$. This set is open and contains y. The preimage of each point $y' \in V$ consists of exactly k points $x'_1, \dots, x'_k$, and $\operatorname{sign} J_f(x'_i) = \operatorname{sign} J_f(x)$. Therefore, $\deg(f, y) = \deg(f, y')$. Consider a similar neighborhood of y for the map g and take the intersection of these two neighborhoods. We obtain an open set $W \ni y$ such that any point $z \in W$ is a regular point for f and g; moreover, $\deg(f, y) = \deg(f, z)$ and $\deg(g, y) = \deg(g, z)$ for $z \in W$.

It follows from Sard's theorem that the map $F\colon M^n \times I \to N^n$ has a regular value z in the open set W. Let us show that $\deg(f, z) = \deg(g, z)$. According to Theorem 5.3 (see p. 186), the set $F^{-1}(z)$ is a one-dimensional submanifold in $M^n \times I$. The connected components of this set are circles

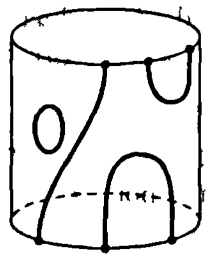

Figure 7. The manifold $F^{-1}(z)$

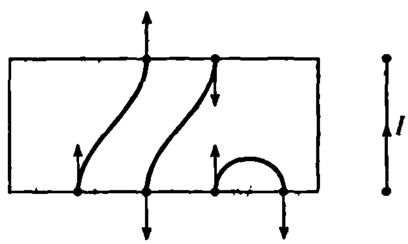

Figure 8. The orientation of the manifold $F^{-1}(z)$

and segments; the endpoints of the segments may belong to one of the sets $M^n \times \{0\}$ and $M^n \times \{1\}$, or they may belong to different sets (see Figure 7). The manifold $F^{-1}(z)$ can be oriented as follows. First, we orient the manifold $M^n \times I$. Then, we choose positively oriented local coordinate systems near $w \in F^{-1}(z)$ and z so that the map F has the form $(x_1, \dots, x_n, x_{n+1}) \mapsto (x_1, \dots, x_n)$ in these local coordinates. The orientation of $F^{-1}(z)$ at the point w is determined by the direction of the coordinate x_{n+1}.

We are interested only in those connected components of $F^{-1}(z)$ that are segments. If the endpoints of such an oriented segment belong to different sets $M^n \times \{0\}$ and $M^n \times \{1\}$, then the orientations at these endpoints have the same sign with respect to the orientation of I, and if the endpoints belong to the same set, then the orientations at its endpoints have opposite signs (see Figure 8). The signs of the Jacobian determinants of the maps $f = F \restriction M^n \times \{0\}$ and $g = F \restriction M^n \times \{1\}$ at each point $w \in F^{-1}(z)$ are completely determined by the sign of the orientation of the manifold $F^{-1}(z)$ at w with respect to the orientation of I. Therefore, the endpoints of an oriented segment correspond to either two points at which the Jacobians of f and g have the same sign or two points at which the Jacobian of the same map (f or g) takes values of opposite signs. This implies $\deg(f, z) = \deg(g, z)$. $\square$

Consider an oriented manifold of the form $W^{n+1} = M^n \times I$. The orientation of W^{n+1} induces opposite orientations on the manifolds $M^n \times \{0\}$ and $M^n \times \{1\}$. Thus, Theorem 5.23 is a special case of the following assertion.

Theorem 5.24. *Suppose that W^{n+1} is a compact oriented manifold with boundary ∂W^{n+1} (endowed with the induced orientation), N^n is a closed oriented manifold, $f: W^{n+1} \to N^n$ is a smooth map, and y is a regular value of $f \restriction \partial W^{n+1}$. Then* $\deg(f \restriction \partial W^{n+1}, y) = 0$.

The proof of Theorem 5.24 is similar to that of Theorem 5.23. Note that the orientability assumption is essential. For example, the degree of the restriction to the boundary of the projection of the Möbius band onto its median circle is nonzero (it equals ± 2). But when degrees are considered modulo 2, Theorem 5.24 holds also for nonorientable manifolds W^{n+1}.

Theorem 5.25. *Suppose that $f: M^n \to N^n$ is a smooth map of closed oriented manifolds, the manifold N^n is connected, and $x, y \in N^n$ are regular values for f. Then* $\deg(f, y) = \deg(f, x)$.

Proof. Let $h: N^n \to N^n$ be a diffeomorphism. A point $x \in N^n$ is a regular value for f if and only if the point $h(x)$ is a regular value for hf. It follows directly from the definition of the degree of a map that if h is orientation-preserving, then $\deg(f, x) = \deg(hf, h(x))$. Therefore, it suffices to prove the existence of a diffeomorphism $h: N^n \to N^n$ with the following properties:

(a) h is orientation-preserving;

(b) $h(x) = y$;

(c) the map hf is smoothly homotopic to f.

Indeed, if such a diffeomorphism exists, then the point y is a regular value for the homotopic maps f and hf, and therefore $\deg(hf, y) = \deg(f, y)$.

Diffeomorphisms h_0 and h_1 are said to be *isotopic* if they are smoothly homotopic and all the intermediate maps h_t are diffeomorphisms. Clearly, any diffeomorphism (of orientable manifolds) isotopic to the identity diffeomorphism is orientation-preserving. Thus, it remains to prove the following lemma.

Lemma (homogeneity of manifolds). *Let N^n be a connected manifold without boundary. Then, for any two points $x, y \in N^n$, there exists a diffeomorphism $h: N^n \to N^n$ that is isotopic to the identity and takes x to y.*

Proof. Let $\lambda: \mathbb{R}^n \to \mathbb{R}$ be a smooth function such that $\lambda(x) > 0$ for $\|x\| < 1$ and $\lambda(x) = 0$ for $\|x\| \geq 1$. Consider the differential equation $\frac{dx}{dt} = \lambda(x)c$, where $c \in \mathbb{R}^n$ is a fixed vector. Let $F_t(x)$ be a solution to this equation with initial condition $F_0(x) = x$. Clearly, $F_{t+s} = F_t \circ F_s$; therefore, F_t is a diffeomorphism isotopic to the identity. The map F_t leaves all points outside the unit disk fixed and shifts all points inside the unit disk in the direction of the vector c. Suppose that $\|x\| < 1$ and $\|y\| < 1$. We set $c = y - x$. For some $t > 0$, the diffeomorphism F_t takes x to y and leaves all points outside the unit disk fixed.

Using this construction, we can obtain the required diffeomorphism $h: N^n \to N^n$ in the case where the points $x, y \in N^n$ belong to the same chart $\varphi: U \to \mathbb{R}^n$, where $\varphi(U)$ is the open unit disk.

We say that points $x, y \in N^n$ are equivalent if there exists a diffeomorphism which is isotopic to the identity and takes x to y. The above considerations show that the equivalence classes are open sets. But the connected manifold N^n does not admit a nontrivial decomposition into a union of pairwise disjoint open sets. This means that there is only one equivalence class. □

According to Theorem 5.25, if M^n and N^n are closed oriented manifolds and N^n is connected, then, for any smooth map $f: M^n \to N^n$, the number $\deg(f, x)$ does not depend on the regular value x. This number is called the *degree* of the smooth map f and denoted by $\deg f$.

Remark 5.4. For closed (but not necessarily orientable) manifolds M^n and N^n, the *degree modulo* 2 can be considered (for nonorientable manifolds, the Jacobian determinant has indefinite sign, but $-1 \equiv 1 \pmod 2$). For the degree modulo 2, Theorems 5.23–5.25 remain valid.

Problem 73. Let M^2 be a sphere with g handles, where $g \geq 1$. Prove that any smooth map $f: S^2 \to M^2$ has degree zero.

Problem 74. Prove that $\deg(fg) = (\deg f)(\deg g)$.

Problem 75. Let $P(z)$ be a polynomial of degree n. Prove that the map $\mathbb{C} \to \mathbb{C}$ defined by $z \mapsto P(z)$ can be extended to a smooth map $\mathbb{C}P^1 \to \mathbb{C}P^1$. Calculate the degree of this map.

Problem 76. Let $R(z)$ be a simple fraction of two polynomials of degrees m and n. Prove that the map defined by $z \mapsto R(z)$ can be extended to a smooth map $\mathbb{C}P^1 \to \mathbb{C}P^1$. Calculate the degree of this map.

Problem 77. Given a map $f: S^n \to S^n$, consider the map $\Sigma f: \Sigma S^n \to \Sigma S^n$ which takes (x, t) to $(f(x), t)$ for all $(x, t) \in S^n \times \{t\}$ and all t. Prove that $\deg f = \deg \Sigma f$.

Problem 78. Let S^{2n-1} be the unit sphere in the space $\mathbb{C}^n$ with coordinates $(r_1 e^{i\varphi_1}, \dots, r_n e^{i\varphi_n})$. Calculate the degree of the map $f: S^{2n-1} \to S^{2n-1}$ defined by

$$(r_1 e^{i\varphi_1}, \dots, r_n e^{i\varphi_n}) \mapsto (r_1 e^{ik_1\varphi_1}, \dots, r_n e^{ik_n\varphi_n}),$$

where $k_1, \dots, k_n$ are integers.

Problem 79. Is the map $f\colon \mathrm{SO}(n) \to \mathrm{SO}(n)$ ($n \geq 2$) defined by $f(A) = A^2$ homotopic to the identity map?

4.2. The Index of a Singular Point of a Vector Field. Suppose that M^n is a manifold without boundary and $v\colon M^n \to TM^n$ is a smooth vector field on M^n. A point $x \in M^n$ is *singular* for v if $v(x) = 0$. A singular point x is said to be *isolated* if it has a neighborhood containing no other singular points.

Let U be an open subset in $\mathbb{R}^n$, and let $v\colon U \to \mathbb{R}^n$ be a smooth vector field with an isolated singular point $x_0 \in U$. For sufficiently small $r > 0$, the disk $\|x - x_0\| \leq r$ contains no other singular points. Consider the map of the sphere $\|x - x_0\| = r$ to the unit sphere defined by $x \mapsto v(x)/\|v(x)\|$. The degree of this map is called the *index* of the singular point x_0. Clearly, the index of a singular point is an integer, and it continuously depends on r (provided that the singular point x_0 under consideration is the only singular point in the disk $\|x - x_0\| \leq r$); therefore, the index does not depend on r at all.

The index of an isolated singular point $x_0 \in M^n$ of a vector field v can be defined as follows. Consider a smooth chart $\varphi\colon U \to \mathbb{R}^n$, where $x_0 \in U$ and φ is a homeomorphism of U onto the entire space $\mathbb{R}^n$. The vector field v induces the vector field $d\varphi(v)$ on $\mathbb{R}^n$, for which $\varphi(x_0)$ is an isolated singular point. The index of the singular point x_0 of v is defined as the index of the singular point $\varphi(x_0)$ of $d\varphi(v)$.

This definition is to be checked for consistency. That is, we must show that if $\psi\colon U \to \mathbb{R}^n$ is another chart,[4] then the index of the singular point $\psi(x_0)$ of the vector field $d\psi(v)$ is equal to the index of the singular point $\varphi(x_0)$ of the vector field $d\varphi(v)$. Consider the diffeomorphism $f = \psi\varphi^{-1}\colon \mathbb{R}^n \to \mathbb{R}^n$ and define $y = \varphi(x)$, $w(y) = d\varphi(v(x))$, and $y_0 = \varphi(x_0)$. We must prove the following assertion.

Lemma 1. *Suppose that y_0 is an isolated singular point of a vector field w and $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is a diffeomorphism. Then the index of the singular point y_0 of the vector field w is equal to the index of the singular point $f(y_0)$ of the vector field $df(w)$.*

We consider orientation-preserving and orientation-reversing diffeomorphisms separately. For orientation-preserving diffeomorphisms, Lemma 1 can easily be derived from the following statement.

Lemma 2. *Any orientation-preserving diffeomorphism $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is isotopic to the identity diffeomorphism.*

[4]We assume that the domain U is the same. Indeed, the index is determined by the behavior of the vector field v in an arbitrarily small neighborhood of x_0, so it cannot depend on the choice of U.

Proof. For any $a \in \mathbb{R}^n$, the map $x \mapsto a + f(x)$ is a diffeomorphism. Therefore, the maps $f_t(x) = (t-1)f(0) + f(t)$ are diffeomorphisms for all t. Moreover, $f_1 = f$ and $f_0(0) = 0$. Thus, we can assume that $f(0) = 0$. According to the lemma on p. 200, f can be represented in the form $f(x) = \sum x_i g_i(x)$, where $g_1, \dots, g_n$ are smooth maps and $g_i(0) = \frac{\partial f}{\partial x_i}(0)$. Setting

$$F(x,t) = x_1 g_1(tx) + \dots + x_n g_n(tx),$$

we obtain an isotopy between f and the linear transformation

$$F(x,0) = x_1 \frac{\partial f}{\partial x_1}(0) + \dots + x_n \frac{\partial f}{\partial x_n}(0).$$

It remains to prove that any orientation-preserving linear transformation is isotopic to the identity transformation. Each matrix with positive determinant can be represented as SU, where S is a symmetric positive definite matrix and U is an orthogonal matrix with positive determinant. There exists a basis in which the transformation S has diagonal matrix with positive diagonal elements, and there exists a basis in which the transformation U has block-diagonal matrix with diagonal elements 1 and $\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$. The isotopies between the transformations S and U and the identity transformation are constructed in the obvious way. □

We proceed to prove Lemma 1. First, suppose that the diffeomorphism $f\colon \mathbb{R}^n \to \mathbb{R}^n$ preserves orientation. Let f_t be an isotopy between f and the identity map. The index of the singular point $f_t(y_0)$ of the vector field $df_t(w)$ does not depend on t; therefore, the index corresponding to $t = 1$ coincides with that corresponding to $t = 0$. But this is exactly what we need.

Now suppose that the diffeomorphism f reverses orientation. Consider the symmetry $s(x_1, x_2, \dots, x_n) = (-x_1, x_2, \dots, x_n)$ about the hyperplane $x_1 = 0$. The diffeomorphism sf preserves orientation. Therefore, it is sufficient to show that the indices of the singular points x_0 and $s(x_0)$ of the vector fields w and $ds(w)$ coincide. If $w(x) = (w_1, w_2, \dots, w_n)$, then

$$ds(w(s(x))) = (-w_1, w_2, \dots, w_n) = sw(x).$$

Hence the index of the singular point x_0 of w is equal to the degree of the map $W\colon S^{n-1} \to S^{n-1}$ defined by $W(x) = w(x)/\|w(x)\|$, and the index of the singular point $s(x_0)$ of $ds(w)$ is equal to the degree of the map $W' = sWs^{-1}$. We have $\deg s = \deg s^{-1} = -1$ and, according to Problem 74, $\deg W' = (\deg s)^2 \deg W = \deg W$.

Problem 80 (Poincaré). Suppose that the trajectories[5] of a vector field v with one singular point in the plane are tangent to some circle C internally

[5]The *trajectory* of a vector field is a curve $\gamma(t)$ such that the velocity vector $\frac{d\gamma}{dt}(t_0)$ is equal to $v(\gamma(t_0))$ for all t_0.

at i points and externally at e points, and C encloses the singular point. Prove that the index of this singular point is equal to $1 + (i - e)/2$.

Problem 81*. Let f be a smooth function in the plane with coordinates (x_1, x_2). Prove that the index of an isolated singular point of the vector field $v = \operatorname{grad} f = (\partial f/\partial x_1, \partial f/\partial x_2)$ can take the values 1, 0, -1, $-2, \ldots$ and cannot take other values.

Suppose that a manifold M^n is embedded in $\mathbb{R}^N$ and $\varphi\colon U \to M^n \subset \mathbb{R}^N$ is a diffeomorphism of a domain $U \subset \mathbb{R}^n$ onto a domain $\varphi(U) \subset M^n$. Let $x = (x_1, \ldots, x_n) \in U$. Then the vectors $e_i(x) = \frac{\partial \varphi}{\partial x_i}(x)$ form a basis in the space $T_{\varphi(x)}M^n$; hence $v(\varphi(x)) = \sum v_i(x)e_i(x)$, where the v_i are smooth functions. Each vector $e_j(x)$ is determined by the curve $\varphi(x_1, \ldots, x_j + t, \ldots, x_n)$. The map v takes it to the curve $\sum v_i(\ldots, x_j + t, \ldots)e_i(\ldots, x_j + t, \ldots)$. The tangent vector to this curve is

$$\sum_i \frac{\partial v_i}{\partial x_j}(x)e_i(x) + \sum_i v_i(x)\frac{\partial e_i}{\partial x_j}(x).$$

In particular, if $\varphi(x)$ is a singular point of v, then this tangent vector belongs to the space spanned by $e_1(x), \ldots, e_n(x)$. This means that the tangent spaces at the singular points of the vector field v are invariant under the map dv.

A singular point y of a vector field v is said to be *nondegenerate* if the linear operator $dv\colon T_yM^n \to T_yM^n$ is nondegenerate.

Theorem 5.26. *Any nondegenerate singular point y of a vector field v is isolated and has index ± 1; the sign of the index coincides with that of the determinant of the operator $dv\colon T_yM^n \to T_yM^n$.*

Proof. Choose local coordinates with origin y. We shall consider v as a map from $\mathbb{R}^n$ to $\mathbb{R}^n$. By assumption, the Jacobian determinant of this map does not vanish at the origin; therefore, by the inverse function theorem, there is a neighborhood U of the origin such that v is a diffeomorphism of U onto its image. (This implies, in particular, that v has exactly one singular point in U.) Identifying the neighborhood U and its image with $\mathbb{R}^n$, we obtain a diffeomorphism $v\colon \mathbb{R}^n \to \mathbb{R}^n$. This diffeomorphism preserves orientation if and only if $\det(dv) > 0$. According to Lemma 2 on p. 225, any orientation-preserving diffeomorphism $\mathbb{R}^n \to \mathbb{R}^n$ is isotopic to the identity diffeomorphism. Thus, if v preserves orientation, then the singular point has index 1. If the diffeomorphism v reverses orientation, then it is isotopic to a symmetry about a hyperplane. In this case, the map $S^{n-1} \to S^{n-1}$ has degree -1, and the index of the singular point is -1. □

One of the most important properties of vector fields with isolated singular points on closed manifolds is that the sum of indices of their singular points is constant. To prove this, we need the following assertion, which is also used in many other theorems.

Theorem 5.27 (tubular neighborhoods). *Suppose that M^n is a closed manifold and $f\colon M^n \to \mathbb{R}^m$ is an embedding. Let M_ε be the set of all points in $\mathbb{R}^m$ which are at distance at most ε from $f(M^n)$. Then the number $\varepsilon > 0$ can be chosen so that each point $y \in M_\varepsilon$ has a unique representation $y = x + \xi$, where $x \in M^n$ and $\xi \perp T_xM^n$.*

Proof. Let N be a set of pairs (x, ξ), where $x \in M^n$ and ξ is a vector orthogonal to $T_xM^n \subset \mathbb{R}^m$. We can define the structure of an m-manifold on the set N as follows. We introduce local coordinates $(u_1, \dots, u_n)$ on M^n and choose an orthonormal system of vectors $\varepsilon_1, \dots, \varepsilon_{m-n}$ orthogonal to T_xM^n at each point x of this local coordinate system; we assume that the vectors ε_i smoothly depend on x. To each pair (x, ξ) we assign the coordinates $(u_1, \dots, u_n, \xi_1, \dots, \xi_{m-n})$, where $\xi = \xi_1\varepsilon_1 + \dots + \xi_{m-n}\varepsilon_{m-n}$. In these coordinates, the map F defined by $(x, \xi) \mapsto x + \xi$ has the form

$$(u_1, \dots, u_n, \xi_1, \dots, \xi_{m-n}) \mapsto x(u_1, \dots, u_n) + \xi_1\varepsilon_1 + \dots + \xi_{m-n}\varepsilon_{m-n} \in \mathbb{R}^m,$$

where $x(u_1, \dots, u_n) \in M^n \subset \mathbb{R}^m$ is the point of M^n with local coordinates $(u_1, \dots, u_n)$. The Jacobian matrix of this map is

$$\left(e_1 + \sum \xi_k \frac{\partial \varepsilon_k}{\partial u_1}, \dots, e_n + \sum \xi_k \frac{\partial \varepsilon_k}{\partial u_n}, \varepsilon_1, \dots, \varepsilon_{m-n}\right),$$

where the $e_i = \frac{\partial x}{\partial u_i}$ $(i = 1, \dots, n)$ are the basis vectors of T_xM^n. In this expression, each vector is represented by the column of its coordinates.

The vectors $e_1, \dots, e_n$ form a basis in the space T_xM^n, and the vectors $\varepsilon_1, \dots, \varepsilon_{m-n}$ form a basis in its orthogonal complement. Therefore, for $\xi = 0$, the map $(x, \xi) \mapsto x + \xi$ is locally one-to-one. Since the manifold M^n is compact, we can choose $\varepsilon > 0$ such that the restriction F_ε of F to the set $N_\varepsilon = \{(x, \xi) \in N : \|\xi\| \le \varepsilon\}$ is one-to-one. The map $F_\varepsilon\colon N_\varepsilon \to \mathbb{R}^m$ is a one-to-one immersion of a compact manifold; therefore, F_ε is an embedding. In particular, $M_\varepsilon = F_\varepsilon(N_\varepsilon)$ is a compact manifold with boundary. Moreover, each point $y \in M_\varepsilon$ has a unique representation in the form $y = x + \xi$, where $x \in M^n$ and $\xi \perp T_xM^n$. □

Theorem 5.28 (Poincaré–Hopf). *The sum of the indices of singular points is the same for all vector fields with isolated singular points on a closed manifold M^n.*

Proof. Consider any embedding $f\colon M^n \to \mathbb{R}^m$. First, suppose that v is a vector field with nondegenerate singular points on M^n (the case of degenerate singular points is considered at the end of the proof). We use the notation from the proof of Theorem 5.27. The vector field v is extended over M_ε as follows. We represent each point $y \in M_\varepsilon$ as $y = x + \xi$ and set $\tilde{v}(y) = v(x) + \xi$. Clearly, $v(x) \perp \xi$ and $\xi = 0$ if and only if $y \in M^n$. Therefore, the vector field $\tilde{v}$ has the same singular points as v. It follows from Theorem 5.26 that the singular points of $\tilde{v}$ have the same indices as the singular points of v (the operator $d\tilde{v}$ is obtained from dv by adding the identity map as a direct summand).

Lemma. *The sum of the indices of singular points of the vector field $\tilde{v}$ is equal to the degree of the map $\partial M_\varepsilon \to S^{m-1}$ defined by $y = x + \xi \mapsto \xi$. In particular, this sum does not depend on $\tilde{v}$.*

Proof. The tangent space $T_y(\partial M_\varepsilon)$, where $y = x + \xi$, is the hyperplane orthogonal to ξ. The vector $v(x)$ belongs to this hyperplane; therefore, $(\tilde{v}(y), \xi) = (\xi, \xi) > 0$. For $t \in [0,1]$ and $y \in \partial M_\varepsilon$, we set $w_t(y) = t\tilde{v}(y) + (1-t)\xi$. We have $(w_t(y), \xi) = t(\tilde{v}(y), \xi) + (1-t)(\xi, \xi) > 0$; in particular, $w_t(y) \neq 0$. Hence the degree of the map $\partial M_\varepsilon \to S^{m-1}$ defined by $y \mapsto w_t(y)/\|w_t(y)\|$ does not depend on t. Thus, we must prove that the degree of the map $\partial M_\varepsilon \to S^{m-1}$ defined by

$$y \mapsto \tilde{v}(y)/\|\tilde{v}(y)\| \tag{1}$$

is equal to the sum of the indices of singular points of the vector field $\tilde{v}(y)$.

Removing small disks $D_1^m, \dots, D_k^m$ containing singular points from the manifold M_ε, we obtain a manifold M'_ε with boundary $\partial M_\varepsilon \cup S_1^{m-1} \cup \dots \cup S_k^{m-1}$. The orientation of S_i^{m-1} induced by the orientation of M'_ε is opposite to that induced by the orientation of D_i^m. This means that if S_i^{m-1} is oriented as the boundary of the manifold M'_ε, then the degree of the map $S_i^{m-1} \to S_i^{m-1}$ defined by (1) is equal to the index of the ith singular point taken with opposite sign.

Formula (1) defines a smooth map of the manifold M'_ε; therefore, according to Theorem 5.24, the restriction of this map to the boundary has degree zero. Hence the degree of the map $\partial M_\varepsilon \to S^{m-1}$ is equal to the sum of the degrees of $S_i^{m-1} \to S_i^{m-1}$ taken with opposite signs, i.e., to the sum of the indices of singular points. □

It remains to consider the case in which the vector field v has degenerate singular points. We change v only in small neighborhoods of degenerate singular points; thus, we can assume that v is a vector field on an open subset of $\mathbb{R}^n$. Let y_0 be an isolated degenerate singular point of v, and let $\lambda\colon \mathbb{R}^n \to [0,1] \subset \mathbb{R}$ be a smooth function taking the value 1 on an open set $U \ni y_0$ and vanishing outside an open set V. We assume that V is sufficiently small for its closure $\overline{V}$ to contain no singular points different from y_0. Let v_0 be a regular value of the map $v\colon U \to \mathbb{R}^n$. We put $v'(y) = v(y) - \lambda(y)v_0$.

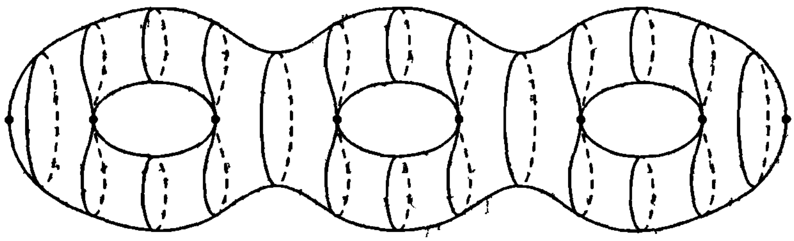

Figure 9. The trajectories of a vector field on the sphere with g handles

Suppose that the minimum value of the function $\|v(y)\|$ on the compact set $\overline{V}\setminus U$ is $\delta > 0$. The regular value v_0 can be chosen so that $\|v_0\| < \delta$ (because the set $v(U)$ contains the vector $v(y_0) = 0$); then $v'(y) \neq 0$ for all $y \in \overline{V}\setminus U$. If $y \in U$, then $\lambda(y) = 1$ and $v'(y) = v(y) - v_0$. Therefore, the singular points of v' belonging to U are preimages of the regular value v_0 of the map $v\colon U \to \mathbb{R}^n$; all these singular points are nondegenerate.

It remains to note that both the index of the singular point y_0 of v and the sum of the indices of the singular points of v' that belong to U are equal to the degree of the map $\partial U \to S^{n-1}$ defined by $y \mapsto v(y)/\|v(y)\|$. □

Example. The sum of indices of any vector field (with isolated singular points) on the sphere with g handles is equal to $2 - 2g$.

Proof. On the sphere with g handles, there exists a vector field with two singular points of index 1 and $2g$ singular points of index -1 (see Figure 9). □

Theorem 5.29. *Suppose M^n and N^n are closed manifolds and $p\colon M^n \to N^n$ is a smooth k-fold covering. If the sum of indices of any vector field with isolated singular points on N^n equals χ, then the sum of indices of any vector field with isolated singular points on M^n equals $k\chi$.*

Proof. Let v be a vector field on N^n. Since the covering p is a local diffeomorphism, we can define the vector field $\tilde{v} = d(p^{-1})(v)$ on M^n; we mean that each vector $\tilde{v}(y)$ is equal to $d(p^{-1})(v(p(y)))$, where p^{-1} is the inverse map to the projection of a neighborhood of y onto a neighborhood of $p(y)$.

To each singular point of v correspond k singular points of $\tilde{v}$ with the same index. □

Corollary. *The sum of indices of any vector field with isolated singular points on the nonorientable surface nP^2 is equal to $2 - n$.*

Proof. The orientation covering of the surface nP^2 is $(n-1)T^2$. □

Remark 5.5. A vector field of index $\chi(M^2)$ can be constructed directly from a triangulation of a two-dimensional surface M^2 as follows. We take

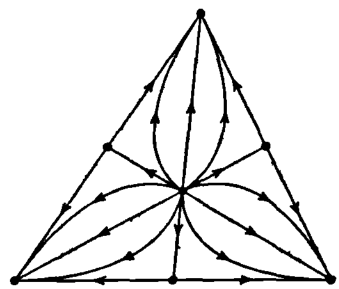

Figure 10. Construction of a vector field from a triangulation

the barycentric subdivision of the given triangulation and construct a vector field on its 1-skeleton such that it exits the vertices corresponding to the centers of faces and enters all vertices of the initial triangulation. This vector field can be extended to a vector field on M^2 (see Figure 10). The only singular points of this vector field are the vertices of the barycentric subdivision. The singular points corresponding to the vertices and faces have index 1, and the singular points corresponding to the edges have index -1. This construction can be generalized to n-manifolds; in this case, the vector field on the 1-skeleton of the barycentric subdivision is constructed so that it is directed from centers of k-faces to centers of l-faces for $k > l$.

Theorem 5.30. *The sum of the indices of a vector field with isolated singular points on a closed manifold of odd dimension is zero.*

Proof. By the Poincaré–Hopf theorem, the sums of the indices of singular points of vector fields v and $-v$ are equal. Therefore, it suffices to prove that if a map $f\colon S^{n-1} \to S^{n-1}$ has degree d, then the map $-f$ has degree $(-1)^n d$, or, in other words, that the degree of the map $x \mapsto -x$ equals $(-1)^n$. But this map is the composition of n maps of the form

$$(\dots, x_{i-1}, x_i, x_{i+1}, \dots) \mapsto (\dots, x_{i-1}, -x_i, x_{i+1}, \dots),$$

each of which has degree -1. □

4.3. The Hopf Theorem. We have already proved that if maps $f, g\colon M^n \to N^n$ are smoothly homotopic, then $\deg f = \deg g$ (this follows from Theorems 5.23 and 5.25). Hopf [58] proved that if $N^n = S^n$, then the converse is true.

Theorem 5.31 (Hopf). *Let M^n be a closed oriented connected manifold.*

(a) *If the degrees of smooth maps $f, g\colon M^n \to S^n$ are equal, then these maps are homotopic.*

(b) *For any integer m, there is a smooth map $f\colon M^n \to S^n$ of degree m.*

Proof. (a) Many of the arguments used below apply only in the case of $n \geq 2$; for $n = 1$, they must be changed, although not substantially. We consider only $n \geq 2$, because the proof for maps $S^1 \to S^1$ is fairy simple.

First, consider the simplest case in which there is a regular point $y_0 \in S^n$ such that the sets $f^{-1}(y_0)$ and $g^{-1}(y_0)$ consist of $|\deg f| = |\deg g|$ elements.

The map g is homotopic to a map g_1 such that $f^{-1}(y_0) = g_1^{-1}(y_0)$; the signs of the Jacobian determinants of f and g_1 at the preimages of y_0 coincide. Indeed, it suffices to show that if $\{a_1, \dots, a_k\}$ and $\{b_1, \dots, b_k\}$ are sets of pairwise different points of a connected manifold without boundary, then there exists a diffeomorphism of this manifold which is isotopic to the identity map and takes each b_i to a_i. A diffeomorphism taking b_1 to a_1 exists by the lemma on the homogeneity of manifolds (see p. 223). To obtain the required diffeomorphism, we remove the point a_1 from the manifold, apply the lemma on the homogeneity of manifolds to the resulting manifold (it is connected if $n \geq 2$), and so on.

In what follows, we assume that $f^{-1}(y_0) = g^{-1}(y_0) = \{a_1, \dots, a_k\}$ and the Jacobians of the maps f and g at the points $a_1, \dots, a_k$ have the same sign. Choose pairwise disjoint neighborhoods $U_i \ni a_i$. The set $f\big(M^n \setminus \bigcup_{i=1}^k U_i\big)$ does not contain y_0; therefore, the map f is homotopic to a map f_1 with the following properties:

- $f_1\big(M^n \setminus \bigcup_{i=1}^k U_i\big) = y_1$, where y_1 is the antipode of y_0 on the sphere S^n;
- for each i, the map f_1 coincides with f in some neighborhood $V_i \subset U_i$ of a_i.

If the neighborhoods V_i are sufficiently small, then the restriction of f to each of them is a diffeomorphism. Therefore, applying an additional homotopy, we can transform the map f_1 into a map f_2 whose restriction to each V_i is a diffeomorphism of $V_i \approx \mathbb{R}^n$ onto $S^n \setminus \{y_1\} \approx \mathbb{R}^n$. According to Lemma 2 on p. 225, any two orientation-preserving (or orientation-reversing) diffeomorphisms $\mathbb{R}^n \to \mathbb{R}^n$ are isotopic. Therefore, the maps f and g are homotopic.

To complete the proof, it remains to consider the case where the preimage of y_0 contains points at which the Jacobians have opposite signs. In this case, the homotopy is constructed as follows. For each pair of points (a_i, b_i), where $a_i \in f^{-1}(y_0)$, $b_i \in g^{-1}(y_0)$, and the Jacobians of f and g have the same signs at a_i and b_i, we perform the above construction and obtain a tube $U_i \times I$ in $\subset M^n \times I$; on this tube, the required map to S^n is defined as above. We join pairs of the remaining points from the preimages of y_0 (under f and g) at which the Jacobians have opposite signs by similar tubes (see Figure 11). The tubes can easily be rendered pairwise disjoint (this is obvious

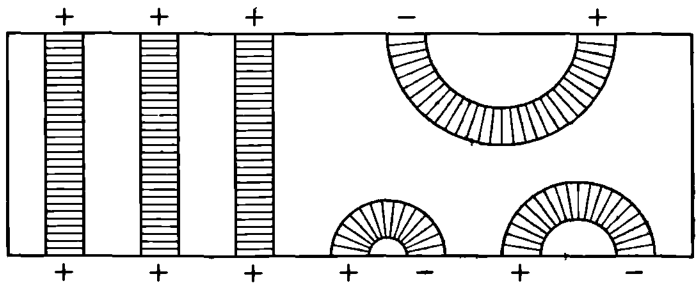

Figure 11. Elimination of preimages with Jacobians of opposite signs

for $n \geq 2$). On the new tubes, the map to S^n is constructed in the same way as on the old ones. The complement of the tubes is mapped to one point y_1.

(b) We choose pairwise disjoint open sets $U_1, \dots, U_{|m|} \subset M^n$ and map them diffeomorphically to $S^n \setminus \{y_1\}$, where y_1 is a fixed point. The diffeomorphisms are chosen so that the signs of their Jacobian determinants coincide with that of m. The remaining part of the manifold is mapped to the point y_1 (if $m = 0$, then the entire manifold M^n is mapped to y_1). □

Remark 5.6. A simplicial map of oriented pseudomanifolds also has a degree, and for these degrees, an analog of the Hopf theorem is valid; such an analog was proved in [**105**].

Problem 82. Prove that if M^n is a closed nonorientable connected manifold, then smooth maps $f, g\colon M^n \to S^n$ are homotopic if and only if their degrees are congruent modulo 2.

4.4. Approximation of Continuous Maps. Any continuous map $M^m \to N^n$ of closed manifolds can be approximated by a smooth map with an arbitrary accuracy. In this section, we prove this assertion by using an embedding of N^n in Euclidean space. Another approach is discussed in [54].

We start with approximation of maps $M^m \to \mathbb{R}^n$. A map to $\mathbb{R}^n$ can be approximated coordinatewise; therefore, it is sufficient to consider the case $n = 1$.

Theorem 5.32. *Suppose that M^m is a closed manifold and $f\colon M^m \to \mathbb{R}$ is a continuous function. Then, for any $\varepsilon > 0$, there exists a smooth function $g\colon M^m \to \mathbb{R}$ such that $|f(x) - g(x)| \leq \varepsilon$ for all $x \in M^m$.*

Proof. Every point $x \in M^m$ has an open neighborhood U_x such that $|f(x) - f(y)| \leq \varepsilon$ for all $y \in U_x$. Take a cover $\{U_{x_1}, \dots, U_{x_k}\}$ of M^n by such neighborhoods and consider a smooth partition of unity $\{\lambda_i\}$ subordinate to this cover. Let $g(x) = f(x_1)\lambda_1(x) + \dots + f(x_k)\lambda_k(x)$. The identity

$\sum \lambda_i(x) = 1$ implies

$$f(x) - g(x) = f(x) - \sum f(x_i)\lambda_i(x) = \sum \lambda_i(x)\big(f(x) - f(x_i)\big).$$

If $x \notin U_{x_i}$, then $\lambda_i(x) = 0$. If $x \in U_{x_i}$, then $|f(x) - f(y)| \le \varepsilon$. In both cases, $|\lambda_i(x)(f(x) - f(x_i))| \le \varepsilon\lambda_i(x)$; therefore, $|f(x) - g(x)| \le \sum \varepsilon\lambda_i(x) = \varepsilon$. □

Using the tubular neighborhood theorem (Theorem 5.27 on p. 228), we can construct a smooth approximation of a map $f\colon M^m \to N^n$ as follows. Suppose that $N^n \subset \mathbb{R}^N$ and $g_1\colon M^m \to \mathbb{R}^N$ is a smooth map for which $\|g_1(x) - f(x)\| < \varepsilon$, where ε is the same as in the tubular neighborhood theorem. Then $g_1(x) = g(x) + \xi(x)$, where $g(x) \in N^n$, $\xi(x) \perp T_{g(x)}N^n$, and $\|\xi(x)\| \le \varepsilon$. Here $g\colon M^m \to N^n$ is a smooth map and $\|g(x) - f(x)\| \le \|g_1(x) - f(x)\| + \|\xi(x)\| < 2\varepsilon$.

Theorem 5.33. *Let M^m and N^n be closed manifolds. Then* (a) *any continuous map $f\colon M^m \to N^n$ is homotopic to a smooth map $g\colon M^m \to N^n$;* (b) *any two homotopic smooth maps $f, g\colon M^m \to N^n$ are smoothly homotopic.*

Proof. (a) Consider an embedding $N^n \hookrightarrow \mathbb{R}^N$ and take $\varepsilon > 0$ from the tubular neighborhood theorem. Let $g\colon M^m \to N^n \subset \mathbb{R}^N$ be a smooth map for which $\|f(x)-g(x)\| < \varepsilon$. Then the segment with endpoints $f(x)$ and $g(x)$ is contained entirely in the tubular ε-neighborhood. We take the orthogonal projection on N^n of the point which divides this segment in the ratio $t : (1-t)$.

(b) Two copies of the interval $[0, 1]$ can be glued together so as to form the circle S^1. Thus, a homotopy between maps f and g can be treated as a map $M^m \times S^1 \to N^n$. Approximating this continuous map by a smooth map, we obtain a smooth homotopy between f_1 and g_1, where f_1 and g_1 are approximations of f and g. The fact that the maps f and f_1 (as well as g and g_1) are smoothly homotopic can be proved by using the construction of (a). □

The theory of homotopy groups considers maps which take base points to base points. We can pass from arbitrary smooth approximations to approximations with this property by using the lemma on the homogeneity of manifolds (see p. 223). It is seen from the proof of this lemma that if a point y belongs to a small neighborhood of a point y_0, then it is possible to construct a diffeomorphism $N^n \to N^n$ taking y to y_0 in such a way that it smoothly depends on y. This allows us to pass from a smooth approximation of a homotopy $H\colon M^m \times I \to N^n$ for which $H(x_0, t) = y_0$ to a smooth homotopy H' for which $H'(x_0, t) = y_0$.

Now, we can prove that $\pi_n(S^n) = \mathbb{Z}$ for $n \geq 2$. We derive this from the Hopf theorem. We have to prove that if maps $f, g\colon S^n \to S^n$ taking x_0 to y_0 are homotopic, then they are homotopic in the class of maps taking x_0 to y_0. The required homotopy is constructed as follows. According to the above considerations, we can assume that the maps f and g and a homotopy H between them are smooth. The path $H(x_0, t)$ is contained in some contractible open set $U \subset S^n$ (this is not so for $n = 1$). Let $F(x, t)$ be a homotopy in the class of maps taking y_0 to y_0 which joins the identity map $S^n \to S^n$ to a map sending U to y_0. The homotopies $\varphi_t(x) = F(f(x), t)$ and $\psi_t(x) = F(g(x), t)$ join the maps f and g to f' and g', respectively, and the homotopy $H'(x, t) = F(H(x, t), 1)$ joins f' to g'.

Problem 83. Prove that if the indices of singular points of some vector field on a closed manifold M^n sum to 0, then there exists a vector field without singular points on M^n.

Let PM^n be the quotient of the space of nonzero tangent vectors to a manifold M^n by the equivalence relation $v \sim \lambda v$ (λ is an arbitrary nonzero number), and let $p\colon PM^n \to M^n$ be the natural projection. The space PM^n has a natural structure of a manifold. The smooth sections of the projection p are called *line element fields* on the manifold M^n. In other words, any line element field on M^n determines a one-dimensional subspace of T_xM^n at each point $x \in M^n$, and these subspaces smoothly depend on x.

Problem 84. Prove that the existence of a line element field on a closed manifold M^n is equivalent to the existence of a vector field without singular points.

4.5. The Pontryagin Construction. The Hopf theorem suggests an interpretation of the elements of the group $\pi_n(S^n)$ in the language of framed manifolds. The Pontryagin construction generalizes this interpretation to the groups $\pi_{n+k}(S^n)$ with $k \geq 0$ and $n \geq 2$.

A smooth closed submanifold $M^k \subset \mathbb{R}^{n+k}$ is said to be *framed* if, at each point $x \in M^k$, an orthonormal set of vectors $v_1(x), \dots, v_n(x)$ orthogonal to T_xM^k and smoothly depending on x is given. The manifold M^k is not required to be connected; it may consist of several connected components of the same dimension k. It is assumed that the empty set is a framed manifold of arbitrary dimension.

Two framed manifolds M_0^k and M_1^k are *framed cobordant* if there exists a submanifold W^{k+1} in $\mathbb{R}^{n+k+1}$ with the following properties:

(a) W^{k+1} is contained in the strip $0 \leq x_{n+k+1} \leq 1$;

(b) the boundary of W^{k+1} consists of the manifolds M_0^k and M_1^k, and these manifolds are contained in the hyperplanes $x_{n+k+1} = 0$ and $x_{n+k+1} = 1$, respectively;

(c) W^{k+1} approaches these hyperplanes orthogonally, i.e., at each point of the boundary, the tangent space to W^{k+1} is orthogonal to the hyperplanes;

(d) the smooth families of orthonormal sets of vectors given on M_0^k and M_1^k can be extended to a smooth family of orthonormal sets of vectors on W^{k+1}.

The set of classes of framed cobordant manifolds of dimension k in $\mathbb{R}^{n+k}$ is denoted by $\Omega_{\mathrm{fr}}^k(n+k)$. The set $\Omega_{\mathrm{fr}}^k(n+k)$ admits the structure of an Abelian group. The sum of two elements is the class of the union of their representatives contained in different half-spaces $\mathbb{R}_+^{n+k}$ and $\mathbb{R}_-^{n+k}$. The zero element is the class containing the empty set. The inverse element is obtained by reversing the orientation of the orthonormal basis (e.g., by changing $v_1(x)$ for $-v_1(x)$). The same argument as the one used in the proof of the Hopf theorem to eliminate preimages with Jacobians of opposite signs shows that this indeed gives an inverse element.

A framed zero-dimensional submanifold in $\mathbb{R}^n$ is a set of m_+ points at which positively oriented bases are given and m_- points at which negatively oriented bases are given. The class of such a framed manifold in $\Omega_{\mathrm{fr}}^0(n)$ is determined by the number $m_+ - m_-$. The Hopf theorem establishes the isomorphism $\Omega_{\mathrm{fr}}^0(n) \cong \pi_n(S^n)$ for $n \geq 2$. (To be more precise, the isomorphism holds for $n = 1$ as well, but the arguments used for $n > 1$ do not apply to $n = 1$.)

Theorem 5.34 (Pontryagin). *For $k \geq 0$ and $n \geq 2$, the group $\Omega_{\mathrm{fr}}^k(n+k)$ is isomorphic to $\pi_{n+k}(S^n)$.*

Proof. To a framed manifold $M^k \subset \mathbb{R}^{n+k}$ we can assign a map $f\colon S^{n+k} \to S^n$ as follows. According to the tubular neighborhood theorem (Theorem 5.27 on p. 228), there exists an $\varepsilon > 0$ such that if $\|a\| < \varepsilon$, then the map $M^k \times \mathbb{R}^n \to \mathbb{R}^{n+k}$ defined by $(x, a) \mapsto x + \sum a_i v_i(x)$ is a homeomorphism of $M^k \times D_\varepsilon^n$ onto the ε-neighborhood of M^k in $\mathbb{R}^{n+k}$. Suppose that $x_0 \in S^{n+k}$ and $y_0 \in S^n$ are base points. We identify $S^{n+k} \setminus \{x_0\}$ with $\mathbb{R}^{n+k}$ and $S^n \setminus \{y_0\}$ with D_ε^n. Then, we map all points of $\mathbb{R}^{n+k}$ not contained in the ε-neighborhood of M^k to y_0, identify the ε-neighborhood of M^k with $M^k \times D_\varepsilon^n$, and take its projection to $D_\varepsilon^n = S^n \setminus \{y_0\}$.

If M_0^k and M_1^k are framed cobordant manifolds, then, applying a similar construction to W^{k+1}, we obtain a map $S^{n+k} \times I \to S^n$ which is a homotopy between the maps $f_0, f_1\colon S^{n+k} \to S^n$.

Now let us assign a framed submanifold $M^k \subset \mathbb{R}^{n+k}$ to a map $f\colon S^{n+k} \to S^n$. We replace the continuous map f by a smooth map g homotopic to it, take a regular value $y_1 \in S^k$ different from the base point $y_0 = g(x_0)$, and let $M^k = g^{-1}(y_1)$. The manifold M^k is framed as follows. At each point $y_1 \in \mathbb{R}^n = S^n \setminus \{y_0\}$, we take an orthonormal basis $e_1, \dots, e_n$ and let $v_i(x)$ be the normal vector to M^k at x for which $dg(v_i(x)) = e_i$. It can be proved in the same way as for the degree that smoothly homotopic maps determine framed cobordant manifolds and the equivalence class of M^k does not depend on the choice of the regular point y_1.

The maps of the groups $\Omega^k_{\mathrm{fr}}(n+k)$ and $\pi_{n+k}(S^n)$ constructed above are mutually inverse and respect the group operations. □

Problem 85. Prove that the Hopf fibration $p \colon S^3 \to S^2$ is the generator of the group $\pi_3(S^2) \cong \mathbb{Z}$ and describe the corresponding framed manifold in $\Omega^1_{\mathrm{fr}}(3)$.

4.6. Homotopy Equivalent Lens Spaces. Take an integer $p > 1$ and integers $q_1, \dots, q_n$, where $n \geq 2$ is coprime to p. The unit sphere $S^{2n-1} \subset \mathbb{C}^n$ admits the action of the group $\mathbb{Z}_p$ defined by

$$\sigma(z_1, \dots, z_n) = \big(\exp(2\pi i q_1/p) z_1, \dots, \exp(2\pi i q_n/p) z_n\big),$$

where σ is a generator of $\mathbb{Z}_p$. The quotient under this action (which has no fixed points) is a manifold. We denote this manifold by $L_p(q_1, \dots, q_n)$ and call it the *lens space*.

The map $\pi \colon S^{2n-1} \to L_p(q_1, \dots, q_n)$ is a p-fold covering with automorphism group $\mathbb{Z}_p$. Therefore, $\pi_1\big(L_p(q_1, \dots, q_n)\big) = \mathbb{Z}_p$.

If a number k is coprime to p, then $L_p(q_1, \dots, q_n) = L_p(kq_1, \dots, kq_n)$, because in the group $\mathbb{Z}_p$ with generator σ, the element $k\sigma$ is also a generator. For $n = 2$ (i.e., for 3-manifolds), we obtain $L_p(q_1, q_2) = L_p(1, q_1^{-1} q_2)$; thus, any three-dimensional lens space has the form $L_p(1, q)$. In the topology of 3-manifolds, the spaces $L_p(1, q)$ are usually denoted by $L(p, q)$.

Theorem 5.35. *Suppose that* $q_1 \cdots q_n \equiv \pm k^n q_1' \cdots q_n' \pmod p$ *for some integer* k. *Then the lens spaces* $L_p(q_1, \dots, q_n)$ *and* $L_p(q_1', \dots, q_n')$ *are homotopy equivalent.*

Proof. Clearly, k is coprime to p. Therefore,

$$L_p(q_1', \dots, q_n') = L_p(kq_1', \dots, kq_n') = L_p(q_1'', \dots, q_n''),$$

where $q_1'' \cdots q_n'' = k^n q_1' \cdots q_n'$. Thus, we must prove that if $q_1 \cdots q_n \equiv \pm q_1'' \cdots q_n'' \pmod p$, then $L_p(q_1, \dots, q_n) \sim L_p(q_1'', \dots, q_n'')$. For simplicity, we take $q_i'' = kq_i'$ for q_i', i.e., assume that $q_1 \cdots q_n = \pm q_1' \cdots q_n'$.

Choose a number k_j such that $k_j q_j \equiv q_j' \pmod p$ and consider the map $\tilde{f} \colon S^{2n-1} \to S^{2n-1}$ defined by

$$\tilde{f}(r_1 e^{i\varphi_1}, \dots, r_n e^{i\varphi_n}) = (r_1 e^{ik_1\varphi_1}, \dots, r_n e^{ik_n\varphi_n}).$$

According to Problem 78, $\tilde{f}$ has degree $k_1 \cdots k_n$. The condition $k_j q_j \equiv q'_j \pmod p$ means that $\tilde{f}$ induces a map of quotient spaces $f\colon L \to L'$, where $L = L_p(q_1, \ldots, q_n)$ and $L' = L_p(q'_1, \ldots, q'_n)$. Indeed, $\tilde{f}$ takes the point with coordinates $r_j e^{i\varphi_j} e^{(2\pi i q_j p)}$ to the point with coordinates $r_j e^{ik_j\varphi_j} e^{(2\pi i q'_j p)}$, because $k_j q_j \equiv q'_j \pmod p$. Thus, points equivalent with respect to the map σ are taken to points equivalent with respect to the map σ'.

The degree of f is equal to the degree of $\tilde{f}$, i.e., $\deg f = k_1 \cdots k_n$.

Consider the maps

$$L \to L \vee S^{2n-1} \xrightarrow{\mathrm{id} \vee (\deg\ d)} L \vee S^{2n-1} \xrightarrow{f \vee \pi'} L'$$

defined as follows. The first map contracts the boundary of a small disk in L to a point. The second map is the identity on L and a map $S^{2n-1} \to S^{2n-1}$ of degree d on S^{2n-1}. The third map coincides with f on L and with the canonical projection $\pi'\colon S^{2n-1} \to L'$ on S^{2n-1}. Let $g\colon L \to L'$ be the composition of these maps. It follows directly from the definition of the degree of a map that $\deg g = \deg f + dp = k_1 \cdots k_n + dp$. But $k_1 \cdots k_n \equiv q'_1 q_1^{-1} \cdots q'_n q_n^{-1} \equiv \pm 1 \pmod p$; therefore, d can be chosen so that $\deg g = \pm 1$. In what follows, we assume d to have this property.

Let $g'\colon L' \to L$ be a map defined similarly. The maps g and g' are the required homotopy equivalences. According to the Whitehead theorem (Theorem 4.21 on p. 179), it is sufficient to verify that the homomorphism $g_*\colon \pi_n(L) \to \pi_n(L')$ is an isomorphism for all $n \geq 1$. The sphere S^{2n-1} is simply connected; hence, for $n = 1$, it suffices to show that the map $f_*\colon \pi_1(L) \to \pi_1(L')$ is an isomorphism. Consider the commutative diagram

$$\begin{tikzcd}
S^{2n-1} \arrow[r, "\tilde{f}"] \arrow[d, "\pi"] & S^{2n-1} \arrow[r, "\tilde{f}'"] \arrow[d, "\pi'"] & S^{2n-1} \arrow[d, "\pi"] \\
L \arrow[r, "f"] & L' \arrow[r, "f"] & L
\end{tikzcd}$$

If follows directly from the definitions that $\tilde{f}'\tilde{f}$ is the identity map; therefore, $f'f$ is the identity map as well. Thus, $f_*\colon \pi_1(L) \to \pi_1(L')$ is an isomorphism.

Now, suppose that $n \geq 2$. Using the universality of the covering $\pi\colon S^{2n-1} \to L$, we construct the commutative diagram

$$\begin{tikzcd}
S^{2n-1} \arrow[r, "\tilde{f}"] \arrow[d, "\pi"] & S^{2n-1} \arrow[d, "\pi'"] \\
L \arrow[r, "f"] & L'
\end{tikzcd}$$

The map $\tilde{g}$, as well as g, has degree ± 1. Therefore, by the Hopf theorem, $\tilde{g}$ is homotopic to either the identity map or the symmetry about the equatorial hyperplane. Hence the maps $\tilde{g}_*: \pi_n(S^{2n-1}) \to \pi_n(S^{2n-1})$ are isomorphisms for all n. The map $g_*: \pi_n(L) \to \pi_n(L')$ is an isomorphism also, because the maps $\pi_*: \pi_n(S^{2n-1}) \to \pi_n(L)$ and $\pi'_*: \pi_n(S^{2n-1}) \to \pi_n(L')$ are isomorphisms for $n \geq 2$. □

For the three-dimensional lens spaces $L(p, q)$, Theorem 5.35 is stated as follows.

Theorem 5.36. *If $q \equiv \pm k^2 q' \pmod p$, then the lens spaces $L(p, q)$ and $L(p, q')$ are homotopy equivalent.*

Indeed, we have $L(p, q) = L_p(1, q)$, i.e., $q_1 = 1$ and $q_2 = q$. Therefore, the equalities $q = \pm k^2 q'$ and $q_1 q_2 = \pm k^2 q'_1 q'_2$ are equivalent.

5. Morse Theory

5.1. Morse Functions. Suppose that M^n is a manifold without boundary and $f: M^n \to \mathbb{R}$ is a smooth function. A point $x \in M^n$ is critical for f if and only if rank $f(x) = 0$, i.e., the map $df: T_x M^n \to \mathbb{R}$ identically vanishes. In local coordinates $(x_1, \dots, x_n)$, this means that $\frac{\partial f}{\partial x_i}(x) = 0$ for $i = 1, \dots, n$.

A critical point x of a function f is said to be *nondegenerate*if its *Hessian matrix* $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right)$ is nonsingular. This definition does not depend on the choice of local coordinates because when the local coordinates are changed to $(y_1, \dots, y_n)$, the Hessian matrix transforms as

$$\left(\frac{\partial^2 f}{\partial y_i \partial y_j}(x)\right) = J^T \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right) J,$$

where $J = \left(\frac{\partial x_i}{\partial y_j}\right)$.

A smooth function $f: M^n \to \mathbb{R}$ is called a *Morse function* if all of its critical points are nondegenerate.

Recall that the index of a quadratic form $\sum a_{ij} x_i x_j$ with symmetric matrix (a_{ij}) is defined as follows. Making a change of variables (over the field $\mathbb{R}$), we can write the quadratic form as $-y_1^2 - \dots - y_q^2 + y_{q+1}^2 + \dots + y_n^2$. The number q is called the *index* of such a quadratic form. The index of a quadratic form can also be defined as the maximal dimension of a subspace on which the form is negative definite.

The *index* of a nondegenerate critical point x of a function f is defined as the index of the Hessian matrix of f at the point x.

Exercise 48. Let $f(x) = -x_1^2 - \dots - x_q^2 + x_{q+1}^2 + \dots + x_n^2$. Prove that the point $x_0 = (0, \dots, 0)$ is critical and has index q.

Theorem 5.37 (Morse lemma). *In a neighborhood of a nondegenerate critical point of index q, there exist local coordinates with origin at this critical* *point such that the function f has the form $f(x_1,\dots,x_n) = f(0) - x_1^2 - \cdots - x_q^2 + x_{q+1}^2 + \cdots + x_n^2$ in these coordinates.*

Proof. We can assume that $f(0) = 0$ and local coordinates are defined in a convex neighborhood of 0 in $\mathbb{R}^n$. According to the lemma on p. 200, there exist smooth functions $g_1,\dots,g_n$ such that $f(x) = \sum x_i g_i(x)$ and $g_i(0) = \frac{\partial f}{\partial x_i}(0)$. By assumption, the point 0 is critical, i.e., $\frac{\partial f}{\partial x_i}(0) = 0$. Applying the same lemma, we obtain $f(x) = \sum x_i x_j h_{ij}(x)$, where $h_{ij}(0) = \frac{\partial^2 f}{\partial x_i \partial x_j}(0)$, i.e., $(h_{ij}(0))$ is the Hessian matrix of f at the critical point. Changing $h_{ij}(x)$ for $\frac{1}{2}(h_{ij}(x) + h_{ji}(x))$, we can assume that the matrix $h_{ij}(x)$ is symmetric, and performing a linear transformation of coordinates, we can also assume that $h_{11}(0) \neq 0$. Finally, decreasing the coordinate neighborhood if necessary, we can assume that $h_{11}(x)/h_{11}(0) > 0$ for all x from this neighborhood. We set

$$y_1 = x_1 + \frac{h_{12}(x)}{h_{11}(x)} x_2 + \cdots + \frac{h_{1n}(x)}{h_{11}(x)} x_n \quad \text{and} \quad y_i = x_i \text{ for } i \geq 2.$$

By the inverse function theorem, $(x_1,\dots,x_n) \mapsto (y_1,\dots,y_n)$ is a diffeomorphism (possibly, in a smaller neighborhood). It is easy to verify that

$$\sum x_i x_j h_{ij}(x) = h_{11}(x) y_1^2 + \sum_{i,j\geq 2} y_i y_j \tilde{h}_{ij}(x).$$

We make the changes $z_1 = y_1\sqrt{|h_{11}(x)|}$ and $z_i = y_i$ for $i \geq 2$, apply similar transformations to the quadratic form in $n-1$ variables, and so on. □

Corollary. *Any nondegenerate critical point is isolated.*

We prove that Morse functions exist on any manifold. We give two different proofs, each with its own advantages. The first proof shows that any smooth function can be turned into a Morse function by a small change; here a "small change" means a small change of the first and second derivatives. The second proof is constructive. Moreover, it shows that there exists a Morse function f for which all sets of the form $\{x \in M^n : f(x) \leq c\}$ are compact; this property is useful when dealing with noncompact manifolds.

Theorem 5.38. *On any manifold M^n, a Morse function exists.*

First proof. Let $g\colon M^n \to \mathbb{R}$ be an arbitrary smooth function (e.g., a constant). We shall construct a Morse function f by successively changing the function g, as in the proof of Theorem 5.17 (see p. 214). Take the same domains $U_{i,1} \subset U_{i,2} \subset U_{i,3}$, charts $\varphi_i\colon U_{i,3} \to \mathbb{R}^n$, and function $\lambda\colon \mathbb{R}^n \to \mathbb{R}$ as in that proof. The function g can be changed to a function having no degenerate critical points in $U_{i,1}$ by using the following lemma.

Lemma 1. *Suppose that $U \subset \mathbb{R}^n$ is an open set and $f: U \to \mathbb{R}$ is a smooth function. Then, for almost all linear functions $A: \mathbb{R}^n \to \mathbb{R}$, the function $f + A$ has only nondegenerate critical points.*

Proof. Consider the map $F: U \to \mathbb{R}^n$ defined by

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

A point x_0 is critical for F if and only if the Hessian matrix of f at x_0 is degenerate. Therefore, x_0 is a degenerate critical point for a function of the form $f(x) - a_1x_1 - \dots - a_nx_n$ if and only if $F(x_0) = (a_1, \dots, a_n)$ and x_0 is a critical point for F, i.e., if $(a_1, \dots, a_n)$ is the image of a critical point of F. It remains to apply Sard's theorem. □

It follows from Lemma 1 that if g_{i-1} is a smooth function on a manifold M^n, then there exists a linear function $A(x) = a_1x_1 + \dots + a_nx_n$ with arbitrarily small coefficients a_i such that the function $g_i(y) = g_{i-1}(y) + \lambda(\varphi_i(y))A(\varphi_i(y))$ has no degenerate critical points on the set $\overline{U}_{i,1}$ (Lemma 1 should be applied to $U = U_{i,2} \supset \overline{U}_{i,1}$; note that $\lambda(\varphi_i(y)) = 1$ for all $y \in \overline{U}_{i,1}$).

We have learned how to correct the function g_{i-1} on the set $U_{i,1}$. Now, let us learn how to do this without loss of what has been achieved at the preceding steps. Namely, suppose that a function g_{i-1} has no degenerate critical points on the set $\bigcup_{j=1}^{i-1} \overline{U}_{j,1}$; we want g_i to have no degenerate critical points on this set as well. The function g_{i-1} is changed only on the compact set $\overline{U}_{i,2}$, and it has no degenerate critical points on the compact set $\overline{U}_{i,2} \cap \left(\bigcup_{j=1}^{i-1} \overline{U}_{j,1}\right)$.

Lemma 2. *Suppose that $f, g: U \to \mathbb{R}$ are smooth functions on an open set $U \subset \mathbb{R}^n$ and f has no degenerate critical points on a compact set $K \subset U$. Then there exists a number $\varepsilon > 0$ such that if all the first and second partial derivatives of the function $f - g$ at all points of K are less than ε in absolute value, then g has no degenerate critical points on K.*

Proof. The function

$$F = \sum \left(\frac{\partial f}{\partial x_i}\right)^2 + \left[\det\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)\right]^2$$

vanishes only at the degenerate critical points of f; therefore, the minimum value δ of F on the compact set K is positive. If ε it sufficiently small, then

$$\sum \left(\frac{\partial f}{\partial x_i}\right)^2 - \sum \left(\frac{\partial g}{\partial x_i}\right)^2 < \delta/2$$

and

$$\left[\det\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)\right]^2 - \left[\det\left(\frac{\partial^2 g}{\partial x_i \partial x_j}\right)\right]^2 < \delta/2;$$

hence

$$\sum\left(\frac{\partial g}{\partial x_i}\right)^2+\left[\det\left(\frac{\partial^2 g}{\partial x_i \partial x_j}\right)\right]^2>0,$$

which means that g has no degenerate critical points on K. □

If the numbers $a_1,\dots,a_n$ are sufficiently small, then the first and second derivatives of the function $\lambda(x)(a_1x_1+\dots+a_nx_n)$ are small also. Thus, the required function g_i can be constructed by applying Lemma 2. □

Second proof. We embed the manifold M^n in $\mathbb{R}^m$, take a point $a\in\mathbb{R}^m$, and define $f(x)=\|x-a\|^2$ for $x\in M^n$. Let $u_1,\dots,u_n$ be local coordinates on the manifold M^n, and let $x_i(u_1,\dots,u_n)$ $(i=1,\dots,m)$ be the coordinates of the point $(u_1,\dots,u_n)$ in $\mathbb{R}^m$. The functions x_i are smooth; therefore, the function f is smooth also. Our purpose is to choose a such that all critical points of f are nondegenerate.

Clearly, $\frac{\partial f}{\partial x_i}=2(x_i-a_i)$ and $\frac{\partial^2 f}{\partial x_i\partial x_j}=2\delta_{ij}$. Therefore,

$$\frac{\partial f}{\partial u_i}=\sum_{k=1}^m\frac{\partial f}{\partial x_k}\frac{\partial x_k}{\partial u_i}=2\sum_{k=1}^m(x_k-a_k)\frac{\partial x_k}{\partial u_i}$$

and

$$\begin{aligned}\frac{\partial^2 f}{\partial u_i\partial u_j}&=\sum_{k,l=1}^m\frac{\partial^2 f}{\partial x_k\partial x_l}\frac{\partial x_k}{\partial u_i}\frac{\partial x_l}{\partial u_j}+\sum_{k=1}^m\frac{\partial f}{\partial x_k}\frac{\partial^2 x_k}{\partial u_i\partial u_j}\\&=2\sum_{k=1}^m\left(\frac{\partial x_k}{\partial u_i}\frac{\partial x_k}{\partial u_j}+(x_k-a_k)\frac{\partial^2 x_k}{\partial u_i\partial u_j}\right).\end{aligned}$$

The vectors $e_i=\left(\frac{\partial x_1}{\partial u_i},\dots,\frac{\partial x_m}{\partial u_i}\right)$, where $i=1,\dots,n$, form a basis in the tangent space T_xM^n; hence a point $x\in M^n$ is critical for f if and only if the vector $\xi=x-a$ is orthogonal to T_xM^n, and this critical point is degenerate if and only if the matrix with elements $g_{ij}+(\xi,l_{ij})$, where $g_{ij}=(e_i,e_j)$ and $l_{ij}=\left(\frac{\partial^2 x_1}{\partial u_i\partial u_j},\dots,\frac{\partial^2 x_m}{\partial u_i\partial u_j}\right)$, is degenerate. Here g_{ij} and l_{ij} depend on the point $x\in M^n$ (and on the local coordinate system).

On p. 228, we considered the m-manifold N consisting of pairs (x,ξ), where $x\in M^n$ and ξ is a vector orthogonal to $T_xM^n\subset\mathbb{R}^m$. Let us show that a point $(x,\xi)\in N$ is critical for the map $(x,\xi)\mapsto x-\xi\in\mathbb{R}^n$ if and only if the matrix with elements $g_{ij}+(\xi,l_{ij})$ is degenerate. This would imply that the function $f(x)=\|x-a\|^2$ on the manifold M^n has a degenerate critical point if and only if a is a critical value for the map $(x,\xi)\mapsto x-\xi$, and, by Sard's theorem, $f(x)=\|x-a\|^2$ is a Morse function for almost every $a\in\mathbb{R}^m$.

We have already calculated the Jacobian matrix of the map $(x,\xi)\mapsto x+\xi$ (see p. 228). The Jacobian matrix J of the map $(x,\xi)\mapsto x-\xi$ is obtained by similar calculations:

$$\left(e_1 + \sum \xi_k \frac{\partial \varepsilon_k}{\partial u_1}, \dots, e_n + \sum \xi_k \frac{\partial \varepsilon_k}{\partial u_n}, -\varepsilon_1, \dots, -\varepsilon_{m-n}\right).$$

The vectors $e_1, \dots, e_n$ form a basis in the space $T_x M^n$, and the vectors $\varepsilon_1, \dots, \varepsilon_{m-n}$ form a basis in its orthogonal complement. Therefore, the matrix $A = (e_1, \dots, e_n, \varepsilon_1, \dots, \varepsilon_{m-n})$ is nonsingular, which means that the rank of J is equal to that of

$$A^T J = \begin{pmatrix} \left(g_{ij} - \left(e_i, \sum \xi_k \frac{\partial \varepsilon_k}{\partial u_j}\right)\right) & 0 \\ * & I_{m-n} \end{pmatrix}.$$

It remains to verify that $-\left(e_i, \sum \xi_k \frac{\partial \varepsilon_k}{\partial u_j}\right) = (\xi, l_{ij})$, where $l_{ij} = \frac{\partial^2 x}{\partial u_i \partial u_j}$. This equality is equivalent to $\left(\varepsilon_k, \frac{\partial^2 x}{\partial u_i \partial u_j}\right) = -\left(\frac{\partial \varepsilon_k}{\partial u_j}, \frac{\partial x}{\partial u_i}\right)$, i.e., $\frac{\partial}{\partial u_j}\left(\varepsilon_k, \frac{\partial x}{\partial u_i}\right) = 0$. But the vectors ε_k and $\frac{\partial x}{\partial u_i} = e_i$ are orthogonal, and hence $(\varepsilon_k, e_i) = 0$. □

Remark 5.7. If a manifold M^n is embedded in $\mathbb{R}^m$ as a closed subset and $f(x) = \|x - a\|^2$, then all sets of the form $\{x \in M^n | f(x) \leq c\}$ are compact.

A Morse function f is said to be *regular* if it takes different values at different critical points.

Theorem 5.39. *On any closed manifold M^n, a regular Morse function exists.*

Proof. Let $x_1, \dots, x_n$ be the critical points of a Morse function $f\colon M^n \to \mathbb{R}$. Choose pairwise disjoint neighborhoods $U_i \ni x_i$ and their open subsets $V_i \ni x_i$ for which there exist smooth functions $\varphi_i\colon M^n \to \mathbb{R}$ taking the value 1 on the set V_i and identically vanishing outside U_i. Consider the function $g(x) = f(x) + \varepsilon_1\varphi_1(x) + \dots + \varepsilon_k\varphi_k(x)$. The minimum value of the function $\left|\frac{\partial f}{\partial x_1}\right|^2 + \dots + \left|\frac{\partial f}{\partial x_n}\right|^2$ on the compact set $\overline{U}_i \setminus V_i$ is positive; therefore, if ε_i is sufficiently small, then $g(x)$ has no critical points belonging to $\overline{U}_i \setminus V_i$. The function g is a regular Morse function if the numbers $\varepsilon_1, \dots, \varepsilon_k$ are sufficiently small and the numbers $g(x_i) = f(x_i) + \varepsilon_i$ are pairwise different. □

5.2. Gradient Vector Fields and the Operation of Attaching Handles. Let f be a smooth function on a manifold M^n. If M^n is endowed with a Riemannian metric, then the function f determines the *gradient vector field* $\operatorname{grad} f$ characterized by the property that $(\operatorname{grad} f, v) = v(f)$ for any smooth vector field v on the manifold M^n; here $v(f)$ is the derivative of f in the direction of the vector field v. If $M^n = \mathbb{R}^n$ and the Riemannian metric is determined by the canonical inner product, then $\operatorname{grad} f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\right)$. This readily implies that the singular points of the vector field $\operatorname{grad} f$ coincide with those of the function f, and nondegenerate singular points correspond to nondegenerate critical points.

Theorem 5.40. *For any Riemannian metric and smooth function f on a manifold M^n, the index of a nondegenerate singular point x_0 of the vector field* grad f *is equal to $(-1)^i$, where i is the index of the critical point x_0 of f.*

Proof. First, we show that the index of a singular point of the vector field grad f does not depend on the choice of a Riemannian metric. Let $(v,w)_0$ and $(v,w)_1$ be two Riemannian metrics on M^n. Then the formula $(v,w)_t = t(v,w)_0+(1-t)(v,w)_1$, where $t \in [0,1]$, defines a continuous family of Riemannian metrics. The indices of singular points of the vector fields grad f (defined for these Riemannian metrics) continuously depend on t and take integer values; therefore, they do not depend on t.

For the function $f(x) = -x_1^2 - \cdots - x_i^2 + x_{i+1}^2 + \cdots + x_n^2$ on the space $\mathbb{R}^n$ with canonical inner product, the vector field grad f has the form $2(-x_1,\dots,-x_i,x_{i+1},\dots,x_n)$. The origin is a singular point of index $(-1)^i$ for this vector field (see the proof of Theorem 5.30 on p. 231). □

Corollary. *Let f be a Morse function on a closed manifold M^n. Then the alternated sum $\sum_{i=1}^n(-1)^i c_i$, where c_i is the number of critical points of index i, does not depend on the choice of f.*

Proof. The alternating sum under consideration is equal to the sum of the indices of singular points of the vector field grad f, and the sums of the indices of singular points are the same for all vector fields on a given closed manifold. □

The topological structure of a closed manifold M^n is largely determined by the set of indices of critical points for a regular Morse function f. We give precise statements below. The main objects of study in Morse theory are the sets $M_a = \{x \in M^n : f(x) \le a\}$ and the level surfaces $f^{-1}(a)$. Morse theory studies what happens to them when a passes through critical values. Note that if a is not a critical value, then M_a is a manifold.

Theorem 5.41. *Suppose that an interval $[a,b]$ contains no critical values of a Morse function f on a closed manifold M^n. Then the manifolds M_a and M_b are diffeomorphic; in particular, the level surfaces $f^{-1}(a)$ and $f^{-1}(b)$ are diffeomorphic. Moreover, the manifold $f^{-1}([a,b])$ is diffeomorphic to $f^{-1}(a) \times [a,b]$.*

Proof. Choose $\varepsilon > 0$ such that the interval $[a-\varepsilon, b+\varepsilon]$ contains no critical values of f. Let $\lambda(s)$ be a smooth function which takes the value 1 at $s \in [a,b]$ and vanishes at $s \notin [a-\varepsilon, b+\varepsilon]$. If $f(x) \in [a-\varepsilon, b+\varepsilon]$, then the vector field grad $f/\|$grad $f\|^2$ exists (we assume that the manifold M^n

is endowed with a Riemannian metric). Consider the vector field

$$v(x) = \frac{\lambda(f(x)) \operatorname{grad} f}{\|\operatorname{grad} f\|^2},$$

which is defined on the entire manifold M^n. Note that $v(x) = 0$ if $f(x) \notin [a-\varepsilon, b+\varepsilon]$.

The smooth vector field v on the compact manifold M^n determines the so-called *integral curves*, i.e., curves $\gamma(x,t)$ for which $\gamma(x,0) = x$ and $\frac{\partial\gamma(x,t)}{\partial t} = v(\gamma(x,t))$. The latter equality means that the tangent vector at the point $\gamma(x,t)$ to the curve $\omega(\tau) = \gamma(x, t+\tau)$ is $v(\gamma(x,t))$. In other words, if $g\colon M^n \to \mathbb{R}$ is an arbitrary smooth function, then, at the point $x \in M^n$, the operator v assigns to this function the number $\frac{\partial g(\gamma(x,t))}{\partial t}\big|_{t=0}$. Thus,

$$\left.\frac{\partial g(\gamma(x,t))}{\partial t}\right|_{t=0} = v(g) = (v, \operatorname{grad} g).$$

Taking the initial function f for g, we see that if $f(x) \in [a,b]$, then

$$\left.\frac{\partial f(\gamma(x,t))}{\partial t}\right|_{t=0} = \left(\frac{\operatorname{grad} f}{\|\operatorname{grad} f\|^2}, \operatorname{grad} f\right) = 1. \tag{1}$$

Consider the map $\varphi_t\colon M^n \to M^n$ defined by $\varphi_t(x) = \gamma(x,t)$. It has the properties $\varphi_0 = \mathrm{id}_{M^n}$ and $\varphi_{t+s} = \varphi_t\varphi_s$. Therefore, φ_t is a diffeomorphism. Formula (1) shows that $\varphi_{b-a}(M_a) = M_b$.

A diffeomorphism between the manifolds $f^{-1}(a) \times [a,b]$ and $f^{-1}([a,b])$ is defined by $(x,t) \mapsto \gamma(x, t-a)$. □

Using Theorem 5.41, we can describe the topological structure of a closed manifold in the case where this manifold admits a Morse function with exactly two critical points (a maximum and a minimum).

Theorem 5.42. *Suppose that on a closed manifold M^n, there exists a Morse function f having exactly two critical points. Then the manifold M^n is homeomorphic to S^n.*[6]

Proof. Let $f_{\max}$ and $f_{\min}$ be the maximum and minimum values of the function f. According to the Morse lemma, there exists a local coordinate system with origin at the point of maximum in which the function f has the form $f(x_1, \dots, x_n) = f_{\max} - x_1^2 - \dots - x_n^2$. Therefore, we can choose $\varepsilon > 0$ so that the level surface $f^{-1}(f_{\max} - \varepsilon)$ is diffeomorphic to S^{n-1} and the inequality $f(x) \geq f_{\max} - \varepsilon$ determines a manifold diffeomorphic to D^n. We assume that ε is such that similar conditions hold at the point of minimum.

[6] Milnor [85] showed that for $n \geq 7$, M^n may not be diffeomorphic to S^n. This means that for $n \geq 7$, the sphere S^n may admit nonisomorphic (or nonstandard) differentiable structures. There are exceptions; for example, the sphere S^{12} admits no nonstandard smooth structures.

The function f has no critical values between the points $f_{\min} + \varepsilon$ and $f_{\max} - \varepsilon$. Therefore, according to Theorem 5.41, the preimage of the interval $[f_{\min} + \varepsilon, f_{\max} - \varepsilon]$ is diffeomorphic to $S^{n-1} \times I$. Hence the manifold M^n is obtained from $S^{n-1} \times I$ by attaching two copies of D^n via some diffeomorphisms $\varphi_1 \colon S^{n-1} \to S^{n-1}$ and $\varphi_2 \colon S^{n-1} \to S^{n-1}$ of the boundaries. It is easy to show that such a manifold is homeomorphic to S^n. Indeed, if $\varphi_1 = \varphi_2 = \mathrm{id}_{S^{n-1}}$, then this assertion is obvious. It remains to show that any diffeomorphism $\varphi \colon S^{n-1} \to S^{n-1}$ can be extended to a homeomorphism $\Phi \colon D^n \to D^n$. For $x \in D^n$, we set

$$\Phi(x) = \begin{cases} x\varphi(x/\|x\|) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

At the point 0, the map Φ is continuous but not differentiable. □

Now consider the case in which there is precisely one critical point between the level surfaces $f^{-1}(a)$ and $f^{-1}(b)$.

Theorem 5.43. *Suppose that x_0 is a nondegenerate critical point of index i of a smooth function f and the interval $[a, b] = [f(x_0) - \varepsilon, f(x_0) + \varepsilon]$ contains no images of other critical points. Then the space M_b is homotopy equivalent to the space obtained from M_a by attaching a disk D^i via a map $\partial D^i \to M_a$.*

Proof. Theorem 5.41 allows us to assume the number ε to be arbitrarily small: the manifolds M_a and M_b can change their structures only when a critical value of f is passed through. Applying the Morse lemma, we choose local coordinates with origin x_0 such that in these coordinates we have $f(x_1, \ldots, x_n) = f(x_0) - x_1^2 - \cdots - x_i^2 + x_{i+1}^2 + \cdots + x_n^2$. In the chosen local coordinate system, the intersection of the level surface $f(x) = f(x_0) - \varepsilon$ with the linear subspace generated by the first i coordinates is the i-disk $x_1^2 + \cdots + x_i^2 \leq \varepsilon$; we denote it by D^i_ε. The coordinates of the points of the surface $f(x) = f(x_0) + \varepsilon$ which are projected to this disk satisfy the inequality $x_{i+1}^2 + \cdots + x_n^2 \leq 2\varepsilon$; we denote the corresponding $(n-i)$-disk by $D^{n-i}_{2\varepsilon}$ (in Figure 12, the case of $n = 2$ and $i = 1$ is shown). We assume that ε is so small that $D^i_\varepsilon \times D^{n-i}_{2\varepsilon}$ is contained entirely in the chosen coordinate neighborhood $U \ni x_0$.

Let us construct a deformation retraction $r \colon M_b \to A$, where $A = M_a \cup D^i_\varepsilon$ (and $M_a \cap D^i_\varepsilon = \partial D^i_\varepsilon$). To simplify the notation, we write coordinates $(x_1, \ldots, x_n)$ in the form (x_-, x_+), where $x_- = (x_1, \ldots, x_i)$ and $x_+ = (x_{i+1}, \ldots, x_n)$. On the set $D^i_\varepsilon \times D^{n-i}_{2\varepsilon} \cap M_b$, the map r is defined as follows. Suppose that u is the vector field that takes the value $(0, -x_+)$ at each point $(x_-, x_+) \in D^i_\varepsilon \times D^{n-i}_{2\varepsilon} \cap M_b$. Then r sends x to the endpoint of the closure of the integral curve of u passing through x (see Figure 12).

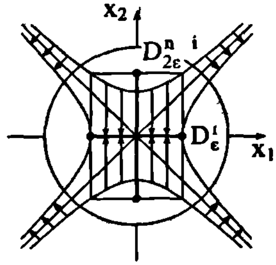

Figure 12. The deformation retraction

Outside the coordinate neighborhood U, the map r is defined similarly by using the vector field $v(x) = -\frac{1}{2}\operatorname{grad} f(x)$. Note that if $x \in U$, then $v(x) = (x_-, -x_+)$.

To define the map r on the entire set M_b, we need to construct a vector field which coincides with u on $D^i_\varepsilon \times D^{n-i}_{2\varepsilon}$ and with v outside U. Let $\lambda\colon M^n \to [0,1]$ be a smooth function vanishing on $D^i_\varepsilon \times D^{n-i}_{2\varepsilon}$ and taking the value 1 outside U. For $x \in U$, we set $w(x) = (\lambda(x)x_-, -x_+)$, and for $x \notin U$, we set $w(x) = v(x)$. On the integral curves of the vector field w, the absolute values of the coordinates x_+ decrease; therefore, all these integral curves reach the level surface $f(x) = f(x_0) - \varepsilon$ (we mean the integral curves which start in M_b outside M_a and outside $D^i_\varepsilon \times D^{n-i}_{2\varepsilon}$).

A homotopy between id_{M_b} and $r\colon M_b \to A \subset M_b$ is constructed as follows. For each point $x \in M_b$, we consider the segment of the integral curve from x to $r(x)$ and take the point dividing it in the ratio $t : (1-t)$. $\square$

Theorems 5.41 and 5.43 imply the following important theorem, which relates the indices of critical points of Morse functions to CW-structures on manifolds.

Theorem 5.44. *Suppose that f is a Morse function on a closed manifold M^n which has c_i critical points of index i ($i = 0, 1, \dots, n$). Then M^n is homotopy equivalent to a CW-complex with c_i i-cells.*

Proof. Suppose that x_0 is a critical point of index i and the interval $[a, b] = [f(x_0) - \varepsilon, f(x_0) + \varepsilon]$ contains no images of other critical points. It is sufficient to show that if the space M_a is homotopy equivalent to a CW-complex X, then the space M_b is homotopy equivalent to the space $X \cup_\chi D^i$, where $\chi\colon S^{i-1} = \partial D^i \to X$ is a cellular map.

According to Theorem 5.43, the space M_b is homotopy equivalent to $M_a \cup_\varphi D^i$, where $\varphi\colon S^{i-1} \to M_a$ is a map. Let $h\colon M_a \to X$ be a homotopy equivalence. For χ we take a cellular approximation of the map

$h\varphi\colon S^{i-1} \to X$. It remains to prove the following two assertions (the space Y in the lemma plays the role of M_a).

Lemma. (a) *If maps $\chi_0, \chi_1\colon S^{i-1} \to X$ are homotopic, then the spaces $X \cup_{\chi_0} D^i$ and $X \cup_{\chi_1} D^i$ are homotopy equivalent.*

(b) *If $\varphi\colon S^{i-1} \to Y$ is a map and $h\colon Y \to X$ is a homotopy equivalence, then the spaces $Y \cup_\varphi D^i$ and $X \cup_{h\varphi} D^i$ are homotopy equivalent.*

Proof. (a) Let χ_t be a homotopy between χ_0 and χ_1. Consider the map $\alpha\colon X_0 = X \cup_{\chi_0} D^i \to X \cup_{\chi_1} D^i = X_1$ defined as follows. For $x \in X$, we set $\alpha(x) = x$. Each point $u \in S^{i-1}$ is identified with $\chi_0(u)$ in the space X_0; therefore, the equalities $\alpha(u) = \alpha(\chi_0(u)) = \chi_0(u)$ must hold. For $t \in [0, 1/2]$ and $u \in S^{i-1}$, we set $\alpha(tu) = 2tu$. Each point $u \in S^{i-1}$ is identified with $\chi_1(u)$ in the space X_1; hence we must define α so that $\alpha(u/2) = u = \chi_1(u)$. To ensure that $\alpha(u/2) = \chi_1(u)$ and $\alpha(u) = \chi_0(u)$, we set $\alpha(tu) = \chi_{2-2t}(u)$ for $t \in [1/2, 1]$ and $u \in S^{i-1}$.

A map $\beta\colon X_1 \to X_0$ is defined similarly. In this case, the equalities $\beta(u/2) = \chi_0(u)$ and $\beta(u) = \chi_1(u)$ must hold, and we set $\beta(tu) = \chi_{2t-1}(u)$ for $t \in [1/2, 1]$ and $u \in S^{i-1}$.

It remains to verify that $\beta\alpha \sim \mathrm{id}_{X_0}$ and $\alpha\beta \sim \mathrm{id}_{X_1}$. Direct calculations give

$$\beta(\alpha(tu)) = \begin{cases} 4tu & \text{if } t \in [0, 1/4], \\ \chi_{4t-1}(u) & \text{if } t \in [1/4, 1/2], \\ \chi_{2-2t}(u) & \text{if } t \in [1/2, 1]. \end{cases}$$

Informally, a homotopy joining this map to the identity map can be described as follows. We uniformly expand the interval $[0, 1/4]$ to $[0, 1]$ and leave the map linear on this interval. Simultaneously, on the remaining part of the interval, we replace $\chi_{k(t)}$ by $\chi_{(1-s)k(t)}$, where $s \in [0, 1]$ (at the endpoints, the map χ_0 is given; we want to obtain it at $s = 1$).

A homotopy between $\alpha\beta$ and the identity map is constructed similarly.

(b) Let $g\colon X \to Y$ be a homotopy inverse to h. Consider the maps $H\colon Y \cup_\varphi D^i \to X \cup_{h\varphi} D^i$ and $G\colon X \cup_{h\varphi} D^i \to Y \cup_{gh\varphi} D^i$ defined by the conditions $H \restriction Y = h$, $H \restriction D^i = \mathrm{id}_{D^i}$ and $G \restriction X = g$, $G \restriction D^i = \mathrm{id}_{D^i}$.

Since $gh \sim \mathrm{id}_Y$, we have $gh\varphi \sim \varphi$. Therefore, according to the assertion (a) proved above, there exists a homotopy equivalence $\alpha\colon Y \cup_{gh\varphi} D^i \to Y \cup_\varphi D^i$. We show that the map αGH is homotopic to the identity. By construction, $\alpha GH(y) = y$ for $y \in Y$ and

$$\alpha GH(tu) = \begin{cases} 2tu & \text{if } t \in [0, 1/2] \text{ and } u \in S^{i-1}, \\ \chi_{2-2t}\varphi(u) & \text{if } t \in [1/2, 1] \text{ and } u \in S^{i-1}; \end{cases}$$

here χ_t is the homotopy between gh and id_Y.

The construction of a homotopy between αGH and the identity map largely repeats the construction from (a). The main difference is that this time, there are no constraints on the image of the right endpoint of the interval $[0, 1]$. Hence for the map of the interval remaining after the expansion of $[0, 1/2]$ we can take the initial map of some left segment of the interval $[0, 1]$. Formally, this map can be defined as

$$\psi_s(tu) = \begin{cases} \frac{2}{1+s}tu & \text{if } t \in \left[0, \frac{1+s}{2}\right] \text{ and } u \in S^{i-1}, \\ \chi_{2-2t+s}\varphi(u) & \text{if } t \in \left[\frac{1+s}{2}, 1\right] \text{ and } u \in S^{i-1}. \end{cases}$$

The linear function $2-2t+s$ is chosen because it takes the value 1 at $t = \frac{1+s}{2}$.

Thus, $\alpha GH \sim \mathrm{id}$. Similarly, $\beta HG \sim \mathrm{id}$, where β is the homotopy inverse of α. Therefore,

$$GH\alpha \sim \beta\alpha(GH\alpha) = \beta(\alpha GH)\alpha \sim \beta\alpha \sim \mathrm{id}.$$

We obtain

$$H\alpha G \sim (\beta HG)H\alpha G = \beta H(GH\alpha)G \sim \beta HG \sim \mathrm{id}.$$

The relations $GH\alpha \sim \mathrm{id}$ and $H\alpha G \sim \mathrm{id}$ mean that the maps H and αG are homotopy inverse to each other; in particular, H is a homotopy equivalence. □

Problem 86. (a) Suppose that Y is a contractible subcomplex in a CW-complex X, i.e., the embedding $Y \to X$ is homotopic to a constant map. Prove that $X/Y \sim X \vee \Sigma Y$.

(b) Suppose that $m < n$ and the sphere S^m is canonically embedded in S^n. Prove that $S^n/S^m \sim S^n \vee S^{m+1}$.

5.3. Examples of Morse Functions.

Torus. Let us represent the torus T^n as the quotient space $\mathbb{R}^n/2\pi\mathbb{Z}^n$. Then the smooth functions on T^n are the smooth functions of n variables that are 2π-periodic in each variable.

Example 1. Let $c_1, \ldots, c_n$ be real numbers. The function $f(x_1, \ldots, x_n) = c_1 \sin x_1 + \cdots + c_n \sin x_n$ is a Morse function on T^n if and only if $c_1, \ldots, c_n \neq 0$. This function has $\binom{n}{k}$ critical points of index k.

Proof. The equalities $\frac{\partial f}{\partial x_1} = \cdots = \frac{\partial f}{\partial x_n} = 0$ imply $c_1 \cos x_1 = \cdots = c_n \cos x_n = 0$. If $c_i \neq 0$, then $x_i = \pm\frac{\pi}{2} + 2m\pi$. Clearly, if $c_i = 0$, then the function f has a nonisolated critical point. If $c_1, \ldots, c_n \neq 0$, then we have

$$\left(\frac{\partial f^2}{\partial x_i \partial x_j}\right) = \mathrm{diag}(-c_1 \sin x_1, \ldots, -c_n \sin x_n) = \mathrm{diag}(\varepsilon_1 c_1, \ldots, \varepsilon_n c_n)$$

at this critical point; here $x_i = -\varepsilon_i \frac{\pi}{2} + 2m\pi$. The index of the critical point is equal to the number of those i for which $\varepsilon_i c_i < 0$. The critical points of index k are determined by k numbers i for which $\varepsilon_i c_i < 0$. □

Note that if S^1 is represented as a CW-complex with one 0-cell and 1-cell, then the torus $T^n = S^1 \times \cdots \times S^1$ is a CW-complex with $\binom{n}{k}$ k-cells. The Morse function from Example 1 determines precisely this partition of the torus into cells.

Sphere. We represent the sphere S^n as the submanifold in $\mathbb{R}^n$ determined by the equation $x_1^2 + \cdots + x_{n+1}^2 = 1$.

Example 2. The function $f(x_1, \dots, x_{n+1}) = x_{n+1}$ is a Morse function on S^n, and it has two critical points, one of index 0 and one of index n.

Proof. The sphere S^n can be covered by $2(n+1)$ charts, each of which is determined by the inequality $x_i > 0$ or $x_i < 0$; the local coordinates for such a chart are $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1})$. If $i \neq n+1$, then the function f is smooth and has no critical points on the corresponding chart. On the chart $x_{n+1} > 0$, f has the form $\sqrt{1 - x_1^2 - \cdots - x_n^2}$; here $x_1^2 + \cdots + x_n^2 < 1$ and the positive value of the root is taken.

A simple calculation shows that $\frac{\partial f}{\partial x_1} = \cdots = \frac{\partial f}{\partial x_n} = 0$ only at the point $(x_1, \dots, x_n) = (0, \dots, 0)$; at this point, the matrix $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)$ is equal to $-I_n$. Thus, the point $(x_1, \dots, x_n, x_{n+1}) = (0, \dots, 0, 1) \in S^n$ has index n.

A similar calculation for the chart $x_{n+1} < 0$ shows that at the point $(0, \dots, 0, -1)$, the Hessian matrix is I_n, i.e., this point has index 0. □

Example 3. Let $c_1, \dots, c_{n+1}$ be real numbers. The function $f(x_1, \dots, x_{n+1}) = c_1 x_1^2 + \cdots + c_{n+1} x_{n+1}^2$ is a Morse function on S^n if and only if the numbers $c_1, \dots, c_{n+1}$ are pairwise distinct. This Morse function has two critical points of each index $0, 1, 2, \dots, n$.

Proof. On the charts $x_{n+1} > 0$ and $x_{n+1} < 0$, the function f has the form

$$\begin{aligned} f &= c_1 x_1^2 + \cdots + c_n x_n^2 + c_{n+1}(1 - x_1^2 - \cdots - x_n^2) \\ &= (c_1 - c_{n+1})x_1^2 + \cdots + (c_n - c_{n+1})x_n^2 + c_{n+1}. \end{aligned}$$

Such a function has only one critical point, namely, $(x_1, \dots, x_n) = (0, \dots, 0)$. This critical point is nondegenerate if and only if all the numbers $c_1, \dots, c_n$ are different from c_{n+1}; the index of a nondegenerate critical point is equal to the number of those $c_1, \dots, c_n$ which are smaller than c_{n+1}. The points $(x_1, \dots, x_n) = (0, \dots, 0)$ on the charts $x_{n+1} > 0$ and $x_{n+1} < 0$ correspond to the points $(x_1, \dots, x_n, x_{n+1}) = (0, \dots, 0, \pm 1) \in S^n$.

Performing similar calculations for the other charts $x_i > 0$ and $x_i < 0$, we see that critical points are nondegenerate if and only if the numbers $c_1, \dots, c_{n+1}$ are pairwise distinct. Clearly, for each $k = 0, 1, \dots, n$, there exists precisely one i for which k of the numbers $c_1, \dots, c_{i-1}, c_{i+1}, \dots, c_{n+1}$ are less than c_i. $\square$

Real projective space. We represent the real projective space $\mathbb{R}P^n$ as the manifold obtained from the sphere S^n by identifying x and $-x$.

Let f be a Morse function on S^n such that $f(-x) = f(x)$. Then the function f can be regarded as a function on $\mathbb{R}P^n$, and it remains a Morse function on $\mathbb{R}P^n$. Moreover, each critical point of f as a function on $\mathbb{R}P^n$ corresponds to two critical points (with the same index) of f as a function on S^n.

The function $f(x_1, \dots, x_{n+1}) = c_1 x_1^2 + \dots + c_{n+1} x_{n+1}^2$ does have the required property; therefore, the following assertion is valid.

Example 4. The function $f(x_1, \dots, x_{n+1}) = c_1 x_1^2 + \dots + c_{n+1} x_{n+1}^2$ is a Morse function on $\mathbb{R}P^n$ if and only if the numbers $c_1, \dots, c_{n+1}$ are pairwise distinct. This Morse function has one critical point of each of the indices $0, 1, 2, \dots, n$.

Complex projective space. We specify points of the complex projective space $\mathbb{C}P^n$ by homogeneous coordinates $(z_1 : \dots : z_n)$.

Example 5. Let $c_1, \dots, c_{n+1}$ be pairwise distinct real numbers. Then the function

$$f(z_1 : \dots : z_{n+1}) = \frac{c_1 |z_1|^2 + \dots + c_{n+1} |z_{n+1}|^2}{|z_1|^2 + \dots + |z_{n+1}|^2}$$

is a Morse function on $\mathbb{C}P^n$. This Morse function has one critical point of each of the indices $0, 2, \dots, 2n$.

Proof. The manifold $\mathbb{C}P^n$ can be covered by $n+1$ charts, each determined by an inequality of the form $z_k \neq 0$. On the chart $z_{n+1} \neq 0$, we define local coordinates by $w_k = z_k / z_{n+1}$ for $k = 1, \dots, n$ (each complex coordinate w_k corresponds to two real coordinates). In these local coordinates, the function f has the form

$$f(w_1, \dots, w_n) = \frac{c_1 |w_1|^2 + \dots + c_n |w_n|^2 + c_{n+1}}{|w_1|^2 + \dots + |w_n|^2 + 1}.$$

It is more convenient to pass to the coordinates $u_k = w_k / \sqrt{\|w\|^2 + 1}$, where $\|w\|^2 = |w_1|^2 + \dots + |w_n|^2$. Using the equalities

$$\frac{1}{\|w\|^2 + 1} = 1 - \frac{\|w\|^2}{\|w\|^2 + 1} = 1 - |u_1|^2 - \dots - |u_n|^2,$$

we obtain

$$f(u_1, \dots, u_n) = (c_1 - c_{n+1})|u_1|^2 + \dots + (c_n - c_{n+1})|u_n|^2 + c_{n+1}.$$

Moreover, $u_k = x_k + iy_k$ and $|u_k|^2 = x_k^2 + y_k^2$.

The rest of the proof is as in Example 3. In the complex case, the indices of critical points are twice the indices in the real case because $|u_k|^2 = x_k^2 + y_k^2$. □

The manifolds SO(n) and U(n). Our construction of Morse functions on SO(n) and U(n) is borrowed from [**133**]. First, we describe the structure of the tangent spaces to SO(n) and U(n) at a point X (by definition, $XX^* = I_n$ is the identity matrix, where $X^* = \overline{X}^T$; in the real case, complex conjugation can be omitted). If ξ is a tangent vector, then the matrix $X + t\xi$ must satisfy the relation $(X + t\xi)(X^* + t\xi^*) = I_n$ up to terms of order t; therefore, $\xi X^* + X\xi^* = 0$. At the point $X = I_n$, this condition takes the form $\xi + \xi^* = 0$; i.e., the matrix ξ is skew-Hermitian (skew-symmetric in the real case) at this point. The dimensions of the spaces of skew-symmetric and skew-Hermitian matrices are easy to calculate; they coincide with those of SO(n) and U(n). Clearly, for the matrix $\tilde{\xi} = \xi X$, the equality $\tilde{\xi} X^* + X\tilde{\xi}^* = 0$ is equivalent to $\xi + \xi^* = 0$. Hence any matrix ξ for which $\xi X^* + X\xi^* = 0$ belongs to the tangent space at X.

Lemma. *Let η be a matrix of order n, i.e., a vector in $\mathbb{R}^{n^2}$. Suppose that the space of matrices in endowed with the inner product $(A, B) = \operatorname{tr}(AB^*)$. Then the orthogonal projection of the vector η on the tangent space to* SO(n) *or* U(n) *at a point X is $\frac{1}{2}(\eta - X\eta^* X)$.*

Proof. First, consider the case of $X = I_n$. In this case, the orthogonal complement of the skew-Hermitian (skew-symmetric) matrices consists of Hermitian (symmetric) matrices. Therefore, it suffices to represent the matrix η as a sum of Hermitian and skew-Hermitian matrices: $\eta = \frac{1}{2}(\eta + \eta^*) + \frac{1}{2}(\eta - \eta^*)$. The matrix $\frac{1}{2}(\eta - \eta^*)$ is the projection of η on the tangent space, and $\frac{1}{2}(\eta + \eta^*)$ is the projection of η on the orthogonal complement of the tangent space.

For an arbitrary point X, we proceed as follows. First, we translate the vector η to the point I_n: $\eta \mapsto \eta X^{-1} = \eta X^*$. Then, we find the projection of ηX^* on the tangent space. It is equal to $\frac{1}{2}(\eta X^* - X\eta^*)$. Finally, we return to the initial tangent space: $\frac{1}{2}(\eta X^* - X\eta^*) \mapsto \frac{1}{2}(\eta X^* - X\eta^*)X = \frac{1}{2}(\eta - X\eta^* X)$. □

Example 6. On the space SO(n) or U(n), consider the function $f_A(X) = \operatorname{Re}\operatorname{tr}(AX)$, where A is a given matrix. A point X is critical for this function if and only if $A^* = XAX$.

Proof. We have $f_A(X) = \sum_{pq}(\alpha_{pq}u_{qp} - \beta_{pq}v_{qp})$, where $a_{pq} = \alpha_{pq} + i\beta_{pq}$ and $x_{pq} = u_{pq} + iv_{pq}$; hence $\operatorname{grad} f = A^*$ is a constant matrix.

The critical points are the points for which $\operatorname{grad} f$ has zero projection on the tangent space. According to the lemma on p. 252, this projection is equal to $\frac{1}{2}(A^* - XAX)$. $\square$

Example 7. Let $A = \operatorname{diag}(a_1, \dots, a_n)$, where $0 \le a_1 < \dots < a_n$. Then $f_A(X)$ is a Morse function on $\mathrm{SO}(n)$ (or on $\mathrm{U}(n)$).

Proof. According to Example 6, a point X is critical if and only if $A^* = XAX$. In this case, $A = (XAX)^* = X^{-1}AX^{-1}$. Therefore,

$$A^2 = (XAX)(X^{-1}AX^{-1}) = XA^2X^{-1},$$

i.e., the matrices A^2 and X commute. Moreover, A is a diagonal matrix with different eigenvalues. Therefore, X is a diagonal matrix (see [100, Problem 39.1a]). By assumption, X is unitary, whence $X = \operatorname{diag}(\pm 1, \dots, \pm 1)$; in the real case, the product of diagonal elements equals 1.

We have found the critical points of the function f_A. Now, we must verify that all of them are nondegenerate.

Let ξ be a tangent vector at the point I_n (i.e., $\xi^* + \xi = 0$), and let $X_0 = \operatorname{diag}(\varepsilon_1, \dots, \varepsilon_n)$, where $\varepsilon_i = \pm 1$, be a critical point. Consider the map $\xi \mapsto X_0 e^{\xi}$; for the local coordinates near X_0 we take the superdiagonal and diagonal elements of the matrix ξ. In these coordinates,

$$\begin{aligned} f_A(X) - f_A(X_0) &= \operatorname{Re}\operatorname{tr} A(X - X_0) = \operatorname{Re}\operatorname{tr} AX_0(e^{\xi} - I_n) \\ &= \operatorname{Re}\operatorname{tr} AX_0\left(\xi + \frac{1}{2}\xi^2 + \cdots\right) \\ &= \operatorname{Re}\left(\sum_p a_p\varepsilon_p\xi_{pp} + \frac{1}{2}\sum_p a_p\varepsilon_p\xi_{pj}\xi_{jp} + \cdots\right). \end{aligned}$$

The numbers ξ_{pp} are purely imaginary; therefore, the real part of the first sum vanishes. Moreover, $\xi_{pj} = -\bar{\xi}_{jp}$; hence all the terms in the second sum are real, and $f_A(X) - f_A(X_0) = -\sum_{1 \le p \le q \le n}(a_p\varepsilon_p + a_q\varepsilon_q)|\xi_{pq}|^2 + \cdots$. By assumption, $a_q > a_p \ge 0$ for $q > p$; therefore, $\operatorname{sign}(a_p\varepsilon_p + a_q\varepsilon_q) = \operatorname{sign}(a_q\varepsilon_q)$ for $q \ge p$. Thus, if $p \le q$, then the sign of the term with $|\xi_{pq}|^2$ is opposite to that of ε_q. In particular, the quadratic part of $f_A(X)$ contains the squares of all coordinates. $\square$

Grassmann manifolds. We use the notation and the properties of Grassmann manifolds from Section 1.5. Of special importance for our purposes in this example is the Schubert symbol $\sigma = (\sigma_1, \dots, \sigma_k)$, where $1 \le \sigma_1 < \dots < \sigma_k \le n$, and the number $d(\sigma) = (\sigma_1 - 1) + (\sigma_2 - 2) + \dots + (\sigma_k - k)$, which is equal to the dimension of the open Schubert cell $e(\sigma)$.

The Schubert symbol is a special name for the multi-index used to describe the cell structure of a Grassmann manifold. Each Schubert symbol

σ is associated with the Plücker coordinate x_σ. The set of Plücker coordinates of a subspace $\Pi \in G(n,k)$ is a point in the projective space $\mathbb{R}P^{\binom{n}{k}-1}$. Let e_σ be the point of this projective space that corresponds to the σ-axis; we have $e_\sigma = (0 : 0 : \cdots : 0 : 1 : 0 : \cdots : 0)$, where 1 occupies the position corresponding to the multi-index σ. Clearly, the point e_σ belongs to the image of the open Schubert cell $e(\sigma)$.

Example 8. Consider the function $f(x) = \sum c_\sigma x_\sigma^2$ on the space $\mathbb{R}P^{\binom{n}{k}-1}$ (it is assumed that $c_\sigma \in \mathbb{R}$ and $\sum x_\sigma^2 = 1$). Let $i\colon G(n,k) \to \mathbb{R}P^{\binom{n}{k}-1}$ be the Plücker embedding. Then the numbers c_σ can be chosen so that fi is a Morse function on $G(n,k)$ with critical points $i^{-1}(e_\sigma)$ of index $d(\sigma)$. (The choice of the numbers c_σ is constructive; it is described in the course of the proof.)

Proof. If the numbers c_σ are pairwise distinct, then, according to Example 4, the function f is a Morse function on $\mathbb{R}P^{\binom{n}{k}-1}$ with critical points e_σ. Therefore, all points $i^{-1}(e_\sigma)$ are critical for fi. Let us find out how the function fi behaves in a neighborhood of the point $i^{-1}(e_\sigma)$.

The multi-index σ corresponds to the chart U_σ on the manifold $G(n,k)$, and the point $i^{-1}(e_\sigma)$ is the origin of this chart. In these local coordinates, the Plücker embedding has the following structure. The Plücker coordinate x_τ is equal to the determinant of the matrix formed by the columns $\tau_1, \ldots, \tau_k$ of the matrix in which the columns $\sigma_1, \ldots, \sigma_k$ constitute the identity matrix and the remaining columns are filled with the numbers $y_1, \ldots, y_{k(n-k)}$. Clearly, x_τ is a homogeneous polynomial in the variables $y_1, \ldots, y_{k(n-k)}$ of degree equal to the number of columns $\tau_1, \ldots, \tau_k$ different from the columns $\sigma_1, \ldots, \sigma_k$ (we have $x_\sigma = 1$). In calculating the Hessian matrix of the function $\sum \frac{c_\tau x_\tau^2}{x_\tau^2}$ at the origin, we are interested only in the linear polynomials x_τ. We say that Schubert symbols σ and τ (of length k) are *neighboring* if they have precisely $k-1$ common elements. It is easy to verify that if σ and τ are neighboring Schubert symbols, then $x_\tau = \pm y_i$. Moreover, for each index i, there is exactly one Schubert symbol $\tau(i)$ neighboring σ.

At the origin, the function f takes the value c_σ, and

$$f(y) - c_\sigma = \frac{c_\sigma + \sum c_\tau x_\tau^2(y)}{1 + \sum x_\tau^2(y)} - c_\sigma = \frac{\sum (c_\tau - c_\sigma) x_\tau^2(y)}{1 + \sum x_\tau^2(y)}.$$

Therefore, the quadratic form that approximates the difference $f(y) - c_\sigma$ is $\sum_{i=1}^{k(n-k)} (c_{\tau(i)} - c_\sigma) x_{\tau(i)}^2(y)$; here the summation is over the Schubert symbols $\tau(i)$ neighboring σ. If the numbers c_τ are pairwise distinct, then the critical point $i^{-1}(e_\sigma)$ is nondegenerate; its index equals the number of the Schubert symbols $\tau(i)$ (neighboring σ) for which $c_{\tau(i)} < c_\sigma$.

We order the Schubert symbols by declaring that $\tau < \sigma$ if $\tau_k = \sigma_k$, $\tau_{k-1} = \sigma_{k-1}, \ldots, \tau_{i+1} = \sigma_{i+1}$, and $\tau_i < \sigma_i$. Take constants c_τ such that

$c_\tau < c_\sigma$ for $\tau < \sigma$. It is easy to verify that for each Schubert symbol σ, there are $d(\sigma)$ neighboring Schubert symbols τ smaller than σ. Indeed, consider an element σ_i of σ. To obtain a Schubert symbol smaller that σ, we should replace it by any positive integer smaller than σ_i and different from $\sigma_1, \dots, \sigma_{i-1}$; the number of such integers equals $\sigma_i - i$.

Thus, for the chosen numbers c_σ, $i^{-1}(e_\sigma)$ is a nondegenerate critical point of index $d(\sigma)$. It remains to show that the function fi has no other critical points.

Consider any k-dimensional subspace Π; let σ be its Schubert symbol. In Π we take the basis $v_1, \dots, v_k$, where $v_i = (v_{i1}, \dots, v_{i\sigma_i - 1}, 1, 0, \dots, 0)$ and $v_{i\sigma_j} = 0$ for $j < i$. We say that a vector v_i has a nonzero coordinate if $v_{is} \neq 0$ for some $s < \sigma_i$ (the coordinate $v_{i\sigma_i} = 1$ is ignored). We assume that $\Pi \neq i^{-1}(e_\sigma)$, i.e., at least one of the vectors $v_1, \dots, v_k$ has a nonzero coordinate. Choose such a vector v_i with largest possible i (this will be convenient in what follows).

Consider the subspace Π_t spanned by the vectors $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k$ and the vector

$$v_i(t) = ((1+t)v_{i1}, \dots, (1+t)v_{i\sigma_i - 1}, 1, 0, \dots, 0).$$

Clearly, if Π is a critical point of the function fi, then $t = 0$ is a critical point of the function $\varphi(t) = fi(\Pi_t)$; therefore, it suffices to show that $\varphi'(0) \neq 0$.

Let $V(t)$ be the matrix whose rows are the coordinates of the vectors $v_1, \dots, v_{i-1}, v_i(t), v_{i+1}, \dots, v_k$, and let $V_\tau(t)$ be the matrix formed by the columns $\tau_1, \dots, \tau_k$ of $V(t)$. For the subspace Π_t, the Plücker coordinate $x_\tau(t)$ equals $\det V_\tau(t)$.

By definition, $\varphi(t) = \sum_\tau c_\tau x_\tau^2(t) / \sum_\tau x_\tau^2(t)$; hence

$$\varphi'(0) = \frac{\left(\sum_\rho c_\rho x'_\rho x_\rho\right)\left(\sum_\tau x_\tau^2\right) - \left(\sum_\rho c_\rho x_\rho^2\right)\left(\sum_\tau x'_\tau x_\tau\right)}{\left(\sum_\tau x_\tau^2\right)^2},$$

where $x_\tau = x_\tau(0)$ and $x'_\tau = x'_\tau(0)$.

The situation is the simplest in the case $i = k$. In this case, the matrix $V_\tau(t)$ has the following structure, depending on τ_k.

If $\tau_k < \sigma_k$, then the last row of $V_\tau(t)$ is obtained from the last row of $V(t)$ by multiplying all of its elements by $1 + t$. In this case, we have $x_\tau(t) = (1+t)x_\tau(0)$ and $x'_\tau(t) = x'_\tau(0)$.

If $\tau_k = \sigma_k$, then the last column of $V_\tau(t)$ consists of the elements $0, \dots, 0, 1$. In this case, $x_\tau(t) = \text{const}$ and $x'_\tau(t) = 0$.

If $\tau_k > \sigma_k$, then the last column of $V_\tau(t)$ is zero. In this case, $x_\tau(t) = 0$.

Thus, $x'_\tau \neq 0$ only if $\tau_k < \sigma_k$, and $x_\rho \neq 0$ only if $\rho_k \leq \sigma_k$. Therefore, for Schubert symbols $\tau = \bar\tau, \tau_k$ and $\rho = \bar\rho, \rho_k$ (here $\bar\tau$ and $\bar\rho$ are shorter Schubert symbols), the nonzero terms in the numerator of the expression for $\varphi'(0)$ take the form

$$\Big(\sum_{\rho,\rho_k<\sigma_k} c_{\rho,\rho_k} x^2_{\rho,\rho_k}\Big)\Big(\sum_{\bar\tau,\tau_k<\sigma_k} x^2_{\tau,\tau_k} + \sum_{\tau} x^2_{\tau,\sigma_k}\Big)$$
$$- \Big(\sum_{\rho,\rho_k<\sigma_k} c_{\rho,\rho_k} x^2_{\rho,\rho_k} + \sum_{\tau} c_{\bar\tau,\sigma_k} x^2_{\tau,\sigma_k}\Big)\Big(\sum_{\bar\tau,\tau_k<\sigma_k} x^2_{\tau,\tau_k}\Big).$$

Canceling the common part $\sum_{\substack{\rho,\rho_k<\sigma_k;\\ \tau,\tau_k<\sigma_k}} c_{\bar\rho,\rho_k} x^2_{\bar\rho,\rho_k} x^2_{\tau,\tau_k}$, we obtain

$$\sum_{\bar\rho,\rho_k<\sigma_k;\bar\tau} (c_{\rho,\rho_k} - c_{\bar\tau,\sigma_k}) x^2_{\rho,\rho_k} x^2_{\tau,\sigma_k}.$$

The constants c_σ are chosen so that $\rho_k < \sigma_k$ implies $c_{\bar\rho,\rho_k} < c_{\bar\tau,\sigma_k}$. Therefore, the equality $\varphi'(0) = 0$ implies $x^2_{\rho,\rho_k} x^2_{\tau,\sigma_k} = 0$ for all $\bar\tau$ and $\bar\rho, \rho_k < \sigma_k$. But if $\bar\tau, \sigma_k = \sigma$, then $x_{\tau,\sigma_k} = x_\sigma = 1$. Hence $x_{\bar\rho,\rho_k} = 0$ for all $\bar\rho, \rho_k < \sigma_k$. On the other hand, if the Schubert symbol ρ consists of j and $\sigma_1, \dots, \sigma_{k-1}$, then $x_\rho = \pm v_{kj}$; this contradicts the assumption that one of the numbers v_{kj} with $j < \sigma_k$ is nonzero.

Now assume that $i < k$. Recall that the number i is maximal in the sense that the vector v_i has a nonzero coordinate and the vectors $v_{i+1}, \dots, v_k$ have only zero coordinates (we ignore the coordinates $v_{j\sigma_j} = 1$). Thus, $x_\tau(t) \neq 0$ only if $\tau_i \leq \sigma_i, \tau_{i+1} = \sigma_{i+1}, \dots, \tau_k = \sigma_k$, and $x'_\tau(t) \neq 0$ only if $\tau_i < \sigma_i, \tau_{i+1} = \sigma_{i+1}, \dots, \tau_k = \sigma_k$; we have $x'_\tau(0) = x_\tau(0)$. We write the Schubert symbol

$$\tau = (\tau_1, \dots, \tau_i, \tau_{i+1}, \dots, \tau_k) = (\tau_1, \dots, \tau_i, \sigma_{i+1}, \dots, \sigma_k)$$

in the form $\bar\tau, \tau_i$, where $\bar\tau = (\tau_1, \dots, \tau_{i-1})$; we ignore the common part $\sigma_{i+1}, \dots, \sigma_k$. In this notation, the nonzero terms in the numerator of the expression for $\varphi'(0)$ have almost the same form as above; the only difference is that the inequalities $\tau_k < \sigma_k$ and $\rho_k < \sigma_k$ are replaced by $\tau_i < \sigma_i$ and $\rho_i < \sigma_i$, respectively. □

Remark 5.8. The first explicit construction of a Morse function on a Grassmann manifold was given in [**147**]. Our exposition follows [**48**]; see also [**3**]. A simpler construction of a Morse function on a Grassmann manifold was suggested in [**133**], but it uses properties of commutators of vector fields.

Fundamental Groups

In Section 2.1, we defined the fundamental group $\pi_1(X, x_0)$ of an arbitrary path-connected space X with base point x_0 and proved the basic properties of fundamental groups. In this section, we use these properties to calculate the fundamental groups of some particular topological spaces.

1. CW-Complexes

The same argument as the one used in the proof of Theorem 1.15 on p. 30 proves that *any finite connected CW-complex is homotopy equivalent to a CW-complex with only one vertex* (*0-cell*) e^0. In what follows, when calculating the fundamental group of a CW-complex X, we always assume that it has exactly one vertex, e^0. We also assume that e^0 is the base point (when the base point changes, the fundamental group is replaced by an isomorphic group; the isomorphism is induced by the map $\omega \mapsto \alpha^{-1}\omega\alpha$, where α is a path between the base points).

1.1. Main Theorem. If the cell structure of a space X is specified explicitly, then the fundamental group of X is easy to calculate. Namely, the 1-cells correspond to generators of the group, and the 2-cells correspond to relations. We explain this in more detail. The 1-skeleton X^1 is a wedge of circles; therefore, $\pi_1(X^1, e^0)$ is the free group with generators $\alpha_1, \dots, \alpha_k$, where k is the number of 1-cells in X. The characteristic map of the ith 2-cell $\chi_i^2 \colon D^2 \to X$ induces a map $\beta_i \colon \partial D^2 \mapsto X^1 \subset X$. Choose a base point s in S^1. The map β_i corresponds to an element of the fundamental group $\pi_1(X^1, \beta_i(s))$. To this element we can assign an element of the group $\pi_1(X^1, e^0)$ by choosing a path from $\beta_i(s)$ to e^0. Thus, to the ith 2-cell we

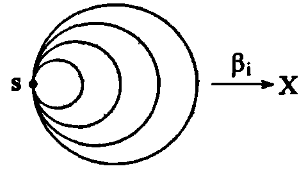

Figure 1. The homotopy

assign an element of the group $\pi_1(X^1, e^0)$. We denote it by the same symbol β_i; it is determined up to conjugation. It is easy to verify that the map $\beta_i\colon S^1 \to X$ corresponds to the identity element of the group $\pi_1(X, \beta_i(s))$; the homotopy is shown in Figure 1.

The natural embedding $X^1 \to X$ takes the generators $\alpha_1, \dots, \alpha_k \in \pi_1(X^1, e^0)$ to elements $a_1, \dots, a_k \in \pi_1(X, e^0)$. The element β_i is a word in the alphabet $\alpha_1^{\pm 1}, \dots, \alpha_k^{\pm 1}$. The corresponding word b_i in the alphabet $a_1^{\pm 1}, \dots, a_k^{\pm 1}$ is the identity element of $\pi_1(X, e^0)$; thus, we obtain the relation $b_i = 1$, where b_i is a word in the alphabet $a_1^{\pm 1}, \dots, a_k^{\pm 1}$. The relations $b_i = 1$ and $ab_ia^{-1} = 1$ are equivalent; therefore, the ambiguity in the choice of β_i is inessential in this respect.

Theorem 6.1. *The group $\pi_1(X, e^0)$ is generated by the elements $a_1, \dots, a_k$, and any relation in this group reduces to the relations $b_1, \dots, b_l$. (This means that if a word β in the alphabet $\alpha_1^{\pm 1}, \dots, \alpha_k^{\pm 1}$ corresponds to a word b in the alphabet $a_1^{\pm 1}, \dots, a_k^{\pm 1}$ that represents the identity element of $\pi_1(X, e^0)$, then $\beta \in N$, where N is the minimal normal subgroup in $\pi_1(X^1, e^0)$ containing the elements $\beta_1, \dots, \beta_l$.)*

Proof. We represent S^1 as a CW-complex with one 0-cell and one 1-cell; the 0-cell is the base point s. According to the cellular approximation theorem, any continuous map $(S^1, s) \to (X, e^0)$ is homotopic to a cellular map $(S^1, s) \to (X^1, e^0)$. Therefore, the embedding $X^1 \to X$ induces an epimorphism $\pi_1(X^1, e^0) \to \pi_1(X, e^0)$ of fundamental groups; i.e., the elements $a_1, \dots, a_k$ generate the group $\pi_1(X, e^0)$.

The same cellular approximation theorem implies that the embedding $X^2 \to X$ induces an isomorphism of fundamental groups. Indeed, since the embedding $X^1 \to X$ induces an epimorphism of fundamental groups, it suffices to verify that if two loops in X^1 (based at e^0) are homotopic in the space X, then they are homotopic in the space X^2. A homotopy in X between two loops in X^1 is a map $H\colon I^2 \to X$. Let us represent the square I^2 as a CW-complex with four vertices, four edges, and one 2-cell. The restriction of H to ∂I^2 is a cellular map; therefore, H is homotopic to

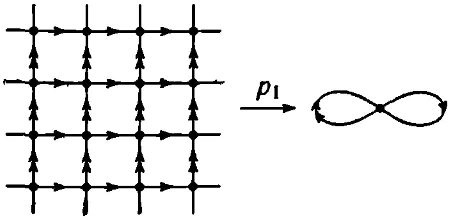

Figure 2. The covering p_1

a cellular map $H'\colon I^2 \to X^2$, and the homotopy is fixed on ∂I^2. The map H' is the required homotopy in the space X^2.

It remains to verify that the kernel of the homomorphism $\pi_1(X^1, e^0) \to \pi_1(X^2, e^0)$ induced by the embedding $X^1 \to X^2$ is the group N. Consider the covering $p_1\colon \widetilde{X}_1 \to X_1$ corresponding to the subgroup $N \subset \pi_1(X^1, e_0)$, i.e., such that $p_{1*}\pi_1(\widetilde{X}^1, \tilde{e}_0) = N$, where $\tilde{e}_0 \in p^{-1}(e_0)$. The covering p_1 in the case where X^2 is the torus and X^1 is its 1-skeleton (the wedge of two circles) is shown in Figure 2.

The group N is normal; therefore, the covering p_1 is regular. According to the definition of p_1, the lifting of each $\beta_i\colon (S^1, s) \to (X^1, e^0)$ starting at $\tilde{e}^0$ is closed. It follows from the regularity of p_1 that any lifting $\tilde{\beta}_{ij}$ of β_i starting at $\tilde{e}^0_j \in p_1^{-1}(e^0)$ is closed as well. Therefore, we can assume that $\tilde{\beta}_{ij}$ is the characteristic map $\tilde{\beta}_{ij}\colon \partial D^1_j \to \widetilde{X}^1$. Attaching the 2-cells to $\widetilde{X}^1$ via the maps $\tilde{\beta}_{ij}$, we obtain a CW-complex $\widetilde{X}^2$, and the covering p_1 can be extended to a covering $p_2\colon \widetilde{X}^2 \to X^2$.

Consider any loop $\omega\colon (S^1, s) \to (X^1, e^0)$ contractible in X^2. The contractibility of ω implies the closedness of its lifting to $\widetilde{X}^2$; moreover, the lifting of ω is entirely contained in $\widetilde{X}^1$. Therefore, the loop ω is the projection of a loop in $\widetilde{X}^1$ based at $\tilde{e}^0$. Hence ω is an element of the group N, as required. $\square$

Corollary. *The group $\pi_1(X, e^0)$ is isomorphic to $\pi_1(X^2, e^0)$, i.e., the fundamental group of a CW-complex is completely determined by the 2-skeleton of this complex.*

1.2. Some Examples.

Example. $\pi_1(\mathbb{C}P^n) = 0$.

Proof. The 2-skeleton of $\mathbb{C}P^n$ is $\mathbb{C}P^1 \approx S^2$. $\square$

Example. The group $\pi_1(nT^2)$ has generators $a_1, b_1, \dots, a_n, b_n$ and one relation $\prod_{i=1}^n (a_i b_i a_i^{-1} b_i^{-1}) = 1$.

Proof. Look at Figure 12 on p. 148. □

Example. The group $\pi_1(nP^2)$ has generators $a_1, \dots, a_n$ and one relation $a_1^2 \cdots a_n^2 = 1$.

We have already proved that the surfaces $S^2, T^2, 2T^2, \dots, P^2, 2P^2, \dots$ are not pairwise homeomorphic (see Theorem 4.5 on p. 146). The use of fundamental groups makes the proof very simple. Recall that the *commutator subgroup* of a group G is defined as the subgroup G' consisting of all products of the form $aba^{-1}b^{-1}$, where $a, b \in G$. The identity

$$xaba^{-1}b^{-1}\gamma x^{-1} = xax^{-1}a^{-1}a(xb)a^{-1}(xb)^{-1}x\gamma x^{-1}$$

shows that the subgroup $G' \subset G$ is normal. Clearly, the quotient group G/G' is commutative.

Exercise 49. (a) Prove that if $G = \pi_1(nT^2)$, then $G/G' \cong \mathbb{Z}^{2n}$.

(b) Prove that if $G = \pi_1(nP^2)$, then $G/G' \cong \mathbb{Z}^{n-1} \oplus \mathbb{Z}_2$.

It is easy to verify that the groups $\mathbb{Z}^{n_1} \oplus \mathbb{Z}_2^{\varepsilon_1}$ and $\mathbb{Z}^{n_2} \oplus \mathbb{Z}_2^{\varepsilon_2}$, where $\varepsilon_i \in \{0, 1\}$, are isomorphic if and only if $n_1 = n_2$ and $\varepsilon_1 = \varepsilon_2$.

Problem 87. (a) Prove that any subgroup of finite index in the group $\pi_1(nT^2)$ is isomorphic to $\pi_1(mT^2)$ for some m; moreover, $m - 1$ is divisible by $n - 1$.

(b) Prove that any subgroup of finite index in the group $\pi_1(nP^2)$ is isomorphic to either $\pi_1(mP^2)$, where $m - 2$ is divisible by $n - 2$, or $\pi_1(mT^2)$, where $2(m - 1)$ is divisible by $n - 2$.

Example. The fundamental group of the surface nT^2 from which $k \geq 1$ disks are removed is the free group of rank $2n + k - 1$.

Proof. The fundamental group of the space under consideration is generated by $a_1, b_1, \dots, a_n, b_n$ and $c_1, \dots, c_k$ and the relation

$$c_1 \cdots c_k \prod_{i=1}^{n} (a_i b_i a_i^{-1} b_i^{-1}) = 1$$

(see Figure 3). The element c_k can be expressed in terms of the remaining elements, which are algebraically independent. □

Corollary. *The surface nT^2 from which $k \geq 1$ disks are removed is homotopy equivalent to the wedge of $2n + k - 1$ circles. (This assertion can easily be proved directly by constructing a deformation retraction.)*

Example. The fundamental group of the surface nP^2 from which $k \geq 1$ disks are removed is the free group of rank $n + k - 1$.

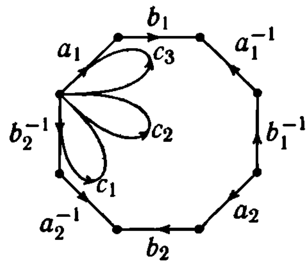

Figure 3. A sphere with handles and holes

Using the properties of fundamental groups, we can give yet another proof that the circle $S^1 = \partial D^2$ is not a retract of the disk D^2. Namely, suppose that there exists a retraction $r\colon D^2 \to S^1$. Then the composition $A \xrightarrow{i} X \xrightarrow{r} A$ is the identity map; therefore, it induces the identity homomorphism of fundamental groups. But the composition $\pi_1(A) \xrightarrow{i_*} \pi_1(X) \xrightarrow{r_*} \pi_1(A)$ cannot be the identity, because $\pi_1(A) = \mathbb{Z}$ and $\pi_1(X) = 0$.

A more delicate algebraic argument proves the following assertion. Recall that on p. 74, we proved the general theorem that the boundary of a compact manifold is never a retract of this manifold.

Problem 88. Let X be a Möbius band with boundary A. Prove that A is not a retract of X.

Problem 89. (a) Let X be the torus T^2 from which the open disk D^2 is removed. Prove that $A = \partial D^2$ is not a retract of X.

(b) Prove the same assertion for a sphere with g handles instead of T^2.

Problem 90. Let X be a closed nonorientable surface from which the open disk D^2 is removed. Prove that $A = \partial D^2$ is not a retract of X.

Now we consider a more complicated example of calculating the fundamental group. The main difficulty is in describing the cell structure of the space.

Example (see [41]). Let M_g^3 be the manifold of unit tangent vectors to a sphere with g handles. Then the group $\pi_1(M_g^3)$ has generators $a_1, \dots, a_g$, $b_1, \dots, b_g, c$ and defining relations $a_i c = c a_i$, $b_i c = c b_i$, $\prod_{i=1}^{g}(a_i b_i a_i^{-1} b_i^{-1}) = c^{2-2g}$.

Proof. Consider any vector field with isolated singular points on the manifold $M_g^2 = gT^2$. We can assume that all singular points of this vector field are inside the disk $D^2 \subset M_g^2$. The indices of the singular points of a vector field on M_g^2 sum to $2-2g$; therefore, the degree of the map $\partial D^2 \to S^1$ defined

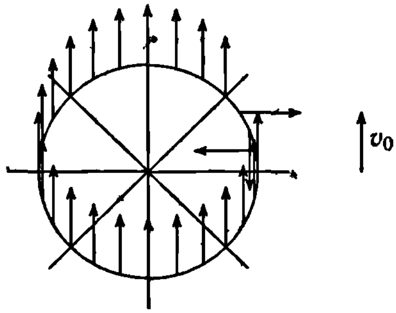

Figure 4. The vector field v

by $x \mapsto v(x)/\|v(x)\|$ is equal to $2-2g$. The map $\partial D^2 \to S^1$ is homotopic to any other map of the same degree. The homotopy can be represented as a vector field on the annulus $D^2 \setminus D^2_\varepsilon$, where D^2_ε is a disk concentric to D^2. Thus, we can construct a vector field on $M^2_g \setminus D^2_\varepsilon$ which consists of unit vectors; moreover, we can assume that on ∂D^2, this vector field is as follows: traversing a small arc, the vector $v(x)$ makes $2-2g$ rotations, and outside this arc, the vector field is constant. We extend this vector field inside D^2_ε along the radii (and leave it undefined at the center of the disk).

Using this vector field v, we define characteristic maps from 1- and 2-cells to M^3_g as follows. We represent the two-dimensional surface M^2_g as a $4g$-gon with sides glued together. We assume that the center of the disk D^2_ε corresponds to the vertices of this polygon (recall that all these vertices are glued to one point) and the arc of ∂D^2_ε outside of which the vector field v is constant is contained entirely inside one of the angles of this polygon (see Figure 4).

Let v_0 be the constant value of v outside the arc specified above. For the 0-cell of M^3_g we take the pair (center of D^2_ε, vector v_0). For the open 1-cells we take the following sets:

- the pairs (center of D^2_ε, any unit vector $w \neq v_0$);
- the pairs (point x, vector $v(x)$), where x is inside the ith side of the $4g$-gon.

Clearly, the continuous extensions of the corresponding characteristic maps $(0,1) \to M^3_g$ take the endpoints of the interval $[0,1]$ precisely to the 0-cell.

Now, we construct the 2-cells. We take a homeomorphism of the interior of the $4g$-gon onto the interior of the $(4g+1)$-gon obtained by blowing up one vertex (namely, the vertex corresponding to the arc on which the vector field is nonconstant); the blow-up is shown in Figure 5. The vector field v naturally determines a map from the closure of the $(4g+1)$-gon to M^3_g.

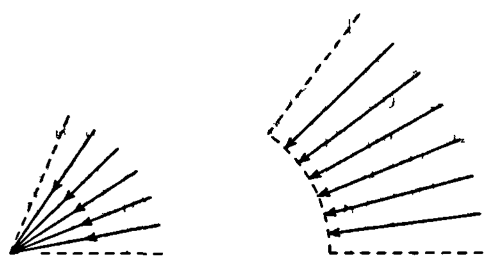

Figure 5. The blow-up

This map takes all vertices to the 0-cell, each interior point x to the pair $(x, v(x))$, and each point of the blown up side to the uniquely determined limit tangent vector at the center of D_ε^2; the remaining sides are mapped to the corresponding 1-cells.

In addition to this 2-cell, we consider the 2-cell consisting of all unit vectors at the points of one of the sides of the $4g$-gon.

It is easy to verify that the complement of the union of all closed 2-cells is an open 3-cell. Indeed, this complement consists of the unit tangent vectors at the interior points of the $4g$-gon, and at each point x, all vectors different from $v(x)$ are taken. Such a set is homeomorphic to the direct product of an open $4g$-gon and $(0,1)$.

The generators $a_1, \dots, a_g, b_1, \dots, b_g$ correspond to sides of the $4g$-gon. The generator c corresponds to the 1-cell over the center of D_ε^2. The relation $\prod_{i=1}^g (a_i b_i a_i^{-1} b_i^{-1}) = c^{2-2g}$ is determined by the $(4g+1)$-gonal cell (the map from the blown up side to the 1-cell c has degree $2-2g$). The relations $a_i c a_i^{-1} c^{-1} = 1$ and $b_i c b_i^{-1} c^{-1} = 1$ are determined by the 4-gonal cells. □

Problem 91. Prove that $M_0^3 \approx \mathbb{R}P^3$. (Here M_0^3 is the manifold of unit tangent vectors to S^2.)

1.3. The Fundamental Group of SO(n). To calculate the fundamental group of a space, it is not necessary to know the cell structure of this space; in many cases, the exact sequence of a fibration is sufficient.

Example. For $n \geq 3$, $\pi_1(\mathrm{SO}(n)) = \mathbb{Z}_2$. The generator of the group $\pi_1(\mathrm{SO}(n))$ is the path consisting of all rotations about a fixed axis.

Proof. Take a unit vector e in $\mathbb{R}^{n+1}$ and consider the map $\mathrm{SO}(n+1) \to S^n$ which assigns the vector Ae to each matrix A. This map is a locally trivial fibration with fiber SO(n). We write the exact sequence for this fibration:

$$\pi_2(S^n) \to \pi_1(\mathrm{SO}(n)) \to \pi_1(\mathrm{SO}(n+1)) \to \pi_1(S^n).$$

If $n \geq 3$, then $\pi_2(S^n) = \pi_1(S^n) = 0$; therefore, $\pi_1(\mathrm{SO}(n)) \cong \pi_1(\mathrm{SO}(n+1))$. Moreover, $\mathrm{SO}(3) \approx \mathbb{R}P^3$ (see the solution of Problem 91) and $\pi_1(\mathbb{R}P^3) = \mathbb{Z}_2$.

Let us represent $\mathbb{R}P^3$ as the disk D^3 with boundary antipodes identified. The homeomorphism $\mathrm{SO}(3) \approx \mathbb{R}P^3$ takes the rotations about a fixed axis to some diameter of this disk. Each path-diameter of D^3 corresponds to a nonzero element of the group $\pi_1(\mathbb{R}P^3)$, because the lifting of this path under the covering $S^3 \to \mathbb{R}P^3$ is nonclosed. The embedding $\mathrm{SO}(n) \to \mathrm{SO}(n+1)$ maps rotations about a fixed axis to rotations about a fixed axis, and it takes the generator of the fundamental group to the generator of the fundamental group. □

The Kakutani theorem. The relation $\pi_1(\mathrm{SO}(3)) = \mathbb{Z}_2$ has the following consequence.

Theorem 6.2 (Kakutani [64]). *Let $S^2 = \{u \in \mathbb{R}^3 : \|x\| = 1\}$, and let $f\colon S^2 \to \mathbb{R}$ be a continuous map. Then there exists an orthonormal basis u_1, u_2, u_3 in $\mathbb{R}^3$ such that $f(u_1) = f(u_2) = f(u_3)$.*

Proof. Take an orthonormal basis e_1, e_2, e_3 and consider the map $\varphi\colon \mathrm{SO}(3) \to \mathbb{R}^3$ defined by $\varphi(A) = (f(Ae_1), f(Ae_2), f(Ae_3))$. We must prove that at least one point $\varphi(A)$ belongs to the line l determined by the equation $x = y = z$. Suppose that $\varphi(\mathrm{SO}(3))$ does not intersect this line. Let ψ be the composition of φ and the projection onto the plane $x+y+z=0$ (this plane is orthogonal to l). By assumption, the origin does not belong to the image of ψ; therefore, we can consider the composition of ψ and the projection from the origin onto the unit circle S^1 in the given plane. As a result, we obtain a map $\bar{\psi}\colon \mathrm{SO}(3) \to S^1$. Consider the restriction of $\bar{\psi}$ to the subgroup $S^1 \hookrightarrow \mathrm{SO}(3)$ consisting of the rotations about l.

Let us find out what the curve $\psi(S^1)$ is. Let t be the parameter on the circle S^1 (the rotation angle about the axis l). The rotation through the angle $t = 2\pi/3$ takes the vectors e_1, e_2, and e_3 to e_2, e_3, and e_1, respectively. In the plane $x + y + z = 0$, this corresponds to the transformation $(x, y, z) \mapsto (y, z, x)$. It is easy to show that this transformation is the rotation through $2\pi/3$. Indeed, the cosine of the angle between the vectors (x, y, z) and (y, z, x), where $z = -x - y$, equals $\frac{xy-(x+y)^2}{x^2+y^2-(x+y)^2} = -\frac{1}{2}$. Therefore, the structure of the curve $\psi(S^1)$ is as follows. As t varies from 0 to $2\pi/3$, the curve goes from the point (x_0, y_0, z_0) to the point (y_0, z_0, x_0). The next segment of the curve, corresponding to t between $2\pi/3$ to $4\pi/3$, is obtained by rotating the preceding one through $2\pi/3$. The remaining part of the curve is this segment rotated through $2\pi/3$. The curve $\bar{\psi}(S^1)$ has similar structure. As t changes from 0 to $2\pi/3$, the point $\bar{\psi}(t)$ rotates through $\frac{2\pi}{3} + 2k\pi$, where k is an integer. Thus, as t changes from 0 to 2π, the point $\bar{\psi}(t)$ rotates through $2\pi(3k+1) \neq 0$. This means that the map $\bar{\psi}\colon \mathrm{SO}(3) \to S^1$ induces a nonzero homomorphism $\pi_1(\mathrm{SO}(3)) \to \pi_1(S^1)$. But any homomorphism $\mathbb{Z}_2 \to \mathbb{Z}$ is zero. □

Corollary. *Any bounded closed convex subset K of $\mathbb{R}^3$ can be inscribed in a cube.*

Proof. For each unit vector $u \in \mathbb{R}^3$, take the two supporting planes of K orthogonal to u. Let $f(u)$ be the distance between these supporting planes. Applying the Kakutani theorem to f, we obtain the required result. $\square$

The spinor group. If $n \geq 3$, then $\pi_1(\mathrm{SO}(n)) = \mathbb{Z}_2$. This means that $\mathrm{SO}(n)$ is doubly covered by a simply connected manifold. We denote this manifold by $\mathrm{Spin}(n)$. This is a group, as well as $\mathrm{SO}(n)$. The group $\mathrm{Spin}(n)$, which is called the *spinor group*, can be constructed as follows.

The *Clifford algebra* C_n is the associative algebra with identity generated by elements $e_1, \dots, e_n$ satisfying the relations $e_i^2 = -1$ and $e_ie_j + e_je_i = 0$ for $i \neq j$. Let $\mathbb{R}^n$ be the linear subspace of C_n spanned by the vectors $e_1, \dots, e_n$. It is easy to verify that $\left(\sum x_ie_i\right)^2 = -\sum x_i^2$; therefore, all nonzero elements of $\mathbb{R}^n \subset C_n$ are invertible. In particular, all elements of the unit sphere $S^{n-1} \subset \mathbb{R}^n \subset C_n$ are invertible. In the multiplicative group of invertible elements of C_n, the elements of the unit sphere S^{n-1} generate a certain subgroup; this subgroup is denoted by $\mathrm{pin}(n)$. The group $\mathrm{Spin}(n)$ is a part of the group $\mathrm{pin}(n)$. To describe this part, we use the decomposition of the linear space C_n into the direct sum $C_n^0 \oplus C_n^1$, where C_n^i is the subspace generated by all products of the form $e_{j_1} \cdots e_{j_{2k+i}}$; the spaces C_n^0 and C_n^1 intersect only in zero, because the application of the relations $e_i^2 = -1$ and $e_ie_j + e_je_i = 0$ cannot change the parity of the number of generators in the product. The group $\mathrm{Spin}(n)$ is the part of $\mathrm{pin}(n)$ that is contained in C_n^0. In other words, the group $\mathrm{Spin}(n)$ consists of the products $u_1u_2 \cdots u_{2k}$, where $u_i \in S^{n-1}$.

To construct a homomorphism $\mathrm{Spin}(n) \to \mathrm{SO}(n)$ that is a double covering, we need the involutive anti-isomorphism of the algebra C_n defined by $e_{i_1} \cdots e_{i_k} \mapsto e_{i_k} \cdots e_{i_1}$; this map leaves the elements e_i^2 and $e_ie_j + e_je_i$ fixed, and therefore this is indeed a self-map of C_n. We denote it by $u \mapsto u^*$. Clearly, $(u^*)^* = u$ and $(uv)^* = v^*u^*$; moreover, $u \in \mathbb{R}^n$ implies $u^* = u$.

To each element $u \in \mathrm{pin}(n)$ we assign the map $\varphi(u)\colon \mathbb{R}^n \to \mathbb{R}^n$ defined by $\varphi(u)x = uxu^*$. First, we must verify that the element uxu^* indeed belongs to $\mathbb{R}^n$, i.e., can be represented as a linear combination of $e_1, \dots, e_n$. It is sufficient to consider the case $u \in S^{n-1}$. Suppose that $u = \sum u_ie_i$ and $x = \sum x_ie_i$. Then $ux = -\sum u_ix_i + \sum_{i\neq j} u_ix_je_ie_j$. The product $uxu^* = uxu$ has a part that belongs to $\mathbb{R}^n$ for obvious reasons. The remaining part is of the form $\sum u_ix_ju_ke_ie_je_k$, where the summation is over the triples of pairwise distinct numbers. But $e_ie_je_k = -e_ke_je_i$; hence this sum vanishes.

Next, we verify that $\varphi(u)$ is an orthogonal transformation. First, note that if $u = u_1 \cdots u_k$, where $u_i \in S^{n-1}$, then $u^*u = uu^* = (-1)^k$. Moreover, as mentioned before, $\left(\sum x_i e_i\right)^2 = -\sum x_i^2 = -\|x\|^2$. Therefore, $\|\varphi(u)x\|^2 = -uxu^*uxu^* = -(-1)^k ux^2u^* = (-1)^k\|x\|^2uu^* = \|x\|^2$.

Theorem 6.3. *The map* φ: $\mathrm{pin}(n) \to \mathrm{O}(n)$ *is a group epimorphism. Moreover,* $\varphi^{-1}(\mathrm{SO}(n)) = \mathrm{Spin}(n)$ *and* $\operatorname{Ker}\varphi \restriction \mathrm{Spin}(n) = \{1, -1\}$.

Proof. Clearly, $\varphi(uv)x = uvx(uv)^* = uvxv^*u^* = \varphi(u)\varphi(v)x$. Therefore, φ is a group homomorphism.

We show that if $u \in S^{n-1}$, then $\varphi(u)$ is the symmetry with respect to the hyperplane orthogonal to u. We represent x as $\lambda u + w$, where $w \perp u$. We must prove that $\varphi(u)x = -\lambda u + w$. But $\varphi(u)x = u(\lambda u + w)u^* = u\lambda uu^* + uwu^* = -\lambda u + uwu^*$. It remains to show that $uwu^* = w$ ($u^* = u$ in the case under consideration). If $\sum u_i w_i = 0$ and $\sum u_i^2 = 1$, then $uw = \sum_{i\neq j} u_i w_j e_i e_j = -\sum_{i\neq j} w_j u_i e_j e_i = -wu$ and $uwu^* = -wuu^* = w$. Any orthogonal transformation can be represented as a composition of symmetries about hyperplanes; therefore, φ is an epimorphism. Clearly, an orthogonal transformation is orientation-preserving if and only if it is a composition of an even number of symmetries about hyperplanes. Therefore, $\varphi^{-1}(\mathrm{SO}(n)) = \mathrm{Spin}(n)$.

Suppose that $u \in \mathrm{Spin}(n)$ and $\varphi(u)x = x$ for all x. Then $ue_iu^* = e_i$, and hence $ue_iu^*u = e_iu$. But $u^*u = 1$ for $u \in \mathrm{Spin}(n)$. Therefore, $ue_i = e_iu$. Conversely, if $ue_i = e_iu$ for all i and $u \in \mathrm{Spin}(n)$, then $\varphi(u)x = x$ for all x. It is easy to verify that $u = \sum a_{i_1\dots i_{2k}} e_{i_1}\cdots e_{i_{2k}}$ implies $e_iue_i^{-1} = -e_iue_i = \sum(-1)^{\varepsilon(i)} a_{i_1\dots i_{2k}} e_{i_1}\cdots e_{i_{2k}}$, where $\varepsilon(i) = 0$ for $i \notin \{i_1, \dots, i_{2k}\}$ and $\varepsilon(i) = 1$ for $i \in \{i_1, \dots, i_{2k}\}$. Thus, if $u \in C_n^0$ and $e_iu = ue_i$ for all i, then $u = \lambda \cdot 1$, where $\lambda \in \mathbb{R}$. Therefore, $\operatorname{Ker}\varphi \restriction \mathrm{Spin}(n) = \{1, -1\}$. □

Thus, the group $\mathrm{Spin}(n)$ is a double covering of $\mathrm{SO}(n)$. It remains to show that this covering is nontrivial, i.e., the space $\mathrm{Spin}(n)$ is connected. It suffices to verify that $\mathrm{Spin}(n)$ contains a path joining 1 and -1. Such a path is given by

$$(e_1\cos t + e_2\sin t)(e_1\cos t - e_2\sin t) = -\cos 2t - e_1e_2\sin 2t,$$

where $t \in [0, \pi/2]$.

2. The Seifert–van Kampen Theorem

2.1. Equivalent Formulations. Suppose that a path-connected topological space X is the union of path-connected sets U_1 and U_2 with path-connected intersection $U_1 \cap U_2$. Choose a point $x_0 \in U_1 \cap U_2$ and consider the fundamental groups $\pi_1(U_1 \cap U_2, x_0)$, $\pi_1(U_1, x_0)$, and $\pi_1(U_2, x_0)$. Suppose that these groups are defined by sets of generators S, S_1, and S_2 and

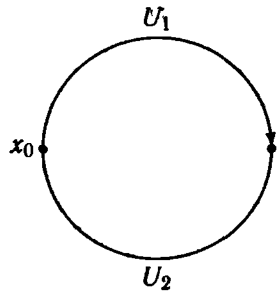

Figure 6. The counterexample

relations R, R_1, and R_2. For brevity, we use the same notation for maps of topological spaces and for the maps of fundamental groups that they induce. The embedding $\psi_i\colon U_i \to X$ takes each element $s_i \in S_i$ to $\psi_i s_i \in \pi_1(X, x_0)$. Thus, $\pi_1(X, x_0)$ contains sets (alphabets) $\psi_1 S_1$ and $\psi_2 S_2$ such that every element of the image of $\psi_i\colon \pi_1(U_i, x_0) \to \pi_1(X, x_0)$ can be written as a word in the alphabet $\psi_i S_i$. Now, take an element $s \in S$ and consider its image in $\pi_1(U_i, x_0)$ under the embedding $\varphi_i\colon U_1 \cap U_2 \to U_i$. The element $\varphi_i s$ can be represented as a word in the alphabet S_i, and the element $\psi_i \varphi_i s$ can be represented as a word in the alphabet $\psi_i S_i$. In the group $\pi_1(X, x_0)$, the relation $\psi_1\varphi_1 s = \psi_2\varphi_2 s$ holds, because $\psi_1\varphi_1 = \psi_2\varphi_2$ (both of these compositions coincide with the embedding of $U_1 \cap U_2$ into X).

Thus, the group $\pi_1(X, x_0)$ contains the elements $\psi_1 S_1$ and $\psi_2 S_2$, and these elements are related by $\psi_1 R_1$, $\psi_2 R_2$, and $\psi_1\varphi_1 s = \psi_2\varphi_2 s$, where $s \in S$. Consider the group G with generators $\psi_1 S_1$ and $\psi_2 S_2$ and defining relations $\psi_1 R_1$, $\psi_2 R_2$, and $\psi_1\varphi_1 s = \psi_2\varphi_2 s$, where $s \in S$. It is natural to expect that G is isomorphic to $\pi_1(X, x_0)$. This is not so without additional assumptions; there exists an example for $X = S^1$ (see Figure 6). But under certain assumptions about the path-connected spaces under consideration, this assertion becomes true. For example, in 1931, Seifert **[120]** proved that it is true in the case where U_1 and U_2 are subcomplexes of a simplicial complex, and in 1933, van Kampen **[67]** proved that it is true if U_1 and U_2 are open in X.

The theorem of Seifert (for CW-complexes) is an obvious consequence of Theorem 6.1 on p. 258 with only one additional assumption: the generators of the fundamental groups are the 1-cells, and the relations are the 2-cells. Thus, before proceeding further, we give an invariant definition of the group G, which does not depend on the choice of generators and relations.

Let G_0, G_1, and G_2 be groups defined by sets of generators S, S_1, and

S_2 and sets of relations R, R_1, and R_2. Suppose that G is a group,

$$\begin{tikzcd} & G_1 \arrow[dr, "\psi_1"] & \\ G_0 \arrow[ur, "\varphi_1"] \arrow[dr, "\varphi_2"'] & & G \\ & G_2 \arrow[ur, "\psi_2"'] & \end{tikzcd}$$

is a commutative diagram of homomorphisms, and G' is another group for which the similar diagram

$$\begin{tikzcd} & G_1 \arrow[dr, "\psi_1'"] & \\ G_0 \arrow[ur, "\varphi_1"] \arrow[dr, "\varphi_2"'] & & G' \\ & G_2 \arrow[ur, "\psi_2'"'] & \end{tikzcd}$$

holds. We show that in this case, there exists a unique homomorphism $\sigma\colon G \to G'$ for which $\sigma\psi_i = \psi_i'$. Let F be the free group on generators $\psi_1 S_1$ and $\psi_2 S_2$, and let N be its minimal normal subgroup containing the words $\psi_1 R_1$, $\psi_2 R_2$, and $\psi_1\varphi_1 s(\psi_2\varphi_2 s)^{-1}$, where $s \in S$. For each element $s_i \in S_i$, we set $\bar{\sigma}\psi_i s_i = \psi_i' s_i$. Any map from the generators of the free group F to an arbitrary group can be uniquely extended to a homomorphism; thus, we obtain a homomorphism $\bar{\sigma}\colon F \to G'$. If $\bar{\sigma}(N) = 1$, then the homomorphism $\bar{\sigma}$ induces a homomorphism $\sigma\colon G \to G'$ for which $\sigma\psi_i = \psi_i'$. Therefore, it suffices to verify that $\bar{\sigma}(N) = 1$ (the homomorphism σ is unique because the elements $\psi_1 S_1$ and $\psi_2 S_2$ generate the group G). If $r_i \in R_i$, then $\bar{\sigma}\psi_i r_i = \psi_i' r_i = 1$, because the word r_i is the identity element of G_i. If $s \in S$, then $\bar{\sigma}\psi_i\varphi_i s = \psi_i'\varphi_i s$; hence the equality $\psi_1'\varphi_1 = \rho' = \psi_2'\varphi_2$ implies $\bar{\sigma}\psi_1\varphi_1 s = \bar{\sigma}\psi_2\varphi_2 s$.

Consider all commutative diagrams of homomorphisms

$$\begin{tikzcd} & G_1 \arrow[dr, "\psi_1"] & \\ G_0 \arrow[ur, "\varphi_1"] \arrow[dr, "\varphi_2"'] & & G \\ & G_2 \arrow[ur, "\psi_2"'] & \end{tikzcd}$$

with fixed homomorphisms φ_1 and φ_2. The group G in these diagrams is called the *amalgam* of the groups G_1 and G_2 with respect to the group G_0 (and the homomorphisms φ_1 and φ_2) if it has the universal property mentioned above: for any other group G' from a similar commutative diagram, there exists a unique homomorphism $\sigma\colon G \to G'$ such that $\sigma\psi_i = \psi_i'$. To obtain an invariant definition of the group G, it remains to show that the amalgam is unique up to isomorphism.

Let G and G' be two amalgams with respect to the same groups. Then there exist homomorphisms $\sigma\colon G \to G'$ and $\sigma'\colon G' \to G$ for which $\sigma\psi_i = \psi_i'$ and $\sigma'\psi_i' = \psi_i$. Consider the homomorphism $\sigma'\sigma\colon G \to G$. It has the property that $\sigma'\sigma\psi_i = \sigma'\psi_i' = \psi_i$. By assumption, there is only one homomorphism $G \to G$ with this property. On the other hand, the identity homomorphism does have this property; therefore, $\sigma'\sigma = \mathrm{id}_G$. Similarly, $\sigma\sigma' = \mathrm{id}_{G'}$, and hence σ is a group isomorphism.

2.2. The Proof.

Theorem 6.4 (van Kampen). *Let U_1 and U_2 be open path-connected subsets of the space $X = U_1 \cup U_2$ such that their intersection $U_1 \cap U_2$ is path-connected. Consider the commutative diagram*

$$\begin{tikzcd}
 & \pi_1(U_1) \arrow[dr, "\psi_1"] & \\
\pi_1(U_1 \cap U_2) \arrow[ur, "\varphi_1"] \arrow[dr, "\varphi_2"'] & & \pi_1(X) \\
 & \pi_1(U_2) \arrow[ur, "\psi_2"'] &
\end{tikzcd}$$

induced by the embeddings (the fundamental groups are with respect to the same base point $x_0 \in U_1 \cap U_2$). Then the group $\pi_1(X)$ is the amalgam of $\pi_1(U_1)$ and $\pi_1(U_2)$ with respect to the group $\pi_1(U_1 \cap U_2)$ and the homomorphisms φ_1 and φ_2.

It is more convenient to prove the theorem of van Kampen in terms of generators and relations rather than in terms of amalgams; we have explained above that the two statements are equivalent.

Proof (see [31]). Consider the commutative diagram

$$\begin{tikzcd}
 & \pi_1(U_1) \arrow[dr, "\psi_1'"] & \\
\pi_1(U_1 \cap U_2) \arrow[ur, "\varphi_1"] \arrow[dr, "\varphi_2"'] & & H \\
 & \pi_1(U_2) \arrow[ur, "\psi_2'"'] &
\end{tikzcd}$$

We must prove that there exists a unique homomorphism $\sigma\colon \pi_1(X) \mapsto H$ for which $\sigma\psi_i = \psi_i'$. The uniqueness of σ is easy to derive from the following assertion.

Step 1 (generators). The images of the groups $\pi_1(U_1)$ and $\pi_2(U_2)$ under the homomorphisms ψ_1 and ψ_2 generate $\pi_1(X)$.

Consider an arbitrary loop $\alpha\colon [0,1] \to X$ based at x_0. The interval $[0,1]$ is covered by two open sets, $\alpha^{-1}(U_1)$ and $\alpha^{-1}(U_2)$. Let δ be the Lebesgue number of this cover. Choose points $0 = t_1 < t_2 < \dots < t_{n+1} = 1$ such that $t_{k+1} - t_k < \delta$. The image of each interval $[t_k, t_{k+1}]$ is entirely contained in one of the sets U_1 and U_2. The loop α is the composition of (not necessarily closed) paths $\alpha_1, \alpha_2, \dots, \alpha_n$, where each α_k is the restriction of α to the interval $[t_k, t_{k+1}]$. To represent α as such a composition of loops, we join each point $\alpha(t_k) \in U_i$ to x_0 by a path β_k contained in U_i; if $\alpha(t_k) \in U_1 \cap U_2$, then the path β_k must be contained in $U_1 \cap U_2$. If the path α_k is contained entirely in U_i, then the loop $\beta_k^{-1}\alpha_k\beta_{k+1}$ is contained entirely in U_i as well; therefore, the class of the loop $\beta_k^{-1}\alpha_k\beta_{k+1}$ is contained in the image of the group $\pi_1(U_i)$ under the homomorphism ψ_i. Moreover, α is the composition of the loops $\alpha_1\beta_2$, $\beta_2^{-1}\alpha_3\beta_3, \dots, \beta_{n-1}^{-1}\alpha_{n-1}\beta_n$, $\beta_n^{-1}\alpha_n$.

Now it is easy to prove that the homomorphism σ is unique. Indeed, let us represent an element $\alpha \in \pi_1(X)$ in the form $\alpha = \prod_{k=1}^n \psi_{i(k)}\gamma_k$, where $\gamma_k \in \pi_1(U_{i(k)})$. The equality $\sigma\psi_i = \psi_i'$ implies $\sigma\psi_{i(k)}\gamma_k = \psi_{i(k)}'\gamma_k$. Therefore,

$$\sigma\alpha = \prod_{k=1}^{n} \psi_{i(k)}'\gamma_k. \tag{1}$$

Relation (1) completely determines the homomorphism σ. It remains to verify that this homomorphism is well defined, i.e., $\alpha = \prod_{k=1}^n \psi_{i(k)}'\gamma_k$ depends only on α and not on the representation $\alpha = \prod_{k=1}^n \psi_{i(k)}\gamma_k$. For this purpose, it suffices to prove the following assertion.

Step 2 (relations). Suppose that $\gamma_k \in \pi_1(U_{i(k)})$ and $\prod_{k=1}^n \psi_{i(k)}\gamma_k = 1$; then $\prod_{k=1}^n \psi_{i(k)}'\gamma_k = 1$.

The equality $\prod_{k=1}^n \psi_{i(k)}\gamma_k = 1$ means that there exists a map $f\colon I^2 \to X$ with the following properties:

- the restriction of f to the bottom edge of the square I^2 represents the class of the loop $\prod_{k=1}^n \psi_{i(k)}\gamma_k$;
- the map f takes the other edges of I^2 to the base point x_0 (see Figure 7).

Let δ be the Lebesgue number of the cover of I^2 by the open sets $f^{-1}(U_1)$ and $f^{-1}(U_2)$. We partition I^2 into rectangles by vertical and horizontal intervals in such a way that the diagonal of each rectangle is less than δ. In the set of such intervals we include all vertical intervals dividing the bottom edge of the square into n equal intervals (it is assumed that each of these equal intervals corresponds to one of the paths γ_i; in particular, their endpoints are mapped to the base point x_0).

By construction, the image of each rectangle is contained in U_1 or U_2. Let a be a vertex of one of the rectangles. We connect the point $f(a)$ to the base

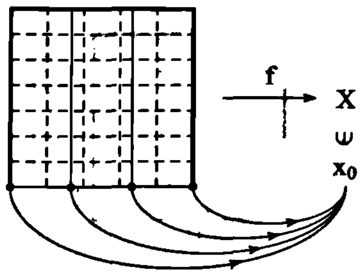

Figure 7. The map of the square

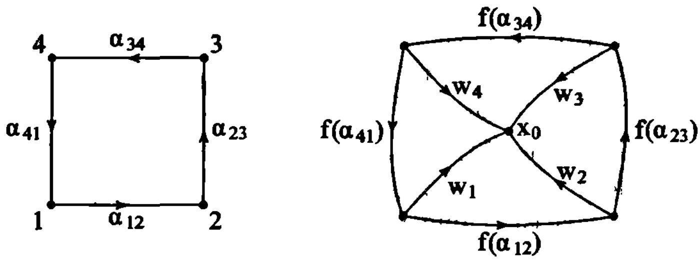

Figure 8. A rectangle from the partition and its image

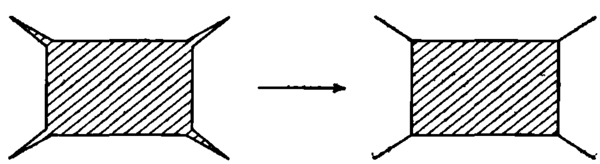

Figure 9. The map of the disk

point x_0 by a path ω_a so that if $a \in U_i$, then $\omega_a \subset U_i$ (in particular, if $a \in U_1 \cap U_2$, then $\omega_a \subset U_1 \cap U_2$) and if $a = x_0$, then $\omega_a = x_0$ (the path is constant).

Consider one of the rectangles in the partition. Let α_{12}, α_{23}, α_{34}, α_{41} be its sides (see Figure 8). It is easy to show that the loop $\beta_{12}\beta_{23}\beta_{34}\beta_{41}$, where $\beta_{pq} = \omega_p^{-1} f(\alpha_{pq})\omega_q$, is contractible in the space X. Indeed, it suffices to prove that the corresponding map $S^1 \to X$ can be extended to a map $D^2 \to X$. The required map $D^2 \to X$ can be constructed as follows. First, we send D^2 to a rectangle with segments attached to its vertices (see Figure 9). To the rectangle we apply the map f, and the segments are mapped to X by using ω_p.

For the element of $\pi_1(X)$ represented by the loop β_{pq} we use the same notation β_{pq}. By construction, the loop β_{pq} is contained in U_1 or U_2; hence $\beta_{pq} = \psi_i \beta_i$, where $\beta_i \in \pi_1(U_i)$. We set $\beta'_{pq} = \psi'_i \beta_i \in H$. We must verify

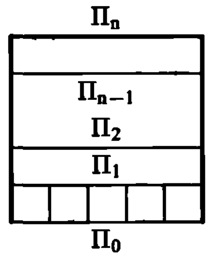

Figure 10. The products Π_i

that the element β'_{pq} is well defined (the loop β_{pq} may lie in $U_1 \cap U_2$). If $\beta_{pq} \subset U_1 \cap U_2$, then $\beta_{pq} = \psi_1\beta_1 = \psi_2\beta_2$, where $\beta_1 = \varphi_1\beta$ and $\beta_2 = \varphi_2\beta$ for some $\beta \in \pi_1(U_1 \cap U_2)$. We have to prove that $\psi'_1\beta_1 = \psi'_2\beta_2$. By assumption, $\psi'_1\varphi_1 = \psi'_2\varphi_2$; therefore, $\psi'_1\beta_1 = \psi'_1\varphi_1\beta = \psi'_2\varphi_2\beta = \psi'_2\beta_2$.

By construction, one of the sets U_1 and U_2 contains not only the loop β_{pq} but also the four loops β_{12}, β_{23}, β_{34}, and β_{41}; hence $\beta_{pq} = \psi_i\delta_{pq}$, where $\delta_{pq} \in \pi_1(U_i)$ and i is the same for all the four loops. The proof of the contractibility of the loop $\delta_{12}\delta_{23}\delta_{34}\delta_{41}$ is similar to that of contractibility of $\beta_{12}\beta_{23}\beta_{34}\beta_{41}$ (with X replaced by U_i). Thus,

$$\beta'_{12}\beta'_{23}\beta'_{34}\beta'_{41} = (\psi'_i\delta_{12})(\psi'_i\delta_{23})(\psi'_i\delta_{34})(\psi'_i\delta_{41}) = \psi'_i(\delta_{12}\delta_{23}\delta_{34}\delta_{41}) = \psi'_i(1) = 1.$$

Let us summarize. To each of the (directed) edges of the rectangles into which the square I^2 is partitioned we have assigned an element of the group H, and the elements assigned to the edges of the same rectangle satisfy the relation $\beta'_{12}\beta'_{23}\beta'_{34}\beta'_{41} = 1$. We must prove that the product of the elements of H that correspond to the bottom edge of the square is equal to 1. We denote this product by Π_0 (see Figure 10). Consider the product of all relations of the form $\beta'_{12}\beta'_{23}\beta'_{34}\beta'_{41} = 1$ over the rectangles adjacent to the bottom edge of the square. We multiply the relations in such a way that the common sides of the rectangles are included with opposite orientations. As a result, we obtain $\Pi_0 = \Pi_1$, where Π_1 is the product of the elements of H that correspond to the next-to-bottom horizontal segment. (The elements corresponding to the two extreme vertical sides are not canceled, but these sides are mapped to the base point x_0 and, therefore, correspond to the identity element of the group.) Similarly, we obtain $\Pi_1 = \Pi_2, \dots, \Pi_{n-1} = \Pi_n$. But the upper horizontal segment is mapped to the base point x_0; hence $\Pi_n = 1$. □

Corollary. *If $n \geq 3$, then $\pi_1(M^n) \cong \pi_1(M^n \setminus \{x\})$, where x is an arbitrary point of the manifold M^n.*

Proof. Take a neighborhood of x homeomorphic to $\mathbb{R}^n$ for U_1 and an arbitrary point of the set $U_1 \setminus \{x\}$ for the base points in the spaces M^n and

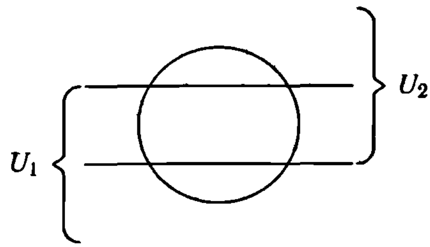

Figure 11. The sets U_1 and U_2

$M^n \setminus \{x\}$. The set $\{x\}$ is closed; hence $U_2 = M^n \setminus \{x\}$ is open. Clearly, $U_1 \cap U_2 = \mathbb{R}^n \setminus \{x\} \sim S^{n-1}$. By assumption, $n \geq 3$; therefore, $\pi_1(U_1 \cap U_2) = 1$. Moreover, $\pi_1(U_1) = 1$. Thus, the groups $\pi_1(M^n)$ and $\pi_1(M^n \setminus \{x\})$ have the same sets of generators and relations. □

Remark 6.1. This assertion can be proved without employing the van Kampen theorem. Indeed, consider a triangulation K of the manifold M^n. We can assume that x is inside some simplex Δ^n of this triangulation. Then the space $M^n \setminus \{x\}$ is homotopy equivalent to the simplicial complex obtained from K by deleting Δ^n. If $n \geq 3$, then the deletion of an n-simplex does not affect the 2-skeleton.

2.3. Knot Groups. A *knot* is the image of the circle S^1 under a continuous map $f\colon S^1 \to \mathbb{R}^3$, and the *group* of a knot K is defined as the group $\pi_1(\mathbb{R}^3 \setminus K, x_0)$, where x_0 is an arbitrary point of $\mathbb{R}^3 \setminus K$. A knot K is said to be *polygonal* if f is a piecewise linear function of the parameter φ on the circle $S^1 = \{e^{i\varphi}\}$. A knot K is *smooth* if f is smooth and $\operatorname{grad} f = \left(\frac{\partial x}{\partial \varphi}, \frac{\partial y}{\partial \varphi}, \frac{\partial z}{\partial \varphi}\right)$ vanishes nowhere.

In this section, we apply the van Kampen theorem to calculate the group of a polygonal (or smooth) knot K. To practice the application of the van Kampen theorem, we start with the simplest example; namely, we calculate the group of the *trivial* knot $S^1 \subset \mathbb{R}^3$ ($S^1 \subset \mathbb{R}^2 \subset \mathbb{R}^3$ is the usual circle).

Example. $\pi_1(\mathbb{R}^3 \setminus S^1) = \mathbb{Z}$.

Proof. For U_1 and U_2 we take the open subsets of $\mathbb{R}^3 \setminus S^1$ shown schematically in Figure 11. Clearly, $U_1 \cap U_2 \sim S^1 \vee S^1$ and $U_i \sim S^1$. We choose generators a and b in the group $\pi_1(U_1 \cap U_2)$ and generators α_1 and α_2 in the groups $\pi_1(U_1)$ and $\pi_1(U_2)$ as shown in Figure 12; we impose no relations on α_1 and α_2. Each homomorphism $\varphi_i\colon \pi_1(U_1 \cap U_2) \to \pi_1(U_i)$ takes both elements a and b to the same element α_i. Therefore, the group $\pi_1(\mathbb{R}^3 \setminus S^1)$ has generators $\psi_1\alpha_1$ and $\psi_2\alpha_2$ and defining relation $\psi_1\alpha_1 = \psi_2\alpha_2$. As a result, we obtain a group on one generator α with no relations. □

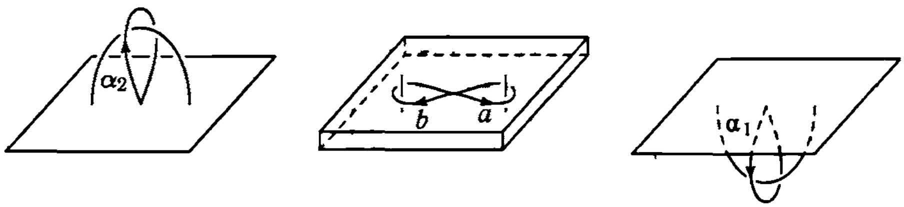

Figure 12. The choice of generators

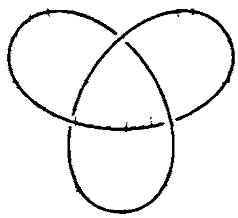

Figure 13. A knot diagram

Problem 92. Prove that $\mathbb{R}^3 \setminus S^1 \sim S^2 \vee S^1$.

For any smooth knot, we can choose a plane such that the projection of the knot on this plane satisfies the following conditions:

- the projection of no tangent to the knot degenerates into a point;
- no point is the projection of more than two different points of the knot;
- the set of *crosses* (i.e., the points in the plane to which two different points of the knot are projected) is finite, and the projections of the tangents to the knot at the two points corresponding to a cross do not coincide.

For a polygonal knot, a projection plane with similar properties can be chosen. The analog of the tangents to a smooth knot is the straight lines containing the edges of the polygonal knot.

A *diagram* of a knot is its projection on a plane satisfying the conditions specified above. For each cross, the diagram must indicate which branch of the knot passes over the other (see Figure 13).

Instead of the knot K, consider any knot K' which coincides with some diagram of K everywhere except in small neighborhoods of the crosses; at each cross, one of the branches passes above the plane of the diagram and the other remains in this plane (see Figure 14). Clearly, the groups of the knots K and K' are isomorphic because the spaces $\mathbb{R}^3 \setminus K$ and $\mathbb{R}^3 \setminus K'$ are homotopy equivalent. In calculating knot groups, we assume that the knots

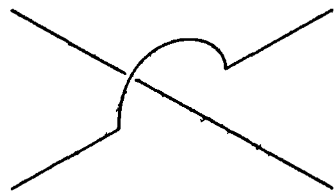

Figure 14. A neighborhood of a cross

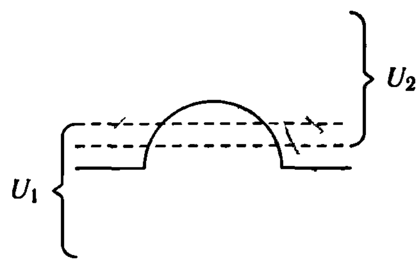

Figure 15. The sets U_1 and U_2

are embedded in space as K'. We consider only one simple example, namely, calculate the group of the *trefoil* knot, which is shown in Figure 13. The group of any other knot can be calculated by the same method.

Example. The group of the trefoil knot is defined by generators x and y and the relation $xyx = yxy$.

Proof. For U_1 and U_2 we take the open subsets of $\mathbb{R}^3 \setminus K$ shown schematically in Figure 15. The set U_i is obtained from the half-space by removing n arcs, where n is the number of crosses in the knot diagram. It is easy to verify that the projections of these arcs on the plane of the knot diagram are pairwise disjoint; therefore, U_i is homotopy equivalent to the wedge of n circles. Clearly, $U_1 \cap U_2$ is homotopy equivalent to the wedge of $2n$ circles.

Take the generators of the fundamental groups $\pi_1(U_1)$, $\pi_1(U_1 \cap U_2)$, and $\pi_1(U_2)$ shown in Figure 16 (for the group $\pi_1(U_1)$, in addition to its generators, the loop $\varphi_1 x_2$ and a loop homotopic to it are shown). Clearly, $\varphi_1 x_1 = x$. Looking at Figure 16, we see that $\varphi_1 x_2 = yzy^{-1}$. In U_2, we have the simpler relation $\varphi_1 x_1 = \varphi_1 x_2 = \bar{x}$. Therefore, $\psi_1 x = \psi_1 \bar{x}$ and $\psi_1(yzy^{-1}) = \psi_1 \bar{x}$. As a result, we obtain the generators x, y, and z and the defining relations $x = yzy^{-1}$, $y = zxz^{-1}$, and $z = xyx^{-1}$ of the trefoil group. The last relation gives an expression for z in terms of x and y; substituting it into the first and second relations, we obtain $yxy = xyx$. ☐

Similar arguments prove the following theorem.

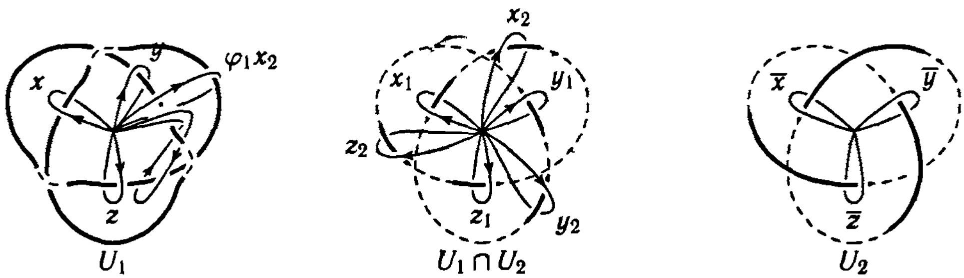

Figure 16. Generators of the fundamental groups

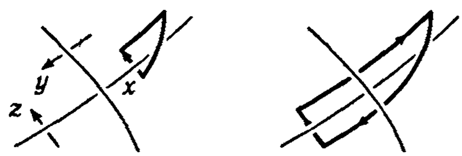

Figure 17. The relation for a cross

Theorem 6.5. *Suppose that each arc in a diagram of a knot K is associated with one generator (the generators are shown by arrows under the diagram; the change of the direction of an arrow corresponds to the replacement of a by a^{-1}), and each cross is associated with a relation of the form $x = yzy^{-1}$ (see Figure 17). Then the group defined by such generators and relations is isomorphic to the group of the knot K.*

Corollary. *The quotient of any knot group modulo its commutator subgroup is isomorphic to $\mathbb{Z}$.*

Proof. If $yz = zy$, then $x = yzy^{-1} = z$. It remains to note that joining the arcs x and z for each cross, we obtain a connected curve. □

In some cases, knot groups can be applied to prove that the spaces $\mathbb{R}^3 \setminus K_1$ and $\mathbb{R}^3 \setminus K_2$, where K_1 and K_2 are knots, are not homeomorphic.

Example. If K is the trefoil knot, then the space $\mathbb{R}^3 \setminus K$ is not homeomorphic to $\mathbb{R}^3 \setminus S^1$.

Proof. Let G be the trefoil group. We already know that it is defined by generators x and y and the relation $xyx = yxy$. It is easy to verify that in the symmetric group S_3, the elements (12) and (23) satisfy the relation

$$(12)(23)(12) = (13) = (23)(12)(23).$$

Hence there exists a homomorphism $h\colon G \to S_3$ for which $h(x) = (12)$ and $h(y) = (23)$. The image of this homomorphism contains the transpositions

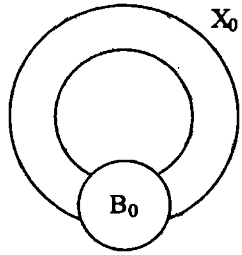

Figure 18. The construction of the space X_0

(12), (23), and (13); therefore, h is an epimorphism. But the group S_3 is non-Abelian, and hence the group G is non-Abelian also. In particular, $G \neq \mathbb{Z}$. □

Problem 93. Suppose that p and q are coprime integers and K is the *toric knot* of type (p, q), i.e., the closed curve on a standardly embedded torus which uniformly winds around the torus p times in the direction of meridians and q times in the direction of parallels. Prove that the group of the knot K is defined by generators x and y and the relation $x^p = y^q$.

2.4. The Alexander Horned Sphere. The *Alexander horned sphere* is the boundary of a set $X \subset \mathbb{R}^3$ such that X is homeomorphic to the disk D^3 but $\mathbb{R}^3 \setminus X$ is not homeomorphic to $\mathbb{R}^3 \setminus D^3$. The set X is constructed as follows. Let B_0 be the standard disk D^3. Attaching the handle $D^2 \times I$ to B_0 as shown in Figure 18, we obtain X_0. Cutting the attached handle along the 2-disk $D^2 \times \{1/2\}$ (this operation can be thought of as removing $D^2 \times I'$, where I' is an interval inside I, from the handle $D^2 \times I$), we obtain a 3-disk B_1. To the cuts (the 2-disks corresponding to the flat surfaces of the removed cylinder $D^2 \times I'$) we attach two linked handles (see Figure 19) and denote the result by X_1. Cutting both handles, we obtain a disk B_2. As above, we attach two pairs of linked handles to the cuts and denote the result by X_2, and so on. We assume that $X_0 \supset X_1 \supset X_2 \supset \cdots$, i.e., every time, a cylinder $D^2 \times I$ is cut out and the handles attached are contained in the cut-out domain.

These constructions can be performed in such a way that there exist homeomorphisms $h_n\colon B_n \to B_{n+1}$ for which $\|x - h_n(x)\| \leq 1/2^n$. Then the limit $\lim_{n\to\infty} h_n \circ h_{n-1} \circ \cdots \circ h_0 = h$ is defined, and the map h is continuous. Clearly, the map $h\colon B_0 \to h(B_0)$ is one-to-one, and hence h is a homeomorphism (see Theorem 3.2 on p. 89). Thus, the space $X = \bigcap_{n=0}^{\infty} X_n = h(B_0)$ is homeomorphic to the disk D^3.

Let us calculate the fundamental group $\pi_1(\mathbb{R}^3 \setminus X)$. We set $Y_n = \mathbb{R}^3 \setminus X_n$. Then $Y_0 \subset Y_1 \subset \cdots$ and $\mathbb{R}^3 \setminus X = \bigcup_{n=0}^{\infty} Y_n$. Clearly, $\pi_1(Y_0) = \mathbb{Z}$ because X_0

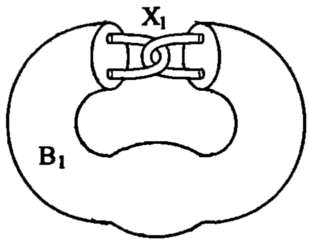

Figure 19. The construction of the space X_1

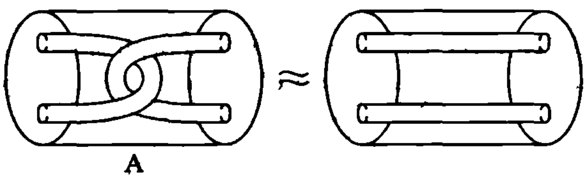

Figure 20. A disk from which two tubes are removed

is the solid torus standardly embedded in $\mathbb{R}^3$. Let us show that $\pi_1(Y_1) = F_2$ is a free group on two generators α_1 and α_2 and the embedding $Y_0 \to Y_1$ induces the monomorphism $\pi_1(Y_0) \to \pi_1(Y_1)$ that takes the generator α of the group $\pi_1(Y_0)$ to the commutator $[\alpha_1, \alpha_2]$.

We represent Y_1 as the union $Y_0 \cup A$, where A is the space shown in Figure 20. To apply the van Kampen theorem, instead of Y_0 we must take $Y_0' = Y_0 \cup Z$, where Z is a small neighborhood of the cylinder $S^1 \times I$, the lateral surface of the set A. But in the situation under consideration, this is inessential from the homotopy point of view. It is easy to see that the space A (a disk from which two linked handles are removed) is homeomorphic to a disk from which two rectilinear handles are removed. Therefore, the fundamental group $\pi_1(A)$ is a free group on two generators α_1 and α_2, which are represented by loops thrown upon the handles removed.

The generator of the group $\pi_1(Y_0)$ is represented by the loop α thrown upon the removed handles as shown in Figure 21. This loop is homotopic to the composition $\alpha_1\alpha_2\alpha_1^{-1}\alpha_2^{-1}$. Therefore, the group $\pi_1(Y_1)$ is generated by α, α_1, and α_2, which are related by $\alpha = [\alpha_1, \alpha_2]$.

Similarly, we can obtain Y_{n+1} from Y_n by attaching 2^n copies of the set A. As a result, arguing by induction, we conclude that $\pi_1(Y_{n+1})$ is a free group on 2^{n+1} generators, and the homomorphism $\pi_1(Y_n) \to \pi_1(Y_{n+1})$ induced by the embedding $Y_n \to Y_{n+1}$ is a monomorphism.

Any loop in $\mathbb{R}^3 \setminus B$ is compact; therefore, it is contained in some set Y_n. The image of a homotopy between two loops is compact also. Therefore, the

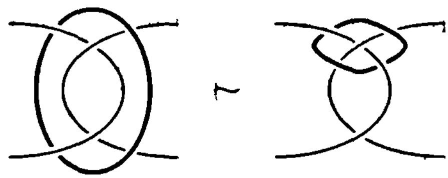

Figure 21. The homotopy of the loop α

group $\pi_1(\mathbb{R}^3 \setminus B)$ is isomorphic to $G = \bigcup_{n=0}^{\infty} \pi_1(Y_n)$. The fundamental group of the space $\mathbb{R}^3 \setminus D^3$ is trivial, while the group G contains, e.g., a free group on two generators. Thus, the space $\mathbb{R}^3 \setminus X$ is not homeomorphic to $\mathbb{R}^3 \setminus D^3$.

Note that the commutator quotient of G is trivial, because each generator is the commutator of two other generators.

3. Fundamental Groups of Complements of Algebraic Curves

A *plane algebraic curve* C is the subset of $\mathbb{C}P^2$ given by an equation of the form $P(x, y, z) = 0$, where P is a homogeneous polynomial of degree $n \geq 1$; the number n is called the *degree* of the curve C. A curve C is said to be *reducible* if $P = P_1 P_2$, where P_1 and P_2 are homogeneous polynomials of positive degrees. A point of a curve C is *singular* if $\operatorname{grad} P = 0$ at this point.

The first groups of the form $\pi_1(\mathbb{C}P^2 \setminus C)$ were calculated in [**66, 148 150**]; since then, they have been studied by many authors.

We start with several examples, in which C is a union of complex lines in $\mathbb{C}P^2$. This corresponds to the situation where the polynomial P decomposes into linear factors.

3.1. The Complement of a Set of Complex Lines. In this section, we calculate the group $\pi_1\bigl(\mathbb{C}P^2 \setminus \bigcup_{i=1}^{n} l_i\bigr)$, where $l_1, \ldots, l_n$ are complex lines in $\mathbb{C}P^2$. This group depends not only on the number of the lines but also on their arrangement. We consider several examples of arrangements. Note that removing one line turns $\mathbb{C}P^2$ into $\mathbb{C}^2$; therefore,

$$\mathbb{C}P^2 \setminus \bigcup_{i=1}^{n} l_i \approx \mathbb{C}^2 \setminus \bigcup_{i=1}^{n-1} \hat{l}_i,$$

where $\mathbb{C}^2 = \mathbb{C}P^2 \setminus l_n$ and $\hat{l}_i = l_i \cap \mathbb{C}^2$. Thus, calculating the fundamental group of the complement of n lines in $\mathbb{C}P^2$ reduces to calculating the fundamental group of the complement of $n-1$ (complex) lines in $\mathbb{C}^2$.

Example. If lines $l_1, \ldots, l_n \in \mathbb{C}P^2$ pass through one point, then

$$\pi_1(\mathbb{C}P^2 \setminus \bigcup_{i=1}^{n} l_i) = F_{n-1}$$

(the free group of rank $n-1$).

Proof. The lines $\hat{l}_1, \ldots, \hat{l}_{n-1}$ do not intersect in the space $\mathbb{C}^2 = \mathbb{C}P^2 \setminus l_n$; hence we can choose complex coordinates z and w in such a way that $\hat{l}_i = \{(z, w) : w = c_i\}$, where c_i is a constant. Therefore,

$$\mathbb{C}^2 \setminus \bigcup_{i=1}^{n-1} \hat{l}_i \sim \mathbb{C} \setminus \{c_1, \ldots, c_{n-1}\} \sim \bigvee_{i=1}^{n-1} S_i^1. \qquad \square$$

Example. If lines $l_1, l_2, l_3 \in \mathbb{C}P^2$ intersect in three distinct points, then $\pi_1(\mathbb{C}P^2 \setminus \bigcup_{i=1}^{3} l_i) = \mathbb{Z} \oplus \mathbb{Z}$.

Proof. In the space $\mathbb{C}^2 = \mathbb{C}P^2 \setminus l_3$, the lines $\hat{l}_1$ and $\hat{l}_2$ intersect at some point; therefore, we can choose complex coordinates z and w in such a way that these lines are determined by the equations $z = 0$ and $w = 0$. The origin does not belong to $\mathbb{C}^2 \setminus (\hat{l}_1 \cup \hat{l}_2)$; thus, applying the projection onto $S^3 = \{(z, w) \in \mathbb{C}^2 : |z|^2 + |w|^2 = 1\}$ along the real rays from the origin in $\mathbb{C}^2 = \mathbb{R}^4$, we can show that the space $\mathbb{C}^2 \setminus (\hat{l}_1 \cup \hat{l}_2)$ is homotopy equivalent to $S^3 \cap (\mathbb{C}^2 \setminus (\hat{l}_1 \cup \hat{l}_2))$. The latter space is the sphere S^3 from which the two circles given by the equations $z = 0$ and $w = 0$ (on S^3) are removed.

Puncturing a 3-manifold does not change its fundamental group (see p. 272). Therefore, we can remove one point from S^3 and pass to the space $\mathbb{R}^3 \setminus (K_1 \cup K_2)$, where K_1 and K_2 are circles. (If $K_1, \ldots, K_n$ are pairwise disjoint images of circles under homeomorphisms $S^1 \to \mathbb{R}^3$, then the set $K_1 \cup \cdots \cup K_n \subset \mathbb{R}^3$ is called an *n-component link*.) The fundamental group of the space $\mathbb{R}^3 \setminus (K_1 \cup K_2)$ is calculated in exactly the same way as the fundamental group of a knot (see the example on p. 275). To construct the diagram of the link $K_1 \cup K_2$, consider the stereographic projection of the sphere S^3 onto the subspace $\operatorname{Re} w = 0$ from the point $(0, 1)$. The projection of K_1 is a circle centered at the origin in the plane $\operatorname{Im} w = 0$, and the projection of K_2 is the line $z = 0$. Therefore, the diagram of the link $K_1 \cup K_2$ is as shown in Figure 22. The group $\pi_1(\mathbb{R}^3 \setminus (K_1 \cup K_2))$ is defined by generators a and b and the relation $b = aba^{-1}$, i.e., $ab = ba$ (the second cross gives the same relation). $\square$

Now, let us explain geometrically how the relation $ab = ba$ arises when we go around the intersection point of two complex lines. The element a is determined by the unit circle in the plane $w = w_0$, and the element b is determined by the unit circle in the plane $z = z_0$. We traverse the circle b and simultaneously translate the circle a (see Figure 23). Formally, such

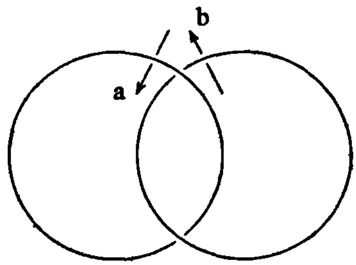

Figure 22. The diagram of the link $K_1 \cup K_2$

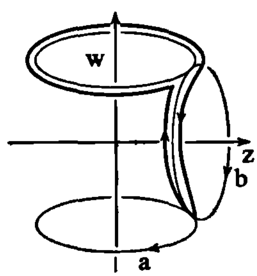

Figure 23. Going around the intersection point of two complex lines

a translation can be described as follows. We assume that $z_0 = w_0 = 1$. Suppose that the point $x(t) = (1, e^{it})$ moves uniformly on the circle b. To each moment of time t we associate the loop that first goes from $x_0 = (1,1)$ to $x(t)$, then passes along the circle (e^{is}, e^{it}), where $t = \text{const}$, and finally returns from $x(t)$ to x_0. After the whole circle b is traversed (i.e., at $t = 2\pi$), we obtain the loop bab^{-1}; this loop is homotopic to the initial loop a.

Problem 94. Given a link $L = K_1 \cup \cdots \cup K_n$ whose diagram consists of n disjoint circles (such a link is called the *trivial n-component link*), prove that the space $\mathbb{R}^3 \setminus L$ is homotopy equivalent to the wedge of n copies of the space $S^2 \vee S^1$.

Problem 95. Given a link $L = K_1 \cup K_2$ with diagram shown in Figure 22, prove that $\mathbb{R}^3 \setminus L \sim T^2 \vee S^2$.

Example. Let $\hat{l}_1, \ldots, \hat{l}_n$ be pairwise distinct lines in $\mathbb{C}^2$ passing through the origin. Then the group $\pi_1\big(\mathbb{C}P^2 \setminus \bigcup_{i=1}^n \hat{l}_i\big)$ is defined by generators $\alpha_1, \ldots, \alpha_n$ and the relations

$$\alpha_1\alpha_2\alpha_3\cdots\alpha_n = \alpha_2\alpha_3\cdots\alpha_n\alpha_1 = \alpha_3\cdots\alpha_n\alpha_1\alpha_2 = \cdots = \alpha_n\alpha_1\alpha_2\cdots\alpha_{n-1}.$$

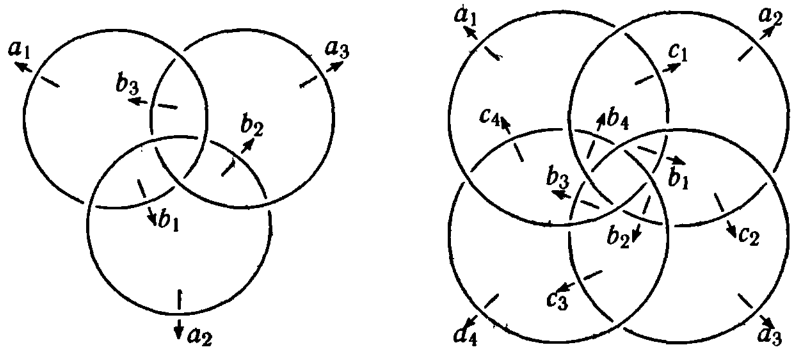

Figure 24. The diagrams of two links

Proof. The case $n = 2$ was considered in the example on p. 3.1. We consider only the cases $n = 3$ and $n = 4$. As in the above-mentioned example, calculating the required fundamental group reduces to calculating the groups of the links shown in Figure 24 (the circles are pairwise linked).

For $n = 3$, the inner crosses give the relations $b_i = a_{i+2}a_ia_{i+2}^{-1}$ and the outer crosses give the relations $a_i = a_{i+1}b_ia_{i+1}^{-1}$. Expressing the generators b_1, b_2, and b_3 in terms of a_1, a_2, and a_3 with the use of the former relations and substituting the obtained expressions into the latter relations, we see that the required group has generators a_1, a_2, and a_3 and defining relations $a_1a_2a_3 = a_2a_3a_1 = a_3a_1a_2$.

For $n = 4$, we obtain the relations $b_i = a_{i+1}a_ia_{i+1}^{-1}$, $c_i = a_{i+2}b_ia_{i+2}^{-1}$, and $a_i = a_{i-1}c_ia_{i-1}^{-1}$. First, we express the generators $\{b_i\}$ in terms of $\{a_i\}$, then we apply the relations obtained to express $\{c_i\}$ in terms of $\{a_i\}$, and finally we substitute the resulting expressions into the relations $a_i = a_{i-1}c_ia_{i-1}^{-1}$. As a result, we obtain the relations $a_1a_4a_3a_2 = a_4a_3a_2a_1 = a_3a_2a_1a_4 = a_2a_1a_4a_3$. It remains to set $\alpha_1 = a_1$, $\alpha_2 = a_4$, $\alpha_3 = a_3$, and $\alpha_4 = a_2$. □

3.2. The van Kampen Theorem. In the paper [148] published in 1929, Zariski proposed a method for calculating the group $\pi_1(\mathbb{C}P^2 \setminus C)$. Soon after, he revealed that his proof used a conjecture which he was not able to prove (if an infinite group is defined by finitely many generators and relations, then the intersection of all of its subgroups of finite index is trivial). In 1951, this conjecture was disproved.

Another method for calculating the group $\pi_1(\mathbb{C}P^2 \setminus C)$ was suggested by van Kampen [66]. The theorem on the structure of the group $\pi_1(\mathbb{C}P^2 \setminus C)$ that he proved is closely related to his other theorem discussed on p. 269 (Theorem 6.4). Both theorems were published in the same issue of the *American Journal of Mathematics*.

Our exposition of the van Kampen theorem on the group $\pi_1(\mathbb{C}P^2 \setminus C)$ follows largely [**27**].

Suppose that a curve C is determined by the equation $P = 0$, where $P = P_1P_2$. Let C_1 and C_2 be the curves given by $P_1 = 0$ and $P_2 = 0$. The equality $\operatorname{grad} P = P_1 \operatorname{grad} P_2 + P_2 \operatorname{grad} P_1$ implies that the singular points of the curve C are the intersection points of the curves C_1 and C_2 and the singular points of C_1 and C_2.

We assume that C is a curve of degree n given by the equation $P = 0$, where $P = P_1 \cdots P_k$ and the polynomials P_i are irreducible and pairwise distinct. Let C_i be the curve $P_i = 0$. The singular points of C are the singular points of the curves C_i and the intersection points of pairs of these curves.

Suppose that $x_0 \in \mathbb{C}P^2 \setminus C$ is a base point. We show that almost all lines passing through x_0 intersect the curve C in exactly n distinct points ("almost all" here means "all but finitely many"). The number of intersection points of pairs of curves $C_1, \dots, C_k$ is finite; therefore, it suffices to consider only straight lines not passing through these points. Such a line intersects the curve C in n points if and only if it intersects each curve C_i in n_i points, where n_i is the degree of the curve C_i. Therefore, it is sufficient to prove the required assertion for an irreducible curve C_i. A straight line intersects the curve C_i in fewer than n_i points if it is tangent to C_i or passes through its singular point. The number of tangent lines to the irreducible curve C_i that can be drawn through a given point is finite (it does not exceed the degree of the dual curve C_i^*), and any irreducible curve has only finitely many singular points.

Choose a line l passing through x_0 and intersecting C in exactly n distinct points. Take a point a on l which is different from x_0 and does not belong to C. Let $l_0, l_1, \dots, l_s$ be the straight lines that pass through a and intersect C in fewer than n distinct points. We choose the point a so that each of the lines $l_0, l_1, \dots, l_s$ contains precisely one singular point or point of tangency. Finally, we draw a line m different from l through the point x_0 and consider the points $a_0, a_1, \dots, a_s$ at which the lines $l_0, \dots, l_s$ intersect m (see Figure 25). We choose m in such a way that it does not pass through the intersection points of C with the lines l_i, i.e., so that the points a_i do not belong to C.

We set $E = \mathbb{C}P^2 \setminus (C \cup l_0 \cup \cdots \cup l_s)$ and $B = m \setminus \{a_0, \dots, a_s\}$. The projection from the point a onto the line m induces a map $p\colon E \to B$.

Lemma. *The map $p\colon E \to B$ is a locally trivial fibration with fiber F homeomorphic to $\mathbb{C}$ with n punctures.*

Proof. We show that for any point $x \in B$, there exists an open set $U \ni x$ such that the map p is a trivial fibration over U. We take any point $x' \neq x$ on

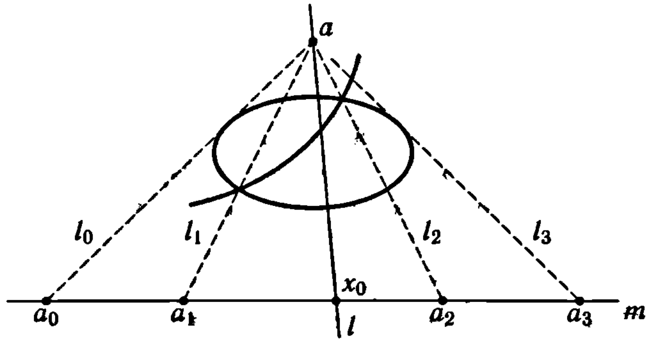

Figure 25. The projection of the curve C onto the line m

m and remove the line ax' from $\mathbb{C}P^2$. On the remaining set, which is homeomorphic to $\mathbb{C}^2$, we take the line m and one of the lines passing through a in $\mathbb{C}P^2$ for coordinate axes. In these coordinates, the projection from a onto m has the form $(z, w) \mapsto (z, 0)$. If the neighborhood $U \subset B$ is sufficiently small, then the branches of the curve C over it are sufficiently well approximated by the lines $w = \alpha_i z + \beta_i$, where $i = 1, \dots, n$. The lines $z = \text{const}$ are not included, because we have eliminated the tangents to C. The lines $w_i = \alpha_i z + \beta_i$ have no intersection points over a sufficiently small domain U, because the curve C has no singular points over it. Hence the branches of the curve C over a small domain $U \subset B$ are sufficiently well approximated by the lines $w = c_i$, where $i = 1, \dots, n$; "sufficiently well" means, in particular, that the set $p^{-1}(U)$ is homeomorphic to $U \times (\mathbb{C} \setminus \{c_1, \dots, c_n\})$, and the homeomorphism is compatible with p. □

The space $B = m \setminus \{a_0, \dots, a_s\}$ is homotopy equivalent to the wedge of s circles. The group $\pi_1(B, x_0)$ is generated by loops $h_1, \dots, h_s$, each enclosing exactly one of the points a_i, where $i = 1, \dots, s$; there are no relations between these generators, but if we add a loop h_0 enclosing a_0, then the relation $h_0 h_1 \cdots h_s = 1$ will arise.

For generators of the group $\pi_1(F, x_0)$ we take loops $g_1, \dots, g_n$ each of which encloses exactly one of the punctures in $\mathbb{C}$. Moreover, we need the composition of these loops in $\mathbb{C}P^1$ to be homotopic to a loop enclosing a (the fiber F is the complex projective line $\mathbb{C}P^1$ from which the point a and the n intersection points of this line with the curve C are removed). In the fundamental group of the space with a glued back in, the relation $g_1 \cdots g_n = 1$ arises.

To calculate the group $\pi_1(E, x_0)$, we use the exact sequence for the fibration under consideration:

$$\pi_2(B, x_0) \xrightarrow{\partial_*} \pi_1(F, x_0) \xrightarrow{i_*} \pi_1(E, x_0) \xrightarrow{p_*} \pi_1(B, x_0) \xrightarrow{\partial_*} \pi_0(F, x_0).$$

Since the fiber F is connected, we have $\pi_0(F, x_0) = 1$. Moreover, the space B is homotopy equivalent to a wedge of circles, and the universal covering space of a wedge of circles is contractible. Therefore, $\pi_2(B, x_0) = 1$. As a result, we obtain

$$1 \to \pi_1(F, x_0) \xrightarrow{i_*} \pi_1(E, x_0) \xrightarrow{p_*} \pi_1(B, x_0) \to 1.$$

The map of the group $\pi_1(F, x_0)$, which is freely generated by $g_1, \dots, g_n$, to $\pi_1(E, x_0)$ is a monomorphism. Hence we can identify $\pi_1(F, x_0)$ with the subgroup $G = i_*\pi_1(F, x_0) \subset \pi_1(E, x_0)$. The group $\pi_1(B, x_0)$ is freely generated by $h_1, \dots, h_s$. In the case under consideration, the space B is contained in E; therefore, $\pi_1(E, x_0)$ has elements represented by the same loops as the elements $h_1, \dots, h_s$ of the group $\pi_1(B, x_0)$; we use the same notation for these elements of $\pi_1(E, x_0)$.

The fact that the group $\pi_1(B, x_0)$ is free substantially facilitates the calculation of the group $\pi_1(E, x_0)$. Indeed, for a free group, any map of its free generators to a group can be (uniquely) extended to a group homomorphism. Hence there exists a unique homomorphism $\varphi\colon \pi_1(B, x_0) \to \pi_1(E, x_0)$ for which $\varphi(h_i) = h_i$. The subgroup $H = \varphi\pi_1(B, x_0) \subset \pi_1(E, x_0)$ is isomorphic to $\pi_1(B, x_0)$, because $p_*\varphi = \mathrm{id}_{\pi_1(B,x_0)}$ and $\varphi p_* \restriction H = \mathrm{id}_H$.

Each element $\omega \in \pi_1(E, x_0)$ has a unique representation in the form $\omega = gh$, where $g \in G$ and $h \in H$ (namely, $h = p_*(\omega)$ and $g = \omega^{-1}h$). Moreover,

$$(gh)(g'h') = (ghg'h^{-1})hh', \text{ where } hg'h^{-1} \in G.$$

Therefore, the group $\pi_1(E, x_0)$ is completely determined by the groups G and H and the action of H on G defined by $h(g) = hgh^{-1} \in G$. Hence the group $\pi_1(E, x_0)$ is defined by the generators $g_1, \dots, g_n$, $h_1, \dots, h_s$ and the relations $h_j g_i h_j^{-1} = \psi_{ij}(g_1, \dots, g_n)$, where $\psi_{ij}(g_1, \dots, g_n)$ is the expression of $h_j g_i h_j^{-1} \in G$ in terms of the generators $g_1, \dots, g_n$.

We have calculated the group $\pi_1(E, x_0)$. The next step is the calculation of the group $\pi_1(E', x_0)$, where

$$E' = \mathbb{C}P^2 \setminus (C \cup l_0) = E \cup \left(E' \cap (l_1 \cup \dots \cup l_s)\right).$$

The set $E' \cap (l_1 \cup \dots \cup l_s)$ is a submanifold of codimension 2 in the manifold E'; therefore, the embedding $E \to E'$ induces an epimorphism $\pi_1(E, x_0) \twoheadrightarrow \pi_1(E', x_0)$. Indeed, any loop in E' is homotopic to a loop disjoint from $l_1, \dots, l_s$. Hence the group $\pi_1(E', x_0)$ has the same generators $g_1, \dots, g_n$, $h_1, \dots, h_s$; to the relations inherited from $\pi_1(E, x_0)$ some new relations may add. For example, the loops h_i are contractible in E' (see Figure 26), which gives the new relations $h_i = 1$ for $i = 1, \dots, s$. We show that no other new relations arise. Consider any homotopy between a loop γ and the constant loop x_0 in E'. We can assume that the loop γ is smooth and avoids the lines $l_1, \dots, l_s$. First, we replace the homotopy under consideration by a smooth

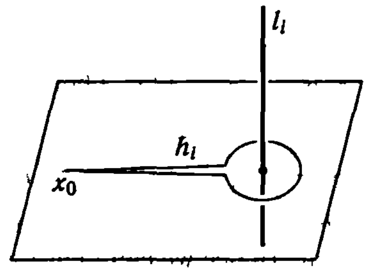

Figure 26. A contractible loop

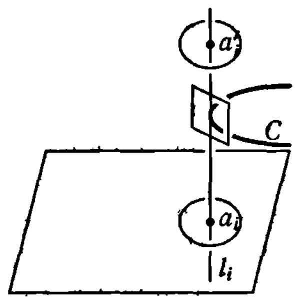

Figure 27. A standard loop

homotopy, and then we slightly change the smooth homotopy so as to obtain a homotopy Ψ such that the points $a_1, \dots, a_s$ are not critical for the map $p\Psi$, where p is the projection from the point a onto the line m. The map $\Psi\colon I^2 \to E'$ has the property that the preimage of $l_1 \cup \dots \cup l_s$ consists of finitely many interior points of the square I^2 (the boundary points of I^2 are mapped to x_0 or to some other points of the loop γ; none of their images belongs to $l_1, \dots, l_s$).

The main difficulty is that each loop in $U \setminus l_i$, where U is a sufficiently small neighborhood of $a_i' \in l_i \setminus C$, must be replaced by a standard loop lying on a circle centered at a_i on the complex line $m \setminus \{a_0\}$ (see Figure 27) by means of a homotopy in the space E; meanwhile, we consider homotopies of loops in the class of all maps $S^1 \to E$, i.e., we allow homotopies to move the base point. To construct such a homotopy, take a path $k(s)$ on the complex line $l_i \setminus \{a\}$ that joins the points a_i' and a_i and avoids the point $l_i \cap C$ (see Figure 28). The required homotopy is constructed as follows. In the space $\mathbb{C}^2 = \mathbb{C}P^2 \setminus l_0$, the projection from a onto the complex line $m \setminus \{a_0\}$ is written as $(z, w) \mapsto (z, 0)$ in some coordinates. Moreover, $a_i = (z_0, 0)$ and $a_i' = (z_0, w_0)$ in these coordinates. The loop in $U \setminus l_i$ defined by $\gamma'(t) = (z(t), w(t))$ is homotopic to the loop $\gamma''(t) = (z(t), w_0)$. Suppose that γ'' is sufficiently small and the minimum distance ρ from the path $k(s)$ to the points $l_i \cap C$

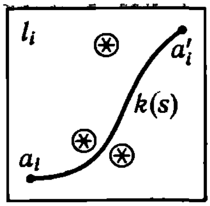

Figure 28. A path from a_i to a'_i

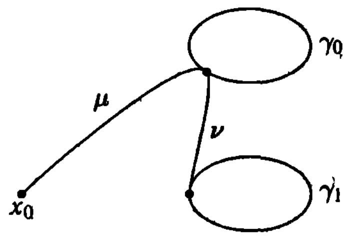

Figure 29. Homotopic loops

is sufficiently large, namely, $\rho > \max|z(t) - z_0|$. Then the formula $\gamma_s(t) = (z(t), k(s))$ defines a homotopy in E between the loop $\gamma_0 = \gamma''$ and a loop γ_1 which is contained entirely in the complex line $m \setminus \{a_0\}$. Moreover, the loop γ_1 lies in a small neighborhood of the point a_i and does not pass through this point. Such a loop is homotopic to a loop lying on a circle centered at a_i.

We can pass from homotopy in the class of all loops to homotopy in the class of loops with fixed base point x_0 as follows. Suppose that $\gamma_s(t)$ is a homotopy between loops $\gamma_0(t)$ and $\gamma_1(t)$, $\mu(\tau)$ is a path from x_0 to $\gamma_0(0)$, and $\nu(s) = \gamma_s(0)$ is a path from $\gamma_0(0)$ to $\gamma_1(0)$. Then the loops $\mu\gamma_0\mu^{-1}$ and $\mu\nu\gamma_1\nu^{-1}\mu^{-1}$ are homotopic (see Figure 29).

We return to the homotopy Ψ constructed on p. 286. In the space E, any loop γ contractible in E' is homotopic to the loop $\prod \mu_k\gamma_k\mu_k^{-1}$, where γ_k is a loop in a small neighborhood of a point of the line $l_{i(k)}$; the homotopy is constructed as shown in Figure 30. In the language of group generators and relations, this means that any word representing the identity element of the group $\pi_1(E', x_0)$ can be reduced to the form $\prod \mu_k\gamma_k\mu_k^{-1}$ by using only the relations between the elements of $\pi_1(E, x_0)$; here we assume that both groups are defined by the same generators (specified before). We have shown that the loop $\mu_k\gamma_k\mu_k^{-1}$ is homotopic in E to the loop $\mu'_k h^r_{i(k)}(\mu'_k)^{-1}$, where $h_{i(k)}$ is a loop from the set of generators $\{h_1, \dots, h_s\}$ and r is an integer. This means that the word $\mu_k\gamma_k\mu_k^{-1}$ can be reduced to the form $\mu'_k h^r_{i(k)}(\mu'_k)^{-1}$ by using only the relations between the generators of $\pi_1(E, x_0)$. Finally, the relation $h_{i(k)} = 1$ implies $\mu'_k h^r_{i(k)}(\mu'_k)^{-1} = 1$. This means that if a word γ

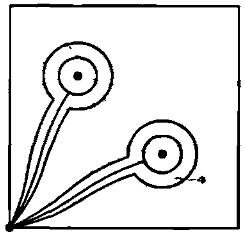

Figure 30. The construction of a homotopy

represents the identity element of $\pi_1(E', x_0)$, then the equality $\gamma = 1$ follows from the relations between the elements of $\pi_1(E, x_0)$ and the relations $h_1 = 1, \ldots, h_s = 1$.

The last step is the calculation of the group $\pi_1(\mathbb{C}P^2 \setminus C)$, where $\mathbb{C}P^2 \setminus C = E' \cup (l_0 \setminus C)$. As mentioned earlier (see p. 284), after adding the point $a \in l_0 \setminus C$, the relation $g_1 \cdots g_n = 1$ arises. The same argument as that used at the preceding step shows that there are no other relations.

We now state the final result. Recall that the words $\psi_{ij}(g_1, \ldots, g_n)$ were defined on p. 285.

Theorem 6.6 (van Kampen [66]). *Suppose that C is a curve of degree n in $\mathbb{C}P^2$, $a \in \mathbb{C}P^2 \setminus C$, and $l_0, l_1, \ldots, l_s$ are straight lines each of which passes through a and is either tangent to the curve C at one point or passes through one singular point of C. Then the group $\pi_1(\mathbb{C}P^2 \setminus C, x_0)$ is defined by generators $g_1, \ldots, g_n$ and the $ns + 1$ relations $g_1 \cdots g_n = 1$, $g_i = \psi_{ij}(g_1, \ldots, g_n)$, where $i = 1, \ldots, n$ and $j = 1, \ldots, s$.*

3.3. Applications of the van Kampen Theorem. The van Kampen theorem gives an algorithm for calculating the groups $\pi_1(\mathbb{C}P^2 \setminus C, x_0)$. Its most difficult part is constructing the expressions $\psi_{ij}(g_1, \ldots, g_n)$. For this reason, we start with a more detailed discussion of their geometric meaning.

Recall the definition of the elements $\psi_{ij}(g_1, \ldots, g_n) \in \pi_1(F, x_0)$, where F is the fiber over the point x_0. The fiber F is the plane $\mathbb{C}$ with n punctures. Let $g_1, \ldots, g_n$ be loops, each going around one of these points (all in the same direction). The base B is $\mathbb{C}$ with s punctures; the loops $h_1, \ldots, h_s$ are defined similarly to $g_1, \ldots, g_n$. The loop $h_j g_i h_j^{-1}$ is homotopic to a loop lying in the fiber F. The element $h_j g_i h_j^{-1}$ can be represented as a word in the alphabet $g_1, \ldots, g_n$; this word is $\psi_{ij}(g_1, \ldots, g_n)$. The homotopy of $h_j g_i h_j^{-1}$ is described as follows (see Figure 31). We take the loop g_i and translate it along h_j (so that at each point of h_j, a loop in the fiber over this point is obtained). After going around the point a_j, we return to the fiber over x_0 and obtain a loop in this fiber. This new loop is $\psi_{ij}(g_1, \ldots, g_n)$.

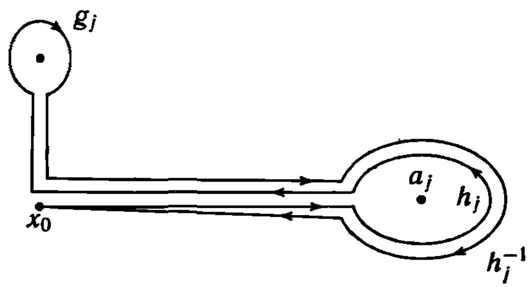

Figure 31. The homotopy of the loop $h_j g_i h_j^{-1}$

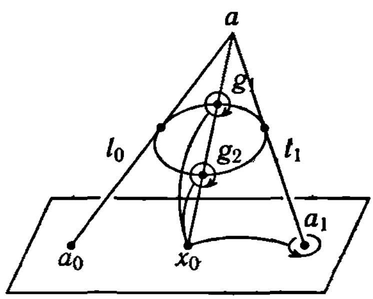

Figure 32. Tangents to the conic

Example. Let C_2 be a nondegenerate conic in $\mathbb{C}P^2$ (e.g., given by the equation $z_1^2 + z_2^2 + z_3^2 = 0$). Then $\pi_1(\mathbb{C}P^2 \setminus C_2) = \mathbb{Z}_2$.

Proof. Through each point $a \notin C_2$ pass exactly two tangents to the conic C_2 (see Figure 32). The group $\pi_1(\mathbb{C}P^2 \setminus C_2)$ is defined by the generators g_1 and g_2 and the relations $g_1 g_2 = 1$ and $g_i = \psi_{i1}(g_1, g_2)$, where $i = 1, 2$.

Instead of going around a tangent to the conic, we can consider a simpler situation, namely, going around the complex line $z = 0$ tangent to the curve $z = w^2$ in $\mathbb{C}^2$. If $z = e^{i\varphi}$, then $w = \pm e^{i\varphi/2}$; therefore, going around the origin in the real plane $w = 0$ interchanges the branches of the function $w(z)$ while preserving the directions of the loops in the real plane $z = \text{const}$ (see Figure 33). This means that $\psi_{11}(g_1, g_2) = g_2$ and $\psi_{21}(g_1, g_2) = g_1$, i.e., $g_1 = g_2$. $\square$

Example. Let C_n be the curve in $\mathbb{C}P^2$ determined by the equation $z_1^n + z_2^n + z_3^n = 0$. Then $\pi_1(\mathbb{C}P^2 \setminus C_n) = \mathbb{Z}_n$.

Proof. In $\mathbb{C}P^2$, the projection onto the real plane $z_3 = 0$ from the point $a = (0 : 0 : 1)$ is given by $(z_1 : z_2 : z_3) \mapsto (z_1 : z_2 : 0)$. A point $(z_1 : z_2 : 0)$ has n preimages on the curve C_n if and only if $z_1^n + z_2^n \neq 0$. The points for which

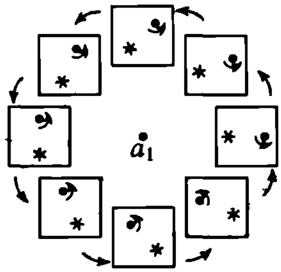

Figure 33. Going around the origin

$z_1 : z_2 = -\varepsilon_k$, where $\varepsilon_k^n = 1$, correspond to tangent lines; moreover, these tangents must have multiplicity n (at each point of tangency, n branches must merge). Going around a tangent of multiplicity n is essentially the same as going around the origin in the plane $w = 0$ for the algebraic function $w(z)$ defined by $z = w^n$; it results in the rotation of the real plane $z = 0$ through an angle of $2\pi/n$ and a cyclic permutation of the branches. Thus, we obtain the relations $g_1 = g_2$, $g_2 = g_3, \ldots, g_{n-1} = g_n$, $g_n = g_1$. We also have the relation $g_1 g_2 \cdots g_n = 1$. As a result, we obtain a group with defining generator g and relation $g^n = 1$. □

Problem 96* ([94]). Let p and q ($p, q \geq 2$) be coprime integers, and let $C_{p,q}$ be the curve in $\mathbb{C}P^2$ given by the equation

$$(z_1^p + z_2^p)^q + (z_2^q + z_3^q)^p = 0.$$

Prove that the group $\pi_1(\mathbb{C}P^2 \setminus C_{p,q})$ is defined by two generators a and b and the relations $a^p = 1$ and $b^q = 1$.

Hints and Solutions

1. Let $\{V_i\}$ be a countable base of X, and let I be the set of those indices i for which V_i is contained in some U_α. To every $i \in I$ we assign one set $U_{\alpha(i)}$ containing V_i. Let us show that $\{U_{\alpha(i)}\}$ is a cover of X. Each point $x \in X$ is contained in some set U_α. The set U_α can be represented as the union of V_i; therefore, $x \in V_{i_0} \subset U_\alpha$ for some i_0. Clearly, $i_0 \in I$ and $x \in U_{\alpha(i_0)}$.

2. The set $S^{n+m-1} \setminus S^{n-1}$ consists of the points $(x_1, \dots, x_{n+m}) \in \mathbb{R}^{n+m}$ for which $x_1^2 + \dots + x_n^2 < 1$ and $x_1^2 + \dots + x_{n+m}^2 = 1$. To each point $(x_1, \dots, x_{n+m}) \in S^{n+m-1} \setminus S^{n-1}$ we assign $(y_1, \dots, y_{n+m})$, where $y_1 = x_1, \dots, y_n = x_n$, $y_{n+1} = x_{n+1}/a, \dots, y_{n+m} = x_{n+m}/a$, and $a = \sqrt{1 - x_1^2 - \dots - x_n^2}$. Clearly,

$$y_{n+1}^2 + \dots + y_{n+m}^2 = \frac{x_{n+1}^2 + \dots + x_{n+m}^2}{1 - x_1^2 - \dots - x_n^2} = 1;$$

thus, we have defined a homeomorphism $S^{n+m-1} \setminus S^{n-1} \to D^n \times S^{m-1}$, where D^n is the open unit disk; it is homeomorphic to $\mathbb{R}^n$.

3. Suppose that the set $A \cup B$ can be represented as the union of nonempty disjoint sets X_1 and X_2 each of which is both open and closed. Take a point $x \in A \cap B$ belonging to X_1. The sets $A_1 = A \cap X_1$ and $A_2 = A \cap X_2$ are open and closed in A, and $A_1 \neq \varnothing$, because $x \in A_1$. Therefore, $A_1 = A$ and $A_2 = \varnothing$. Similarly, $B \cap X_1 = B$ and $B \cap X_2 = \varnothing$. Thus, $X_1 = A \cup B$ and $X_2 = \varnothing$.

4. Consider the function $F\colon K \to \mathbb{R}$ defined by $F(x) = \rho(x, f(x))$. Since K is compact, this function attains its minimum at some point x_0. If $F(x_0) = 0$, then x_0 is fixed. Suppose that $F(x_0) = d > 0$. Then $F(f(x_0)) = \rho(f(x_0), f(f(x_0))) < \rho(x_0, f(x_0)) = d$, which contradicts the assumption.

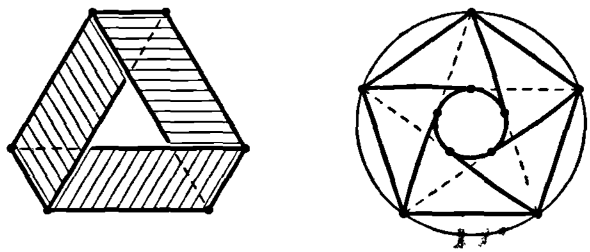

Figure H.1. Embeddings of the graphs $K_{3,3}$ and K_5

5. Yes, this is possible. The embeddings are shown in Figure H.1.

6. Consider a straight line parallel to none of the lines joining a vertex of one cycle to a vertex of the other cycle. A translation of one of the cycles along this line does not change the intersection number. Indeed, each vertex of a cycle is incident to exactly two edges; therefore, as the vertex passes though an edge, the number of intersection points changes by ± 2, and its remainder after division by 2 does not change.

When the cycle under consideration is moved far enough to become disjoint from the other cycle, the intersection number vanishes.

7. (a) First, suppose that the vertices of the graph are fixed and only the arrangement of the edges may change. Two positions of the same edge form a cycle. According to Problem 6, the intersection number of this cycle and the cycle constituted by the nonadjacent edges is 0. Therefore, the self-intersection number does not depend on the position of one edge, and hence on the positions of all other edges.

The self-intersection number does not depend on the arrangement of the vertices either, because, for any n points, there exists a homeomorphism of the plane which takes them to any other n points.

(b) The graphs $K_{3,3}$ and K_5 have the property specified in (a); therefore, for each of these graphs, the self-intersection number does not depend on how the graph is drawn in the plane. Hence we can calculate the self-intersection number for any plane image of the graph. It equals 1 for both graphs. In particular, there is always a self-intersection point.

8. (a) It is sufficient to prove that any cycle C in the graph G has even length. The cycle C encloses several faces. Removing an interior face that shares an edge with C, we replace a part of C consisting of n_1 edges by a part consisting of n_2 edges, where $n_1 + n_2$ is the number of edges in the removed face, which is even. Thus, such a transformation does not affect the parity of the cycle length. Several such transformations give a cycle bounding one face. Its length is even.

(b) Let us choose one point in each domain and join the points belonging to domains adjacent along arcs by edges. As a result, we obtain a graph G in which every face contains an even number of edges. (The faces correspond to the self-intersections points of the curve γ; to a point of multiplicity k correspond faces with $2k$ edges.)

9. Any face of the graph K_5 must contain at least three edges. Thus, if K_5 were planar, then the inequality $e \leq 3v - 6$ would hold; on the other hand, $e = 10$ and $v = 5$ for K_5. In the graph $K_{3,3}$, any face must contain at least four edges, which implies $e \leq 2v - 4$ for a planar graph, while for $K_{3,3}$, we have $e = 9$ and $v = 6$.

10. The required homotopy is defined by

$$h_t(A, B) = \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix},$$

where $0 \leq t \leq \pi/2$.

11. Let X be a path-connected space. Suppose that X can be represented as the union of two nonempty disjoint open sets A and B. Take points $a \in A$ and $b \in B$ and consider a path $f\colon [0,1] \to X$ for which $f(0) = a$ and $f(1) = b$. The interval $[0,1]$ is the union of the nonempty disjoint open sets $A_0 = f^{-1}(A)$ and $B_0 = f^{-1}(B)$. We know that $1 \in B_0$. Let t_0 be the least upper bound for the numbers from A_0. If $t_0 \in A_0$, then A_0 cannot be open, and if $t_0 \in B_0$, then B_0 cannot be open.

12. Each path-connected component of X is open. Indeed, every point $x \in X$ is contained in an open path-connected set U. Clearly, U is contained in the path-connected component of x. Thus, any path-connected component is a union of open sets.

The complement of a path-connected component is the union of the other path-connected components, which are open. Therefore, every path-connected component is both open and closed.

13. For each point $x \in U$, we can choose an open ball centered at x and contained in U. Each ball is path-connected. By Problem 12, U is path-connected.

14. Each of the sets X_1 and X_2 is path-connected and, therefore, connected. Let A be a clopen (i.e., both open and closed) set in X. Then $A_1 = X_1 \cap A$ is a clopen subset of X_1. Hence $A_1 = X_1$ or $\varnothing$. Similarly, $A \cap X_2 = X_2$ or $\varnothing$. Thus, only X_1 and X_2 can be nontrivial clopen subsets of X. But the set X_2 is not closed, because its closure contains X_1. Therefore, the space X is connected.

Let us prove that X is not path-connected. Consider any continuous path $\gamma\colon [0,1] \to X$ for which $\gamma(0) \in X_1$. Let t_0 be the least upper bound for all t such that $\gamma(t) \in X_1$. By continuity, $\gamma(t_0) \in X_1$, i.e., $\gamma(t_0) = (0, y_0)$.

Choose $\varepsilon > 0$ such that if $|t-t_0| < \varepsilon$, then $\gamma(t) = (x, y)$, where $|y-y_0| < 1/2$. Consider the projection on the Ox-axis of the intersection of X with the set defined by $|y - y_0| < 1/2$. Between the origin and any other point of the projection, there is a point not belonging to the projection. Therefore, if $|t - t_0| < \varepsilon$, then $\gamma(t) \in X_1$, and the path $\gamma(t)$ is entirely contained in X_1.

15. It is sufficient to prove that in the spaces under consideration each matrix can be joined by a path to the identity matrix I_n.

Any nonsingular matrix A can be represented as $A = SU$, where S is a symmetric positive definite matrix and U is an orthogonal matrix. Let V be an orthogonal matrix such that $S = VDV^{-1}$, where D is a diagonal matrix with positive diagonal elements. Then $A = VDW$, where $W = V^{-1}U$ is an orthogonal matrix. The path $V[(1-t)D+tI]W$ joins A with the orthogonal matrix VW in $\mathrm{GL}(n)$. Moreover, if $\det A > 0$, then $\det(VW) > 0$.

Now, let us prove that any matrix $U \in \mathrm{SO}(n)$ can be joined with the identity matrix I by a path in $\mathrm{SO}(n)$. In some orthonormal basis, the matrix U is the direct sum of the identity matrix and matrices of the form $\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$. The required path can be constructed by considering the family of matrices $\begin{pmatrix} \cos t\varphi & -\sin t\varphi \\ \sin t\varphi & \cos t\varphi \end{pmatrix}$, where $t \in [0, 1]$.

For the unitary matrices, we can use the fact that any unitary matrix is diagonal in some orthonormal basis and the diagonal elements have the form $e^{i\varphi}$. The required path consists of the diagonal matrices with diagonal elements $e^{i\varphi t}$, where $t \in [0, 1]$, in this basis. For the space $\mathrm{SU}(n)$, the same construction can be used, because if $e^{i\varphi_1} \ldots e^{i\varphi_n} = 1$, then $e^{i\varphi_1 t} \ldots e^{i\varphi_n t} = 1$.

16. (a) Suppose that the covering $p\colon K_n \to G$ is of degree $2m$. Then the preimage of any vertex v of the graph G consists of $2m$ vertices $v_1, \ldots, v_{2m}$, and these vertices generate a subgraph K_{2m} with $m(2m-1)$ edges in K_n. Each of these edges is projected to a loop based at v. Suppose that the number of the loops thus obtained equals l. The preimage of every loop consists of $2m$ edges; therefore, $2ml = m(2m-1)$, i.e., $l = (2m-1)/2$, which is impossible.

(b) The covering $p\colon K_{2m+1} \to G$, where G is a graph consisting of one vertex and m loops, is of odd degree.

17. Suppose that a map f is null-homotopic. Let $x_0 \in S^1$ be the base point and $y_0 = f(x_0)$. Consider the loop $\omega(t) = x_0 \exp(2\pi it)$ at x_0. The map f takes this loop to a contractible loop $\widetilde{\omega}(t)$. Take a point $z_0 \in \mathbb{R}$ for which $\exp(2\pi i z_0) = y_0$ and consider the lifting $\Omega(t)$ of $\widetilde{\omega}(t)$ starting at z_0. The loop $\widetilde{\omega}(t)$ is contractible; therefore, the path $\Omega(t)$ is closed. This means that the formula $f_2(\omega(t)) = \Omega(t)$ defines a map $f_2\colon S^1 \to \mathbb{R}$. It remains to set $f_1(t) = \exp(2\pi it)$.

Figure H.2. A covering with automorphism group $\mathbf{Z}$

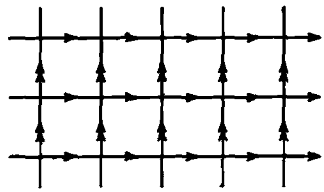

Figure H.3. A covering with automorphism group $\mathbf{Z} \oplus \mathbf{Z}$

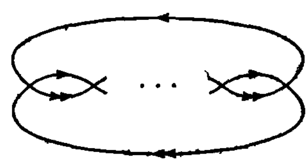

Figure H.4. A covering with automorphism group $\mathbf{Z}_n$

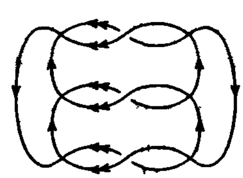

Figure H.5. A covering with automorphism group $\mathbf{Z}_2 \oplus \mathbf{Z}_3$

18. The required coverings are shown in Figures H.2 H.5; each of the figures, except the first, shows only the covering space.

19. Consider the wedge of $\operatorname{rk} G$ circles, i.e., a one-dimensional complex with one vertex and $\operatorname{rk} G$ edges. Let us construct a covering of this complex whose automorphism group H is a subgroup of G. The group H is isomorphic to the fundamental group of the covering space $\widetilde{X}$. The space $\widetilde{X}$ is homeomorphic to a wedge of circles; therefore, the group H is free.

The covering under consideration is k-fold; hence $\widetilde{X}$ has k vertices and $k(\operatorname{rk} G)$ edges. Any maximal tree in $\widetilde{X}$ contains $k-1$ edges; therefore, contracting a maximal tree to a point, we obtain a one-dimensional complex with one vertex and $k(\operatorname{rk} G) - (k-1) = (\operatorname{rk} G - 1)k + 1$ edges. This number is equal to the rank of H.

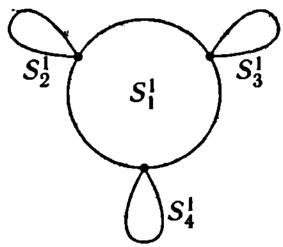

Figure H.6. The covering space

20. To construct a monomorphism $F_n \to F_2$, it suffices to construct a covering for which the base is $S^1_a \vee S^1_b$ and the covering space is homotopy equivalent to $S^1_1 \vee \dots \vee S^1_n$. This covering can be constructed, e.g., as follows. Let us uniformly arrange circles $S^1_2, \dots, S^1_n$ on S^1_1 (see Figure H.6); the space thus obtained is homotopy equivalent to $S^1_1 \vee \dots \vee S^1_n$. Consider a map which is an $(n-1)$-fold covering of S^1_a on S^1_1 and identically maps each of the remaining circles $S^1_2, \dots, S^1_n$ onto S^1_b.

In the algebraic language, the map $F_n \to F_2$ is described as follows. Let $x_1, \dots, x_n$ be generators of the group F_n, and let a and b be generators of F_2. Then $x_1 \mapsto a^{n-1}$, $x_2 \mapsto b$, $x_3 \mapsto aba^{-1}$, $x_4 \mapsto a^2ba^{-2}$, $\dots$, $x_n \mapsto a^{n-2}ba^{n-2}$. For the group F_∞, the map is $x_k \mapsto a^kba^{-k}$.

21. No, this is not true. For example, if $A = \{0\}$, $B = [0,1]$, and $C = \{1\}$, then $d(A,B) = d(B,C) = 0$ and $d(A,C) = 1$.

22. Suppose that $d_H(A,B) = \beta$ and $d_H(B,C) = \gamma$. Take $\varepsilon > 0$. For any point $a \in A$, we can choose a point $b \in B$ such that $\|a-b\| \le \beta+\varepsilon$. For $b \in B$, we can choose a point $c \in C$ such that $\|b-c\| \le \gamma+\varepsilon$. Therefore, $\|a-c\| \le \beta+\gamma+2\varepsilon$. Similarly, for a point $c \in C$, we can choose a point $a \in A$ such that $\|a-c\| \le \beta+\gamma+2\varepsilon$.

23. Let $\varepsilon > 0$. For each point $x \in f(A)$, consider the set $U_x = A \cap f^{-1}(D^n_{x,\varepsilon/2})$, where $D^n_{x,\varepsilon/2}$ is the open disk of radius $\varepsilon/2$ centered at x. Such sets form an open cover of the topological space A. Let $\delta > 0$ be the Lebesgue number of this cover. If $a_1, a_2 \in A$ and $|a_1 - a_2| < \delta$, then $a_1, a_2 \in U_x$ for some point x, and the points $f(a_1)$ and $f(a_2)$ belong to the open disk of radius $\varepsilon/2$ centered at x, i.e., $|f(a_1) - f(a_2)| < \varepsilon$.

24. Let $r\colon A \to X$ be a retraction. If $f\colon A \to Y$ is a continuous map, then fr is an extension of f to X. On the other hand, if each continuous map $f\colon A \to Y$ can be extended to X, then the identity map $\mathrm{id}_A\colon A \to A$ can be extended to $r\colon A \to X$. This r is the required retraction.

25. Let $f\colon A \to A$ be a continuous map. According to Problem 24, this map can be extended to a map $F\colon X \to A \subset X$. By assumption, F has a fixed point x_0. We have $x_0 = F(x_0) \in A$ and $f(x_0) = F(x_0) = x_0$.

26. Let $\{U_\alpha\}$ be an open cover of C. Choose open sets U'_α in K for which $U_\alpha = U'_\alpha \cap C$. The sets U'_α, together with the open set $U = K \setminus C$, form a cover of K. This cover has a finite subcover $U'_1, \dots, U'_n, U$. Clearly, the sets $U_1, \dots, U_n$ cover C.

27. (a) Let us introduce an equivalence relation on X. We declare that $x_1 \sim x_2$ if the images of x_1 and x_2 under any continuous map from X to a Hausdorff space coincide. Then $X^H = X/\sim$, and σ is the natural projection of X onto $X/\sim$.

(b) Consider the surjective map $\mathrm{Mat}_n(\mathbb{C}) \to \mathbb{C}^n$ that assigns the coefficients of the polynomial $\det(A + \lambda I)$, where I is the identity matrix, to each matrix A. This map is constant on the orbits; therefore, it induces a map $c\colon X/G \to \mathbb{C}^n$. The map c corresponds to a surjective map $c^H\colon (X/G)^H \to \mathbb{C}^n$.

If A and B are diagonal matrices, then $c^H(A) = c^H(B)$ if and only if A and B belong to the same orbit.

Any orbit contains an upper triangular matrix

$$A = \begin{pmatrix} \lambda_1 & \dots & * \\ \dots & \dots & \dots \\ 0 & \dots & \lambda_n \end{pmatrix}.$$

Consider the diagonal matrix $\Delta_m = \mathrm{diag}(1, m, m^2, \dots, m^{n-1})$. We have

$$\lim_{m\to\infty} \Delta_m A \Delta_m^{-1} = \mathrm{diag}(\lambda_1, \dots, \lambda_n).$$

Since the space $(X/G)^H$ is Hausdorff, it follows that the matrices A and $\mathrm{diag}(\lambda_1, \dots, \lambda_n)$ represent the same point in this space.

(c) The map f induces a map $\bar{f}\colon X/G \to \mathbb{C}$. To the map $\bar{f}$ corresponds a map $F\colon (X/G)^H = \mathbb{C}^n \to \mathbb{C}$.

28. Let $v_0, v_1, \dots, v_n$ be the vertex set (ordered somehow) of the simplex Δ^n. Corresponding to it is the simplex in the barycentric subdivision which is determined by the inequalities $x_0 \geq x_1 \geq \dots \geq x_n$ in barycentric coordinates. The vertices of this simplex are v_0, the barycenter of $[v_0, v_1]$, the barycenter of $[v_0, v_1, v_2]$, $\dots$.

29. Obviously, any complete subcomplex does have the required property. Suppose that each simplex of K whose boundary is contained in L is itself contained in L. Take a simplex from K such that all of its vertices belong to L. Then all the edges of this simplex belong to L, and therefore all of its 2-faces belong to L, etc.

30. Each simplex Δ^n of K' is uniquely determined by a set of simplices $\sigma_0 \subsetneqq \sigma_1 \subsetneqq \dots \subsetneqq \sigma_n$ from the complex K (the vertices of Δ^n are the barycenters of these simplices). Suppose that all the vertices of Δ^n belong to L'!

Then, in particular, the barycenter of σ_n belongs to L'. This means that the simplex σ_n itself belongs to L. The simplex Δ^n is one of the simplices in the barycentric subdivision of σ_n; therefore, it belongs to L'.

31. (a) [140] We assume that $I = [-1,1]$. Then ∂I^n consists of the points $(x_1,\dots,x_n)$ such that $|x_i| \le 1$ for all i and $x_i = \pm 1$ for some i. We set

$$(I^n)_+^{n-1} = \{(x_1,\dots,x_n) \in I^n : x_i = +1 \text{ for some } i\}.$$

It is easy to verify that $\partial((I^n)_+^{n-1})$ consists of the points $(x_1,\dots,x_n) \in I^n$ such that $x_i = +1$ and $x_j = -1$. Let us define $(I^n)_+^{n-2}$ as the union of the $(n-2)$-faces of the cube determined by the relations $x_i = +1$ and $x_j = -1$ for $i < j$. Similarly, we define $(I^n)_+^{n-j}$ as the union of the $(n-j)$-faces of the cube determined by $x_{a_1} = +1$, $x_{a_2} = -1$, $x_{a_3} = +1, \dots$, $x_{a_j} = (-1)^{j+1}$ for some $1 \le a_1 < a_2 < a_3 < \dots < a_j \le n$. It is easy to show that $\partial((I^n)_+^{n-j}) = (I^n)_+^{n-j-1} \cup (-(I^n)_+^{n-j-1})$.

Let $S_k(i_0, i_1, \dots, i_k)$ be the number of k-simplices with labels $i_0, \dots, i_k$ in $(I^n)_+^k$. We derive a relation between $S_1(i,j)$ and $S_0(i)$ by calculating in two different ways the number $N(i)$ of pairs, each consisting of a 1-simplex belonging to $(I^n)_+^1$ and its vertex with label $i \ge 1$. First, consider the sum over the 1-simplices from $(I^n)_+^1$. We obtain

$$N(i) = 2S_1(i,i) + S_1(i,-i) + \sum_{j\ne i, j\ge 1} (S_1(i,j) + S_1(i,-j)); \tag{1}$$

we have $S_1(i,-i) = 0$ by assumption. Now, consider the sum over the vertices of the triangulation which belong to $(I^n)_+^1$. We obtain

$$N(i) = 2K + S_0(i) + S_0(-i), \tag{2}$$

where K is the number of interior vertices in $(I^n)_+^1$ labeled by i. Indeed, the boundary of $(I^n)_+^1$ consists of $(I^n)_+^0$ and $-(I^n)_+^0$, and if a vertex $v \in (I^n)_+^0$ has label k, then the vertex $-v \in -(I^n)_+^0$ has label $-k$. Comparing (1) with (2) and summing over i between 1 and n, we obtain

$$\sum_{1\le i<j\le n} (S_1(i,-j) + S_1(-i,j)) \equiv \sum_{i=1}^{n} (S_0(i) + S_0(-i)) \pmod 2;$$

the terms of the form $S_1(i,j)$ are canceled because each of them occurs twice, once in the expression for i and once in the expression for j. Clearly, $\sum(S_0(i)+S_0(-i)) = 1$ because $(I^n)_+^0$ consists of one point $(+1,-1,+1,\dots)$.

The further calculations are modulo 2. Let us count the pairs in $(I^n)_+^2$, each consisting of a 2-simplex and its 1-face with labels i and $-j$, where $1 \le i < j$. First, we count them by taking a sum over the 2-simplices.

We obtain (modulo 2)

$$\sum_{k\neq i,j,k\geq 1} (S_2(i,-j,k) + S_2(i,-j,-k)).$$

Summation over the 1-simplices with labels i and $-j$ gives

$$S_1(i,-j) + S_1(-i,j);$$

the interior 1-simplices cancel each other modulo 2, and only the boundary simplices remain. Let us equate the obtained expressions to each other, take the sum over all pairs $i < j$, and reduce the result modulo 2. We obtain

$$\sum_{1\leq i_0<i_1<i_2} (S_2(i_0,-i_1,i_2) + S_2(-i_0,i_1,-i_2)) = \sum_{1\leq i<j} (S_1(i,-j) + S_1(-i,j))$$
$$= \sum_{1\leq i\leq n} (S_0(i) + S_0(-i)) = 1.$$

Continuing the calculations in a similar way, we come to

(3) $\qquad S_{n-1}(1,-2,\dots,\pm n) + S_{n-1}(-1,2,\dots,\mp n) \equiv 1 \pmod 2.$

Every time we use the fact that if two neighboring numbers i_a and i_{a+1} are of the same sign, then the expression $S_k(i_0,\dots,i_a,i_{a+1},\dots,i_k)$ occurs twice.

To obtain a contradiction, we count the pairs consisting of n-simplices and their $(n-1)$-faces with labels $1,-2,3,\dots,\pm n$ in the entire cube I^n. The vertex opposite to such a face cannot be labeled by $-1,2,-3,\dots,\mp n$ because if it were, then we would have an edge with labels i and $-i$. Thus, precisely one of the labels $1,-2,3,\dots,\pm n$ occurs twice. Therefore, the required number of pairs is even (we count the pairs by summing over the n-simplices). On the other hand, counting the same pairs by summing over the $(n-1)$-simplices with labels $1,-2,3,\dots,\pm n$, we obtain the expression on the left-hand side of (3), which is odd.

(b) Suppose that there exists a continuous map $f\colon I^n \to \partial I^n$ that takes antipodal points of ∂I^n to antipodal points. Consider a sufficiently fine triangulation I^n such that it is symmetric on ∂I^n and if v_1 and v_2 are adjacent vertices, then $\|f(v_1) - f(v_2)\| < 2$; in particular, the images of adjacent vertices cannot belong to opposite faces of the cube. We label each vertex $v \in I^n$ of the triangulation by the number of the face of I^n that contains $f(v)$; it is assumed that opposite faces have numbers i and $-i$. According to Tucker's lemma, some adjacent vertices of the triangulation are labeled by i and $-i$. This contradicts the definition of this triangulation.

32. The Borsuk–Ulam theorem implies $n \leq m$. Take an arbitrary $(m+1-n)$-dimensional linear subspace Π of $\mathbb{R}^{m+1} \supset S^m$. Let us show that the set $\Pi \cap \varphi(S^n)$ contains at least two points. Let $\Pi^\perp$ be the orthogonal complement of Π, and let $p\colon \mathbb{R}^{m+1} \to \Pi^\perp$ be the orthogonal projection.

Then the map $p \circ \varphi \colon S^n \to \Pi^\perp \cong \mathbb{R}^n$ is odd, and hence, by the Borsuk Ulam theorem, there exists a point $x \in S^n$ for which $p(\varphi(x)) = 0$, i.e., $\varphi(x) \in \Pi$. We have $\varphi(-x) = -\varphi(x) \in \Pi$ and $\varphi(x) \neq 0$ because $\varphi(x) \in S^m$.

Almost all $(m+1-n)$-dimensional linear subspaces of $\mathbb{R}^m$ intersect any fixed $(n+1)$-dimensional linear subspace in a straight line; i.e., they intersect the sphere S^n standardly embedded in $\mathbb{R}^{n+1} \subset \mathbb{R}^m$ in exactly two points. Thus, almost all $(m+1-n)$-dimensional subspaces have at least as many intersection points with $\varphi(S^n)$ as with S^n. Therefore, the n-dimensional volume of $\varphi(S^n)$ is not less than that of S^n. To show this, we introduce an invariant measure μ on the set $G(m, m+1-n)$ of all $(m+1-n)$-dimensional linear subspaces of $\mathbb{R}^m$ and, for each $X \subset S^m$, consider the integral $\int_{G(m,m+1-n)} |X \cap \Pi| \, d\mu$. If the set X has n-volume, then this integral is finite and proportional to the n-volume of X.

33. (a) The direct application of the Borsuk Ulam theorem does not give the desired result because opposite rays emanating from an interior point of the polyhedron P may intersect ∂P at points that belong to intersecting faces. We must consider the symmetrized polyhedron P, i.e., the set

$$Q = \{x = z - w : z, w \in P\}.$$

Clearly, Q is a convex polyhedron symmetric with respect to the origin.

We define a map $h \colon Q \to P$ as follows. Take $z = (z_1, \ldots, z_{n+1})$ and $z' = (z'_1, \ldots, z'_{n+1})$. We say that $z > z'$ if $z_1 = z'_1, \ldots, z_{k-1} = z'_{k-1}$ and $z_k > z'_k$ (possibly, $k = 1$). We define

$$h(x) = \max\{z : x = z - w, \text{ where } z, w \in P\}.$$

Geometrically, the map h can be described as follows. If $x = z - w$, then $z = x + w \in x + P$; thus, $h(x)$ is the (unique) maximal point of the set $P \cap (x + P)$. To find this maximal point, we first determine all points for which the coordinate z_1 is maximal, then choose the points with maximal coordinate z_2 among them, etc.

It is easy to show that $x = h(x) - h(-x)$. Indeed, if $x = h(x) - w_0$, then $w_0 = h(x) - x \in P \cap (x + P) - x = (-x + P) \cap P$. Moreover, since $h(x)$ is the maximal point of $P \cap (x + P)$, it follows that w_0 is the maximal point of $(-x + P) \cap P$.

Now let us prove that the map h is continuous. Suppose that $x_n \in Q$ and $\lim_{n\to\infty} x_n = x \in Q$. Let us represent x_n as $x_n = z_n - w_n$, where $z_n = h(x_n)$. Take an arbitrary convergent subsequence z_{n_i}. The subsequence w_{n_i} converges also. We set $z = \lim_{i\to\infty} z_{n_i}$ and $w = \lim_{i\to\infty} w_{n_i}$. Then $x = z - w$, whence $z \leq h(x)$. Suppose that $z < h(x)$. Consider the points $z' = z_{n_i} + \varepsilon(h(x) - z)$ and $w' = w_{n_i} + \varepsilon(h(-x) - w)$, where $\varepsilon > 0$. Since $h(x) - h(-x) = x = z - w$, it follows that $z' - w' = z_{n_i} - w_{n_i} = x_{n_i}$. Clearly, $z > z_{n_i}$, because

$h(x) - z > 0$ and $\varepsilon > 0$. Let us show that the numbers $\varepsilon > 0$ and n_i can be chosen so that $z', w' \in P$. For $z \in P$, we can find $\delta > 0$ such that if $\|z - v\| \leq \delta$ and v is a point on the ray from z to $t \in P$, then $v \in P$. Let C_δ be the set of all such points v. Take $\varepsilon > 0$ such that $\varepsilon\|h(x) - z\| \leq \delta/2$ and choose n_i for which $\|z_{n_i} - z\| \leq \delta/2$. The points z_{n_i} and $z + \varepsilon(h(x) - z)$ belong to $C_{\delta/2}$. Since the set $C_{\delta/2}$ is convex, it follows that the midpoint of the segment joining these points belongs to $C_{\delta/2}$ also, whence $z' = z_{n_i} + \varepsilon(h(x) - z) \in C_\delta \subset P$. Similarly, $w' \in P$. But if $z', w' \in P$ and $z' - w' = x_{n_i}$, then $z' \leq h(x_{n_i}) = z_{n_i}$, which contradicts the inequality $z' > z_{n_i}$. This contradiction shows that $z = h(x)$. Thus, any convergent subsequence of the sequence $\{z_n\}$ converges to $h(x)$. The set P, which contains the points z_n, is compact; therefore, only finitely many points z_n can be outside an arbitrarily small neighborhood of $h(x)$. Hence $\lim_{n\to\infty} z_n = h(x)$, i.e., the map h is continuous.

Suppose that $a \neq 0$ is an arbitrary vector and $\max_{x\in Q}(a, x) = (a, x_0)$. The equality $x_0 = h(x_0) - h(-x_0)$ implies

$$\begin{aligned}(a, h(x_0)) + (-a, h(-x_0)) &= (a, x_0) = \max_{x\in Q}(a, x)\\ &= \max_{z,w\in P}(a, z - w) = \max_{z\in P}(a, z) - \max_{w\in P}(-a, w).\end{aligned}$$

Therefore, $(a, h(x_0)) = \max_{z\in P}(a, z)$ and $(-a, h(-x_0)) = \max_{w\in P}(-a, w)$. This means that the points $h(x_0)$ and $h(-x_0)$ belong to two different supporting hyperplanes of the polyhedron P. In particular, the points $h(x_0)$ and $h(-x_0)$ belong to disjoint faces of P.

Now we can apply the Borsuk Ulam theorem to the map $g(x) = f(h(x))$. It says that there exists a point $x_0 \in \partial Q$ for which $g(x_0) = g(-x_0)$. The point x_0 belongs to some supporting plane of the polyhedron Q; hence there exists a vector $a \neq 0$ for which $\max_{x\in Q}(a, x) = (a, x_0)$. The points $z = h(x_0)$ and $w = h(-x_0)$ belong to disjoint faces of the polyhedron P, and

$$f(z) = f(h(x_0)) = g(x_0) = g(-x_0) = f(h(-x_0)) = f(w),$$

as required.

(b) Let B and C be disjoint faces of the simplex Δ^{n+1} such that $f(B) \cap f(C) \neq \varnothing$. The face Δ_i^n contains all but one vertex of the simplex Δ^{n+1}. This vertex cannot belong to both faces B and C; therefore, $B \subset \Delta_i^n$ or $C \subset \Delta_i^n$ (or $B, C \subset \Delta_i^n$), which means that $f(\Delta_i^n) \supset f(B) \cap f(C)$.

Remark 1. For linear maps $f\colon \partial\Delta^{n+1} \to \mathbb{R}^n$, assertion (b) is a special case of the *Helly theorem*. But this special case is the induction step in the inductive proof of the theorem. Thus, it is essentially equivalent to the Helly theorem.

Remark 2. Another solution of Problem 2.9 and its generalization are given in [75].

34. Consider the characteristic map $f\colon D^{2n} \to \mathbb{C}P^n$ defined on p. 121. Its restriction to $\operatorname{int} D^{2n}$ is a homeomorphism of $\operatorname{int} D^{2n}$ onto $\mathbb{C}P^n \setminus \mathbb{C}P^{n-1}$. Clearly, the specified identification of points of the sphere S^{2n-1} turns the sphere into $\mathbb{C}P^{n-1}$.

35. We can assume that S^∞ consists of the points $x = (x_1, x_2, \dots) \in \mathbb{R}^\infty$ such that they have only finitely many nonzero coordinates and $\sum x_i^2 = 1$. Consider the maps $\varphi(x) = (0, x_1, x_2, \dots)$ and $h_t(x) = (1-t)x + t\varphi(x)$. It is easy to show that $h_t(x) \neq 0$ for $x \neq 0$. Therefore, the formula $x \mapsto h_t(x)/\|h_t(x)\|$ defines a homotopy between the identity map id_{S^∞} and $\varphi \restriction S^\infty$. Let $g_t(x) = (1-t)\varphi(x) + (t, 0, 0, \dots)$. Again, $g_t(x) \neq 0$ for $x \neq 0$. Therefore, the formula $x \mapsto g_t(x)/\|g_t(x)\|$ defines a homotopy between the map $\varphi \restriction S^\infty$ and the constant map to the point $(1, 0, 0, \dots)$.

36. Let K be a compact set. For every open cell $\operatorname{int} e_\alpha^n$ which K intersects, choose a point $x_\alpha^n \in K \cap \operatorname{int} e_\alpha^n$. We must prove that the set $T = \{x_\alpha^n\}$ is finite.

By condition (c), a closed cell may intersect only finitely many open cells; therefore, the intersection of any set $T' \subset T$ with any closed cell is finite and, hence, closed. By condition (w) (see p. 119), any set $T' \subset T$ is closed, which means that T is discrete. On the other hand, the set T is compact, as a closed subset of a compact space. It remains to note that all discrete compact sets are finite.

37. The sphere S^n can be represented as a CW-complex with one 0-cell and one n-cell. Hence $S^p \times S^q$ can be represented as a CW-complex with cells of dimensions 0, p, q, and $p+q$. The cells of dimensions 0, p, and q form a subcomplex $S^p \vee S^q$. Contracting this subcomplex to a point, we obtain a CW-complex with cells of dimensions 0 and $p+q$, i.e., a $(p+q)$-sphere.

38. Let $f_t(x) = (x, t, y_0)$, where y_0 is a point of Y. Then $f_0 = f$ and f_1 takes X to the point y_0.

39. (a) If n is even, then $nP^2 \approx \frac{n-2}{2}T^2 \# 2P^2$, and if n is odd, then $nP^2 \approx \frac{n-1}{2}T^2 \# P^2$. Therefore, it suffices to consider the surfaces $2P^2$ and P^2, for which the required curves are constructed in an obvious way.

(b) We must prove that if cutting along a closed curve renders the surface nP^2 orientable, then the boundary of the cut surface has two components for n even, and one component for n odd. Such a cutting does not change the Euler characteristic of the surface. If the boundary has two components, then we attach the handle $S^1 \times I$, and if it has one component, then we attach the disk D^2. In both cases, we obtain a closed orientable surface (with even Euler characteristic). In the former case, the Euler characteristic does not change, and in the latter, it increases by 1.

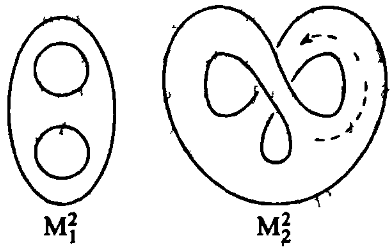

Figure H.7. The surfaces M_1^2 and M_2^2

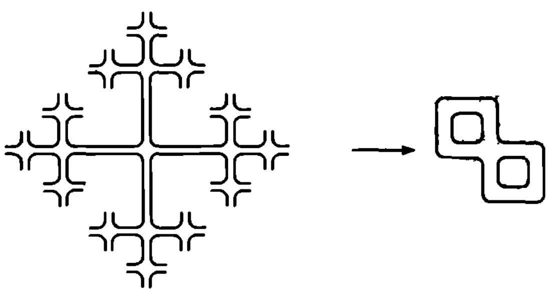

Figure H.8. The universal covering of the plane with two punctures

40. Yes, $M_2^2 \times I$ and $M_1^2 \times I$ can be homeomorphic. The surfaces M_1^2 and M_2^2 shown in Figure H.7 are not homeomorphic because the boundary of M_1^2 has three connected components and that of M_2^2 is connected, but the spaces $M_2^2 \times I$ and $M_1^2 \times I$ are homeomorphic because the "handle" can be dragged along the dashed line.

41. The surface nP^2 can be represented as the sphere S^2 in which n disks are replaced by n disjoint Möbius bands.

Suppose that there are p disjoint Möbius bands on the surface nP^2. Let us cut the surface along the boundaries of these Möbius bands and paste the holes with disks. As a result, we obtain a closed surface M^2 with Euler characteristic $\chi(nP^2) + p = 2 - n + p$. But $\chi(M^2) \leq 2$; therefore, $p \leq n$.

42. (a) The universal covering of the plane with two punctures is shown in Figure H.8. Clearly, the universal covering space is homeomorphic to the plane. For the plane with an arbitrary (finite) number of punctures, the proof is similar.

(b) Let $f_{a_1 \ldots a_n} \colon \mathbb{C} \to \mathbb{C} \setminus \{a_1, \ldots, a_n\}$ be the universal covering. Consider the map $(w_1, \ldots, w_n) \mapsto (z_1, \ldots, z_n)$, where

$$z_1 = w_1, \quad z_2 = f_{z_1} w_2, \quad z_3 = f_{z_1 z_2} w_3, \quad z_4 = f_{z_1 z_2 z_3} w_4, \quad \ldots.$$

This map is a covering $\mathbb{C}^n \to \Sigma$.

43. If a covering p can be represented in the form $p = p_2p_1$, then the sets $I_1 = p_1^{-1}(y_1), \dots, I_n = p_1^{-1}(y_n)$, where $\{y_1, \dots, y_n\} = p_2^{-1}(x)$, are as required. Indeed, if γ is a closed path in X, then its lifting to Y joins some points y_i and y_j. Therefore, the lifting of γ to $\widetilde{X}$ joins points of the sets I_i and I_j.

Now assume that $I_1 = \{t_{11}, \dots, t_{1m}\}, \dots, I_n = \{t_{n1}, \dots, t_{nm}\}$ is a partition of $p^{-1}(x)$ with the properties specified in the statement of the problem. Take a point $x_1 \in X$. Let γ be a path from x to x_1, and let J_k be the set of the ends of those liftings of γ which start in I_k. The enumeration of the sets J_k depends on the path γ, but these sets themselves do not depend on γ. Indeed, suppose that one lifting of γ' starting in I_k ends in J_k and another lifting starting in I_k ends in J_l, where $l \neq k$. Then one lifting of the closed path $\gamma'\gamma^{-1}$ starting in I_k ends in I_k and another lifting starting in I_k ends in I_l, where $l \neq k$. This is impossible.

The covering $p_1 \colon \widetilde{X} \to Y$ maps each set J_k to one point. The construction of the covering p_2 is obvious.

44. It is easy to calculate the Euler characteristic of the deleted products of the graphs $K_{3,3}$ and K_5. All of their faces are quadrangular, and each edge belongs to exactly two faces. Therefore, $2e = 4f$. The number v of vertices in the deleted product equals $n^2 - n$, where n is the number of vertices in the graph. Hence $v = 30$ for the deleted product of the graph $K_{3,3}$ and $v = 20$ for that of K_5. The number f of faces is equal to the number of ordered pairs of disjoint edges. Therefore, $f = 36$ for the deleted product of $K_{3,3}$ and $f = 30$ for that of K_5.

It remains to show that the deleted products of $K_{3,3}$ and K_5 are orientable. This can be verified directly, but such a verification is fairly tedious, because one surface consists of 36 quadrangular faces and the other consists of 30 faces. It can be obviated as follows. The deleted product of a graph is naturally embedded in the deleted join of this graph. Indeed, each pair of nonadjacent edges in the deleted join corresponds to a tetrahedron, and each pair of nonadjacent edges in the deleted product corresponds to a parallelogram; this parallelogram can be regarded as a section of a tetrahedron (see Figure H.9). The deleted joins of the graphs $K_{3,3}$ and K_5 are homeomorphic to S^3. For the graph K_5, this assertion is a special case (for $n = 1$) of Theorem 3.38 on p. 133. For the graph $K_{3,3}$, it easily follows from the observation that the deleted join of a join is a join of deleted joins (see the proof of Theorem 3.39 on p. 134). Indeed, the graph $K_{3,3}$ is the join $\mathrm{sk}_0\,\Delta^2 * \mathrm{sk}_0\,\Delta^2$; therefore, $J_2^2(K_{3,3}) = J_2^2(\mathrm{sk}_0\,\Delta^2 * \mathrm{sk}_0\,\Delta^2) = J_2^2(\mathrm{sk}_0\,\Delta^2) * J_2^2(\mathrm{sk}_0\,\Delta^2) \approx S^1 * S^1 \approx S^3$ because $J_2^2(\mathrm{sk}_0\,\Delta^2) \approx S^1$.

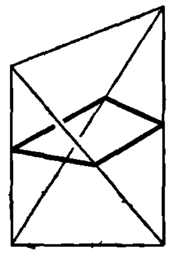

Figure H.9. The embedding of the deleted product in a deleted join

Thus, the deleted products of the graphs $K_{3,3}$ and K_5 can be embedded in S^3. On the other hand, no closed nonorientable surface can be embedded in S^3 (see the corollary of Theorem 5.21 on p. 218).

45. The map $\Delta^i_\alpha \times \Delta^j_\beta \longrightarrow \Delta^j_\beta \times \Delta^i_\alpha$ determines an involution on the deleted product; this involution has no fixed points because the simplices Δ^i_α and Δ^j_β are disjoint.

The Euler characteristic of any two-dimensional surface that admits a fixed-point-free involution is even.

46. (a) The map id_Y is homotopic to the map $Y \to y_0$. Consider the map $f_0 = \mathrm{id}_X$ on X and a homotopy between the maps id_Y and $Y \to y_0$ on Y. Extending this homotopy to a homotopy f_t of f_0, we obtain a map $f_1\colon X \to X$ such that it is homotopic to id_X and $f_1(Y) = y_0$. This map induces a map $q\colon X/Y \to X$ for which $qp = f_1$, where $p\colon X \to X/Y$ is the canonical projection. Thus, $qp = f_1 \sim \mathrm{id}_X$. It remains to show that $pq \sim \mathrm{id}_Y$. By construction, $f_t(Y) \subset Y$. Therefore, each map f_t determines a map $g_t\colon X/Y \to X/Y$; we have $g_0 = \mathrm{id}_Y$ and $g_1 = pq$.

(b) In the CW-complex $X \cup CY$, the subcomplex CY is contractible; therefore, $X \cup CY \sim (X \cup CY)/CY = X/Y$.

47. We prove the required assertion by induction. Suppose that X is an n-connected CW-complex with exactly one vertex and no k-cells, where $1 \le k \le n-1$ (for $n = 0$, no assumptions are made). We must "annihilate" the n-cells of X. Let $\varphi\colon S^n \to X$ be the characteristic map of some n-cell of the complex X. Since X is n-connected, it follows that the map φ can be extended to a map $\bar\varphi\colon D^{n+1} \to X$. (For $n = 0$, we assume that $\varphi\colon S^0 \to X$ takes one point of S^0 to a fixed vertex x_0 and the other point to a given vertex x_i of X; then $\bar\varphi$ is a path from x_0 to x_i.) We think of S^n as the equator of the sphere $S^{n+1} = \partial D^{n+2}$, and of D^{n+1} as a half of the sphere S^{n+1} (see Figure H.10). Attaching D^{n+1} to X via $\bar\varphi$, we obtain a CW-complex $X \cup Y$, which is homotopy equivalent to X (the complex Y for $n = 0$ is hatched in Figure H.11). The complex Y has the contractible subcomplex Y' which corresponds to the "upper" hemisphere D^{n+1} (this

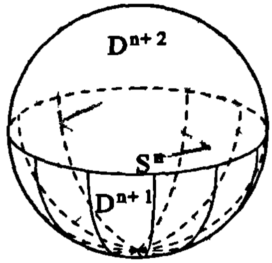

Figure H.10. The disk D^{n+1}

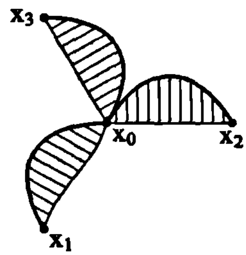

Figure H.11. The complex Y

subcomplex is shown by bold lines in Figure H.11). According to Problem 46, $(X \cup Y)/Y' \sim X \cup Y \sim X$. Clearly, the CW-complex $(X \cup Y)/Y'$ has no n-cells (and cells of smaller positive dimensions). For $n = 0$, this construction gives a CW-complex with one vertex.

48. According to the cellular approximation theorem, any map $S^n \to X$ is homotopic to a map $S^n \to X^n \subset X$; in the case under consideration, the n-skeleton X^n consists of one point.

49. The space $\Sigma(A \wedge B)$ is obtained from $A * B$ by contracting two cones CA and CB with common generatrix $\{a_0\} \times \{b_0\} \times I$ to a point (Figure H.12). Clearly, the space $CA \cup CB$ is contractible (the cones can be contracted to points one after another). Therefore, according to Problem 46, $A * B \sim A * B/(CA \cup CB) \approx \Sigma(A \wedge B)$.

50. (a) According to Problem 47, we have $X \sim X'$, where X' is a CW-complex with one vertex x_0 and no k-cells for $1 \leq k \leq n$. Clearly, $\Sigma X \sim \Sigma X'$. The complex $\Sigma X'/\Sigma x_0 \sim \Sigma X'$ has no cells of dimensions between 1 and $n+1$. By Problem 48, it is $(n+1)$-connected.

(b) We can assume that X and Y have no cells of positive dimensions greater than or equal to n and m, respectively. Then the cells of positive dimensions in the CW-complex $X \times Y$ that do not belong to $X \vee Y$ are products of cells $\sigma^p \times \sigma^q$, where $p \geq n+1$ and $q \geq m+1$; the quotient

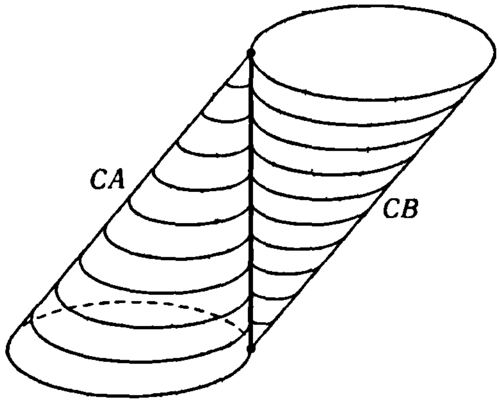

Figure H.12. The two cones

of $X \times Y$ modulo $X \vee Y$ contains only the 0-cell and such product cells. Therefore, $X \wedge Y$ has no k-cells for $1 \leq k \leq n+m+1$.

(c) According to Problem 49, we have $X * Y \sim \Sigma(X \wedge Y)$. It remains to apply (a) and (b).

51. Assertion (a) is a special case of (b). Thus, we prove (b).

The join $X * Y$ contains the distinguished subspaces X and Y; let us attach CX and CY to them. Contracting each of the cones CX and CY to a point, we obtain $\Sigma(X \times Y)$. These cones are contractible subspaces; therefore, $X * Y \cup CX \cup CY \sim \Sigma(X \times Y)$.

Let $x_0 \in X$ and $y_0 \in Y$ be base points. Consider the subspace Z of $X * Y$ which consists of $\{x_0\} * Y$ and $X * \{y_0\}$. The space Z is contractible because the space obtained by contracting the segment $[x_0, y_0]$ to a point in Z is homotopy equivalent to the wedge of two cones. Clearly, contracting the subspace Z to a point in $X * Y \cup CX \cup CY$, we obtain $\Sigma X \vee \Sigma Y \vee \Sigma(X \wedge Y)$.

52. Suppose that there exists a homotopy $h_t \colon S^n \to B$ such that $h_0 = p$ and h_1 takes S^n to a point $b_0 \in B$. The map $\mathrm{id}_{S^n} \colon S^n \to S^n$ covers $p = h_0$; therefore, for the homotopy h_t, we can consider the lifting $k_t \colon S^n \to S^n$, where $k_0 = \mathrm{id}_{S^n}$ and k_1 takes S^n to $p^{-1}(b_0)$. Since B consists of more than one point, it follows that $S^n \setminus p^{-1}(b_0) \neq \varnothing$, and the map $k_1 \colon S^n \to S^n$ is homotopic to a constant. Thus, the map id_{S^n} is homotopic to a constant, which is impossible.

53. By the cellular approximation theorem, any map $S^k \to X$, where $k \leq n$, is homotopic to a map $S^k \to X^n \subset X$; moreover, we can assume that the homotopy leaves the point $x_0 \in X^n$ fixed. For $k \leq n-1$, any map $S^k \times I \to X$ that is a homotopy between two maps $f_0, f_1 \colon S^k \to X$ is

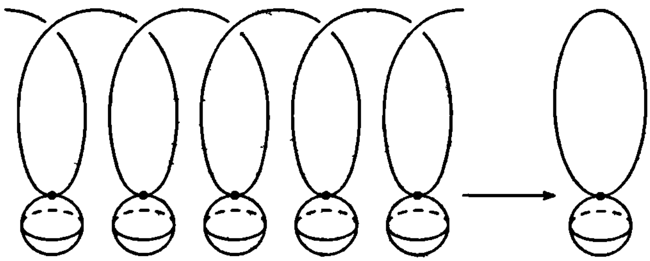

Figure H.13. The universal covering space

homotopic to a map $S^k \times I \to X^n \subset X$, and it coincides with f_0 and f_1, respectively, for $t = 0$ and $t = 1$.

54. For $n \geq 2$, the universal covering space of $S^n \vee S^1$ is the line $\mathbb{R}$ to which n-spheres are attached at integer points (see Figure H.13). This space is homotopy equivalent to the wedge of countably many n-spheres.

For $n \geq 2$, the nth homotopy groups of the base and of the covering space are isomorphic.

55. The group $\pi_n(S^n \vee S^1, x_0)$ is the free group on countably many generators α_k ($k \in \mathbb{Z}$). The action of the generator of the fundamental group $\pi_1(S^n \vee S^1, x_0)$ takes α_k to $\alpha_{k\pm1}$; thus, this action is nontrivial.

56. Any quaternion q can be represented in the form

$$q = x_1 + x_2 i + x_3 j + x_4 ij = x_1 + x_2 i + (x_3 + x_4)j = z_1 + z_2 j.$$

The group S^1 acts on S^3 by the rule $z(z_1 + z_2 j) = zz_1 + zz_2 j$. This action coincides with the one used to construct the Hopf fibration.

57. The change of variables $x_1 = u_1 + u_2$, $x_4 = u_1 - u_2$, $x_2 = u_3 + u_4$, $x_3 = u_3 - u_4$ determines a homeomorphism of the sphere under consideration onto the sphere $u_1^2 + u_2^2 + u_3^2 + u_4^2 = \frac{1}{2}$ and transforms the equation $x_1x_4 - x_2x_3 = 0$ into $u_1^2 - u_2^2 - u_3^2 + u_4^2 = 0$, i.e., $u_1^2 + u_4^2 = u_2^2 + u_3^2$.

58. As explained on p. 120, $\mathbb{C}P^2$ can be obtained by attaching $D^4 = \{(z_1, z_2) \in \mathbb{C}^2 : |z_1|^2 + |z_2|^2 \leq 1\}$ to $\mathbb{C}P^1 = \{(z_1 : z_2 : z_3) \in \mathbb{C}P^2 : z_3 = 0\}$ via the map $f\colon S^3 \to \mathbb{C}P^1$ defined as $f(z_1, z_2) = (z_1 : z_2)$. But this map f coincides with p.

59. It follows from Problem 58 that any retraction $r\colon \mathbb{C}P^2 \to \mathbb{C}P^1$ determines a map $\bar{r}\colon D^4 \to \mathbb{C}P^1$ such that $\bar{r}(x) = p(x)$ for $x \in S^3$. Therefore, $\bar{r}$ is an extension of p to D^4; thus, p is homotopic to a constant map. But the map p induces an isomorphism $p_*\colon \pi_3(S^3) \to \pi_3(S^2)$, where $\pi_3(S^3) = \mathbb{Z}$ (see p. 235); hence it cannot be homotopic to a constant map.

60. Let us represent S^3 as a union of two solid tori. The complement of the Hopf link in S^3 is obtained by removing the central circles of these solid tori. It is easy to construct a homotopy between a solid torus without the central circle and its boundary T^2. The two solid tori under consideration have common boundary.

61. The required isomorphism follows from the exact sequence for a pair,

$$\pi_n(CX) \to \pi_n(CX, X) \to \pi_{n-1}(X) \to \pi_{n-1}(CX),$$

because the cone CX is a contractible space.

62. Let us represent S^m and S^n as $I^m/\partial I^m$ and $I^n/\partial I^n$, respectively. Thereby, we represent $S^m \times S^n$ as the cube I^{n+m} in which some points of ∂I^{n+m} are identified; the identification turns ∂I^{n+m} into $S^m \vee S^n$. Let $a \in I^{n+m} \setminus \partial I^{n+m}$. It is clear how to construct a deformation retraction of $I^{n+m} \setminus \{a\}$ onto ∂I^{n+m}. This gives a deformation retraction of the punctured product $S^m \times S^n$ onto $S^m \vee S^n$.

63. The set $N^n \subset M^n$ is closed (because N^n is compact) and open (because N^n has no boundary and the manifolds N^n and M^n have equal dimensions). Therefore, it coincides with M^n.

64. (a) The manifold $G_+(n, 1)$ is diffeomorphic to S^{n-1}, and $G_+(3, 2) \approx G_+(3, 1)$. In what follows, we assume that $k \geq 2$ and $n \geq 4$. Let us calculate $\pi_1\big(G(n, k)\big)$ by using Theorem 6.1. We are interested only in the 2-skeleton of $G(n, k)$. The 2-skeleton includes Schubert cells of the forms $\left(\begin{smallmatrix} 1 & 0 & 0 & 0 \\ 0 & t & 1 & 0 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} x & 1 & 0 & 0 \\ y & 0 & 1 & 0 \end{smallmatrix}\right)$, and $\left(\begin{smallmatrix} 1 & 0 & 0 & 0 \\ 0 & x & y & 1 \end{smallmatrix}\right)$; it is assumed that the remaining part of the matrix consists of zeros and ones, and so it is of no interest to us. Let us show that both 2-cells are attached to the 1-cell in precisely the same way as in $\mathbb{R}P^2$. As $x, y \to \infty$, $\left(\begin{smallmatrix} 1 & 0 & 0 & 0 \\ 0 & x & y & 1 \end{smallmatrix}\right) \to \left(\begin{smallmatrix} 1 & 0 & 0 & 0 \\ 0 & x/y & 1 & 0 \end{smallmatrix}\right)$ and

$$\begin{pmatrix} x & 1 & 0 & 0 \\ y & 0 & 1 & 0 \end{pmatrix} \sim \begin{pmatrix} x & 1 & 0 & 0 \\ 0 & -y/x & 1 & 0 \end{pmatrix} \hookrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -y/x & 1 & 0 \end{pmatrix}.$$

In both cases, any antipodal boundary points of the 2-cell are attached to one point of the 1-cell.

Thus, the group $\pi_1\big(G(n, k)\big)$ is defined by one generator α and the relation $\alpha^2 = 1$ (the two 2-cells determine the same relation). Hence $\pi_1\big(G(n, k)\big) = \mathbb{Z}_2$.

The space $G_+(n, k)$ doubly covers $G(n, k)$; hence $\pi_1\big(G_+(n, k)\big) = 0$. This implies the orientability of $G_+(n, k)$.

(b) Since $k < n$, for any two sets of k orthonormal vectors in $\mathbb{R}^n$, there exists an element of $\mathrm{SO}(n)$ which transforms one of them into the other; therefore, the group $\mathrm{SO}(n)$ acts transitively on $G(n, k)$, i.e., any k-dimensional subspace Π^k of $\mathbb{R}^n$ can be transformed into any other k-dimensional subspace by an element of $\mathrm{SO}(n)$.

The stationary subgroup of each point $\Pi^k \in G(n,k)$, which consists of the transformations taking Π^k to itself, is isomorphic to $(\mathrm{O}(k) \times \mathrm{O}(n-k)) \cap \mathrm{SO}(n)$. Indeed, any orthogonal transformation taking Π^k to itself is the direct sum of an orthogonal transformation of Π^k and an orthogonal transformation of $(\Pi^k)^\perp$.

Choose an orientation at a point $\Pi^k \in G(n,k)$ and try to extend it to the whole manifold $G(n,k)$ by means of the action of $\mathrm{SO}(n)$. This can be done if the action of the entire stationary subgroup does not change the orientation at the point Π^k.

Take $U_1 \in \mathrm{O}(k)$ and $U_2 \in \mathrm{O}(n-k)$. The pair of matrices (U_1, U_2) acts on the matrix (I_k, X), where I_k is the identity matrix of order k and X is a $k \times (n-k)$ matrix, as follows: $(I_k, X) \mapsto (U_1, XU_2) \sim (I_k, U_1^{-1}XU_2)$. This determines a map $X \mapsto U_1^{-1}XU_2$ on the space of $k \times (n-k)$ matrices X. Take elements of X as coordinates. In these coordinates the map has matrix $U_2^T \otimes (U_1^{-1})^T = A$, with determinant $(\det U_2)^k (\det U_1^{-1})^{n-k}$ (see [100]).

Consider the case in which the matrix $U_1 \oplus U_2$ belongs to $\mathrm{SO}(n)$, i.e., has determinant 1. This means that $\det U_1 = \det U_2 = \pm 1$. If $\det U_1 = \det U_2 = 1$, then $\det A = 1$. If $\det U_1 = \det U_2 = -1$, then $\det A = (-1)^k(-1)^{n-k} = (-1)^n$. Therefore, if n is even, then the stationary subgroup preserves orientation, and if n is odd, then some elements of the stationary subgroup change orientation. Now, for odd n, we can easily construct a loop such that the orientation changes when transferred along it. Namely, in the connected group $\mathrm{SO}(n)$, take a path $\gamma(t)$ from I_n to $U_1 \oplus U_2$, where $U_1 \in \mathrm{O}(k)$ and $U_2 \in \mathrm{O}(n-k)$ are matrices with determinant -1. The loop $\gamma(t)\Pi^k$ in $G(n,k)$ determined by this path has the required property.

65. To each pair of vectors $(v_{11}, v_{12}, v_{13}, v_{14})$ and $(v_{21}, v_{22}, v_{23}, v_{24})$ we assign the six numbers $x_{ij} = \begin{vmatrix} v_{1i} & v_{1j} \\ v_{2i} & v_{2j} \end{vmatrix}$, where $i < j$. These numbers are the Plücker coordinates of the plane spanned by the given vectors. They are connected by the single Plücker relation

$$x_{12}x_{34} - x_{23}x_{14} + x_{13}x_{24} = 0. \tag{1}$$

Thus, it suffices to prove that the intersection of the hypersurface (1) with the 5-sphere $\sum x_{ij}^2 = 1$ is diffeomorphic to $S^2 \times S^2$.

Let us introduce the new coordinates $y_1, \dots, y_6$ defined by $2x_{12} = y_1 + y_4$, $2x_{34} = y_1 - y_4$, $2x_{23} = y_2 + y_5$, $2x_{14} = y_2 - y_5$, $2x_{13} = y_3 + y_6$, and $2x_{24} = y_3 - y_6$. In these coordinates, equation (1) takes the form $y_1^2 + y_2^2 + y_3^2 = y_4^2 + y_5^2 + y_6^2$, and the equation of the 5-sphere is $\sum y_k^2 = 2$. The obtained system of equations is equivalent to $y_1^2 + y_2^2 + y_3^2 = 1$, $y_4^2 + y_5^2 + y_6^2 = 1$. This system determines $S^2 \times S^2$.

66. Let $\Pi^2 \in G_+(n,2)$ be an oriented plane in $\mathbb{R}^n$. In this plane, take vectors v_1 and v_2 such that $|v_1| = |v_2|$, $v_1 \perp v_2$, and the basis v_1, v_2 has positive orientation. If the plane Π^2 is identified with $\mathbb{C}$, then the pair v_1, v_2 is determined up to multiplication by a nonzero complex number.

To each pair of vectors v_1, v_2 we assign the vector $v_1 + iv_2$ in $\mathbb{C}^n$, and to this vector we assign the corresponding point in $\mathbb{C}P^{n-1}$. Thereby, we define a one-to-one map from $G_+(n,2)$ onto a subset of $\mathbb{C}P^{n-1}$. Let us show that this set is the quadric $z_1^2 + \cdots + z_n^2 = 0$.

Suppose $v_1 = (x_1, \ldots, x_n)$ and $v_2 = (y_1, \ldots, y_n)$. Then $\sum (x_k + iy_k)^2 = \sum x_k^2 - \sum y_k^2 + 2i \sum x_k y_k = 0$, i.e., the point $v_1 + iv_2$ belongs to the quadric under consideration. Conversely, if $z_k = x_k + iy_k$ and $\sum z_k^2 = 0$, then the vectors $v_1 = (x_1, \ldots, x_n)$ and $v_2 = (y_1, \ldots, y_n)$ have equal lengths and are orthogonal.

Under complex conjugation, the basis v_1, v_2 is replaced by $v_1, -v_2$; the plane remains invariant, but its orientation changes.

67. Each point of TS^n is determined by two vectors $x, y \in \mathbb{R}^{n+1}$ for which $(x,x) = 1$ and $(x,y) = 0$. To this pair of vectors we assign the point $x\left(\sqrt{1 + \|y\|^2}\right) + iy \in \mathbb{C}^{n+1}$. Let $z_k = x_k\left(\sqrt{1 + \|y\|^2}\right) + iy_k$. Then

$$\begin{aligned}\sum \left(x_k(\sqrt{1 + \|y\|^2}) + iy_k\right)^2 \\ = (1 + \|y\|^2) \sum x_k^2 + 2i\sqrt{1 + \|y\|^2} \sum x_k y_k - \sum y_k^2 = 1\end{aligned}$$

because $\sum x_k^2 = 1$ and $\sum x_k y_k = 0$.

Conversely, take a point $u + iv \in \mathbb{C}^{n+1}$ for which $\sum (u_k + iv_k)^2 = 1$, i.e., $\|u\|^2 - \|v\|^2 = 1$ and $(u,v) = 0$. To the point $u + iv$ we assign the pair of vectors $x = \frac{u}{\sqrt{1+\|v\|^2}}$ and $y = v$. We have $\|x\|^2 = \frac{\|u\|^2}{1+\|v\|^2} = 1$ and $(x,y) = 0$.

68. Let $v_0 = (x_1, x_2, x_3, x_4) \in S^3$. Consider $v_1 = (-x_2, x_1, -x_4, x_3)$, $v_2 = (-x_3, x_4, x_1, -x_2)$, and $v_3 = (-x_4, -x_3, x_2, x_1)$. We have $(v_i, v_j) = 0$ for $i \neq j$. Therefore, v_1, v_2, and v_3 are pairwise orthogonal unit vectors tangent to S^3 at v_0.

The three linearly independent vector fields on S^{4n+3} are constructed similarly: we divide the coordinates into $n+1$ quadruples and perform the same operations on each quadruple.

69. (a) Let $v(x)$ be a vector field without singular points on S^{2n+1}. We can assume that $\|v(x)\| = 1$. Consider $H(t,x) = (\cos \pi t)x + (\sin \pi t)v(x)$. We have $\|H(t,x)\| = 1$, i.e., $H(t,x) \in S^{2n+1}$. Moreover, $H(0,x) = f(x)$ and $H(1,x) = g(x)$.

(b) Suppose that $x_{2k+1}(t) = x_{2k+1}\cos\pi t + x_{2k+2}\sin\pi t$ and $x_{2k+2}(t) = x_{2k+1}\sin\pi t + x_{2k+2}\cos\pi t$. Let $H(t,x) = \big(-x_0, x_1(t), \dots, x_{2n}(t)\big)$. Then $H(0,x) = g(x)$ and $H(1,x) = f(x)$.

70. (a) Consider the projection of the vector $f(x)$ to the tangent space at a point $x \in S^{2n}$. If $f(x) \neq \pm x$ for all x, then this projection is a vector field without singular points on S^{2n}; such a field does not exist.

(b) Every point $x \in \mathbb{R}P^{2n}$ can be regarded as a pair $\pm x \in S^{2n}$. Assigning the pair $\pm f(x)$ to each of the points $\pm x$, we obtain either two maps $\tilde f_{1,2}\colon S^{2n} \to S^{2n}$ (then $\tilde f_2 = -\tilde f_1$) or one map $\tilde f\colon S^{2n} \to \tilde S^{2n}$, where the space $\tilde S^{2n}$ doubly covers S^{2n}. The latter is impossible because $\pi_1(S^{2n}) = 0$.

If the map f has no fixed points, then the map $\tilde f_1$ has the property that $\tilde f_1(x) \neq \pm x$ for all $x \in S^{2n}$. According to (a), there is no such map.

71 (see [97]). We can assume that K is identified with $\mathbb{R}^n$ as a linear space. First, let us prove that the number n is even. Consider a path $\gamma(t)$ joining the points e and $-e$ and not passing through 0 in $\mathbb{R}^n$. Corresponding to each point $\gamma(t)$ of this path is a nondegenerate linear transformation $A_{\gamma(t)}\colon x \mapsto \mu\big(x, \gamma(t)\big)$. To the points e and $-e$ correspond the transformations I_n and $-I_n$ with determinants $\det I_n = 1$ and $\det(-I_n) = (-1)^n$. Suppose that n is odd. Then $\det(-I_n) = -1$. On the other hand, $\det(A_{\gamma(t)}) \neq 0$ for all t. Therefore, if $\det(A_{\gamma(t)}) > 0$ at the starting point of the path, then $\det(A_{\gamma(t)}) > 0$ at the end point. We have obtained a contradiction.

To every vector $v \in \mathbb{R}^n \setminus \{0\}$ we can assign a vector $f(v) \in \mathbb{R}^n \setminus \{0\}$ for which $\mu\big(v, f(v)\big) = e$ and, moreover, $v = \alpha e$ if and only if $f(v) = \beta e$. Thus, f is a self-homeomorphism of $\mathbb{R}^n \setminus \{te\}$, where $t \in \mathbb{R}$. It takes each ray tv, $t > 0$, to the ray $tf(v)$, $t > 0$, because $f(tv) = t^{-1}f(v)$ for $t \neq 0$. Therefore, we can consider the map $\tilde f\colon S^{n-1} \setminus \{\pm e\} \to S^{n-1} \setminus \{\pm e\}$ that takes each point v to the intersection point of the ray $tf(v)$, $t > 0$, with the sphere S^{n-1}. We use this map to define the map $g\colon S^{n-2} \to S^{n-2}$ that takes each point v to the intersection point of the sphere S^{n-2} formed by the unit vectors orthogonal to e with the great circle passing through $\pm e$ and $\tilde f(v)$.

The number $n-2$ is even; therefore, according to Problem 70, we have $g(v) = \pm v$ for some $v \in S^{n-2}$. This means that $f(v) = \alpha v + \beta e$, where $\alpha, \beta \in \mathbb{R}$, and $\alpha \neq 0$. By definition, $e = \mu\big(v, f(v)\big) = \mu(v, \alpha v + \beta e) = \alpha\mu(v,v) + \beta v$, whence $\mu(v,v) = \alpha^{-1}e - \alpha^{-1}\beta v$. Thus, the subspace spanned by the vectors e and v is a subalgebra. This subalgebra is associative and commutative, and it has a two-sided identity element and no divisors of zero; thus, it is a field. But any field which has dimension 2 as a real space is isomorphic to $\mathbb{C}$.

72. In a small neighborhood of any point of the sphere $S^{n-1} = \{x \in \mathbb{R}^n\colon \|x\| = 1/2\}$, the map f is (up to linear terms) a symmetry about the hyperplane tangent to S^{n-1} at the given point.

73. The map f is homotopic to a constant map because $\pi_2(M^2) = 0$. This implies $\deg f = 0$.

74. Suppose that x_0 is a regular value of the map fg, $f^{-1}(x_0) = \{a_1, \dots, a_k\}$, $g^{-1}(a_i) = \{b_{i1}, \dots, b_{il(i)}\}$, $\varepsilon_i = \operatorname{sign} J_f(a_i)$, $\varepsilon_{ij} = \operatorname{sign} J_g(b_{ij})$. Then $\varepsilon_1 + \dots + \varepsilon_k = \deg f$ and $\varepsilon_{i1} + \dots + \varepsilon_{il(i)} = \deg g$ for all $i = 1, \dots, k$. Therefore, $\deg(fg) = \sum \varepsilon_i \varepsilon_{ij} = \sum \varepsilon_i (\deg g) = (\deg f)(\deg g)$.

75. Let $P(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_0$, where $a_n \neq 0$. The map $z \mapsto P(z)$ extends to the map $\widehat{P}\colon \mathbb{C}P^1 \to \mathbb{C}P^1$ defined by

$$(z : w) \mapsto (a_n z^n + a_{n-1} z^{n-1} w + \dots + a_0 w^n : w^n).$$

Let $(u_0 : 1)$ be a regular value of $\widehat{P}$. The degree of $\widehat{P}$ is calculated as follows. First, we take the preimages of the point $(u_0 : 1)$, i.e., the points $z_0, \dots, z_k$ such that $P(z_j) = u_0$. Regularity implies $P'(z_j) \neq 0$. Therefore, the polynomial $P(z) - u_0$ has no multiple roots, i.e., $k = n$. Next, we determine the sign of the Jacobian determinant for the map $z \mapsto P(z)$ at each point z_j. If $P'(z_j) = a + bi$, then the Jacobian matrix of this map at z_j is $\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right)$. Its determinant equals $a^2 + b^2 > 0$; therefore, the Jacobian determinant is positive at all points $z_1, \dots, z_n$. Hence the degree of the map is equal to n.

76. Suppose that $R(z) = P(t)/Q(t)$, where $\deg P = m$ and $\deg Q = n$. Then the required smooth extension $\mathbb{C}P^1 \to \mathbb{C}P^1$ is defined by

$$(z : w) \mapsto \big(w^{m+n} P(z/w) : w^{m+n} Q(z/w)\big).$$

As in Problem 75, the Jacobian determinant is positive at all the preimages of regular values; thus, we only have to determine the number of preimages. The generic equation $P(z) = cQ(z)$ has $\max\{m, n\}$ roots.

77. We regard $\Sigma S^n \approx S^{n+1}$ as a smooth manifold. Take the equatorial sphere $S^n = S^n \times \{1/2\}$ in S^{n+1} and choose a regular value x_0 for f on it; this value is also regular for the map Σf. Moreover, if J is the Jacobian matrix of f at the point $y_0 \in f^{-1}(x_0)$, then $\left(\begin{smallmatrix} J & 0 \\ 0 & 1 \end{smallmatrix}\right)$ is the Jacobian matrix of Σf.

78. Let us show that $\deg f = k_1 \cdots k_n$. If one of the numbers $k_1, \dots, k_n$ is zero, then this assertion is obvious. In what follows, we assume that $k_1 \cdots k_n \neq 0$. The preimage of any point $(r_1 e^{i\psi_1}, \dots, r_n e^{i\psi_n})$ consists of the $|k_1 \cdots k_n|$ points

$$\big(r_1 e^{i(\psi_1 + 2\pi l_1)/k_1}, \dots, r_n e^{i(\psi_n + 2\pi l_n)/k_n}\big),$$

where $0 \le l_i \le |k_1| - 1$. Clearly, the Jacobian determinant of f has the same sign as $k_1 \cdots k_n$.

79. It follows from $f(-A) = f(A)$ that the degree of f is even. Therefore, the map f is not homotopic to the identity map.

80. We can assume that C is the unit circle. Instead of a vector field v on C, it is more convenient to consider the vector field w obtained by rotating the vector v at each point $(\cos\varphi, \sin\varphi)$ through $-\varphi$. For example, if v is tangent to C, then w has constant direction (at an angle of $\pm 90°$). Clearly, $\operatorname{ind} v = \operatorname{ind} w + 1$; thus, it suffices to prove that $2 \operatorname{ind} w = i - e$.

At the points of tangency of the circle C with the integral trajectories, the vector w is normal to the Ox-axis. It is easy to show that at the points of external tangency, w rotates in the direction opposite to that in which the circle is traversed. This means that the Jacobian determinant of the map that takes each point of C to the vector w at this point is negative at the points of external tangency. At the points of internal tangency, the Jacobian is positive. The points of external and internal tangency are the preimages of the two points corresponding to the two directions normal to the Ox-axis. Therefore, one point has $(i-e)/2$ preimages (with signs taken into account).

83. The case $n = 1$ is obvious; thus, we assume that $n \geq 2$. Consider a vector field v with nondegenerate singular points on M^n. The singular points can be divided into pairs consisting of points whose indices have opposite signs. Let us show how to decrease the number of singular points (if they exist) by two. Take two singular points x_+ and x_- with indices 1 and -1. Let γ be a path between them not passing through other singular points, and let γ_ε be the ε-neighborhood of γ. If ε is sufficiently small, then $\gamma_\varepsilon \approx D^n$ and γ_ε contains no singular points except x_+ and x_-. We can assume that the set γ_ε is covered by one chart, and all the vectors $v(x)$, where $x \in \partial\gamma_\varepsilon$, have unit length. Consider the map $\partial\gamma_\varepsilon \to S^{n-1}$ defined by $x \mapsto v(x)$. The degree of this map is equal to the sum of the indices of the singular points x_+ and x_-, i.e., vanishes. Therefore, the map $\partial\gamma_\varepsilon \to S^{n-1}$ is homotopic to a constant map. The homotopy can be regarded as a vector field w on γ_ε which consists of unit vectors and coincides with the initial vector field v on $\partial\gamma_\varepsilon$. Consider the vector field that coincides with v outside γ_ε and with w on γ_ε. This vector field has two fewer singular points than v. In this way, all singular points can be killed.

84. Clearly, if M^n admits a vector field without singular points, then it admits also a line element field. Suppose that there is a line element field on M^n. Let us endow M^n with a Riemannian metric and take both unit vectors in each of the 1-dimensional subspaces determined by the field. The set of all such vectors is a closed manifold $\widetilde{M}^n$, which doubly covers M^n and is endowed with a vector field without singular points. (The manifold $\widetilde{M}^n$ is either connected or has two connected components diffeomorphic to M^n.) Let χ be the sum of indices of the singular points of the vector field on M^n. According to Theorem 5.29 on p. 230, the indices of singular points of the vector field on $\widetilde{M}^n$ sum to 2χ. But $\widetilde{M}^n$ admits a vector field without singular points; therefore, $2\chi = 0$, whence $\chi = 0$. Now, the existence of a vector field without singular points on M^n follows from Problem 83.

85. In Example 4.3 on p. 171, it was shown that the Hopf fibration $p\colon S^3 \to S^2$ generates the group $\pi_3(S^2) \cong \mathbb{Z}$. Let us describe the corresponding framed manifold in $\Omega^1_{\mathrm{fr}}(3)$. Take two near points on the sphere S^2. Their preimages are two circles forming a Hopf link. Therefore, the corresponding framed manifold is the circle S^1 with framing defined as follows. Consider a Hopf link in $\mathbb{R}^3$ such that one of its components is the circle S^1 under consideration (we assume that it is standardly embedded in $\mathbb{R}^3$), the second component lies on the boundary of the ε-neighborhood of S^1, and each disk of radius ε orthogonal to S^1 intersects the second component in exactly one point. The end of the vector e_1 of the framing moves on the second component of the Hopf link ($\|e_1\| = \varepsilon$); the vector e_2 lies in the plane normal to S^1 and is orthogonal to e_1.

86. (a) The contractibility of the cone CY implies $X/Y = (X \cup CY)/CY \sim X \cup CY$ for any subcomplex $Y \subset X$. It is easy to show that the proof of Lemma 5.2 remains valid for Y instead of S^{i-1} and CY instead of D^i. Thus, if $f, g\colon Y \to X$ are homotopic maps, then $X \cup_f CY \sim X \cup_g CY$. By assumption, the embedding $f\colon Y \to X$ is homotopic to the constant map $g\colon Y \to x_0 \in X$. But $X \cup_f CY = X \cup CY$ and $X \cup_g CY = X \vee \Sigma Y$.

(b) The equatorial sphere S^k is contractible in S^{k+1}. Therefore, for $n > m$, the sphere S^m canonically embedded in S^n is contractible in S^n. It remains to apply (a).

88. It is easy to see that the map $i_*\colon \pi_1(A) \to \pi_1(X)$ is given by $a \mapsto 2a$, where a is the generator of the groups $\pi_1(A)$ and $\pi_1(X)$. Therefore, r_*i_* cannot be the identity map.

89. (a) It is seen from the standard representation of the torus T^2 as a square with sides glued together that the map $i_*\colon \pi_1(A) \to \pi_1(X)$ is defined by $a \mapsto \alpha\beta\alpha^{-1}\beta^{-1}$, where a is the generator of $\pi_1(A)$ and α and β are generators of the free group $\pi_1(X)$. Thus, for the commutator subgroups of the fundamental groups, we obtain the zero map. Therefore, the map r_*i_* cannot be the identity.

(b) In this case, the map $i_*\colon \pi_1(A) \to \pi_1(X)$ is defined by

$$a \mapsto \alpha_1\beta_1\alpha_1^{-1}\beta_1^{-1} \cdots \alpha_g\beta_g\alpha_g^{-1}\beta_g^{-1},$$

and the map of commutator subgroups is again zero.

90. The map of commutator subgroups that is induced by $i_*\colon \pi_1(A) \to \pi_1(X)$ has the form $a \mapsto 2\alpha_1 + 2\alpha_2 + \cdots + 2\alpha_g$. Therefore, r_*i_* cannot be the identity.

91. The tangent vector at $x \in S^{n-1} \subset \mathbb{R}^n$ is orthogonal to x. Therefore, each point of the manifold M^3_0 is an ordered pair of orthogonal unit vectors $e_1, e_2 \in \mathbb{R}^3$. This pair admits a unique extension to a positively oriented orthonormal basis e_1, e_2, e_3. Therefore, $M^3_0 \approx \mathrm{SO}(3)$.

The homeomorphism $SO(3) \approx \mathbb{R}P^3$ is established as follows. Any transformation from $SO(3)$ has an eigenvector; hence it is a rotation through an angle φ about an axis l passing through the origin. To each vector $e_3 \in \mathbb{R}^3$ of length φ ($0 < \varphi \leq \pi$) we assign the rotation through the angle φ about the axis e_3 in the direction for which the basis e_1, e_2, e_3, where e_1 is a vector orthogonal to e_3 and e_2 is the image of e_1 under the rotation, is positively oriented. The zero vector is assigned the identity transformation. We have established a correspondence between the points of a disk D^3 of radius π and the transformations from $SO(3)$. Any two antipodal points of the disk correspond to the same transformation.

92. The space $\mathbb{R}^3 \setminus S^1$ is homotopy equivalent to $S^2 \cup I$, where I is a diameter of the sphere S^2. Let I_1 be an arc on the sphere joining the endpoints of I. Then $S^2 \cup I \sim (S^2 \cup I)/I_1 \sim S^2 \vee S^1$.

94. The space $\mathbb{R}^3 \setminus L$ is homotopy equivalent to the wedge of n copies of the space $D^3 \setminus S^1$, where $S^1 \subset D^3$ is the standardly embedded circle (trivial knot). According to Problem 92, we have $D^3 \setminus S^1 \sim S^2 \vee S^1$.

95. In discussing the properties of the Hopf fibration, we have represented the sphere S^3 as the union of the two solid tori $T_1 = D_1^2 \times S^1$ and $T_2 = D_2^2 \times S^1$. The circles $\{0\} \times S^1$ in these solid tori form the link under consideration. Since $\mathbb{R}^3 = S^3 \setminus *$, it follows that the space $\mathbb{R}^3 \setminus L$ is obtained from the solid torus T_1 by removing the circle $\{0\} \times S^1$ and one point. Such a space is homotopy equivalent to $T^2 \vee S^2$.

Bibliography

[1] M. Adachi, *Embeddings and immersions.* Amer. Math. Soc., Providence, RI, 1993.

[2] A. A. Albert, *Non-associative algebras.* Ann. Math. **43** (1942), 685–707.

[3] J. C. Alexander, *Morse functions on Grassmannians.* Illinois J. Math. **15** (1971), 672–681.

[4] P. S. Alexandroff, *Über stetige Abbildungen kompakter Räume.* Math. Ann. **96** (1927), 555–571.

[5] K. Appel and W. Haken, *Every planar map is four colorable.* I. *Discharging.* Illinois J. Math. **21** (1977), 429–490.

[6] ———, *Every planar map is four colorable.* Amer. Math. Soc., Providence, RI, 1989.

[7] K. Appel, W. Haken, and J. Koch, *Every planar map is four colorable. II. Reducibility.* Illinois J. Math. **21** (1977), 491–567.

[8] D. Archdeacon and J. Širáň, *Characterizing planarity using theta graphs.* J. Graph Theory **27** (1998), 17–20.

[9] V. I. Arnold, *Lectures on partial differential equations.* Fazis, Moscow, 1997; English transl., Springer-Verlag, Berlin, 2004.

[10] M. L. Balinski, *On the graph structure of convex polyhedra in n-space.* Pacific J. Math. **11** (1961), 431–434.

[11] T. F. Banchoff, *Global geometry of polygons.* I. *The theorem of Fabricius Bjerre,* Proc. Amer. Math. Soc. **45** (1974), 237–241.

[12] E. G. Bajmóczy and I. Bárány, *On a common generalization of Borsuk's theorem and Radon's theorem.* Acta Math. Hungar. **34** (1979), 347 350.

[13] I. Bárány and L. Lovász, *Borsuk's theorem and the number of facets of centrally symmetric polytopes.* Acta Math. Hungar. **40** (1982), 323–329.

[14] D. W. Barnette and B. Grünbaum, *On Steinitz's theorem concerning convex 3-polytopes and on some properties of planar graphs.* The Many Facets of Graph Theory, Springer-Verlag, Berlin, 1969, pp. 27–40.

[15] P. Bohl, *Über die Bewegung eines mechanischen Systems in der Nähe einer Gleichgewichtslage.* J. Reine Angew. Math. **127** (1904), 179–276.

[16] J. A. Bondy and U. S. R. Murty, *Graph theory with applications.* Macmillan, London, 1976.

[17] K. Borsuk, *Drei Sätze über die n-dimensionale euklidische Sphäre.* Fund. Math. **20** (1933), 177 190.

[18] R. Bott, *Two new combinatorial invariants for polyhedra.* Portugaliae Math. **11** (1952), 35–40.

[19] R. Bott and L. W. Tu, *Differential forms in algebraic topology.* Springer-Verlag, New York, 1989.

[20] N. Bourbaki, *Éléments de mathématique.* Fasc. II. Livre III: *Topologie générale.* Hermann, Paris, 1965.

[21] G. E. Bredon and J. W. Wood, *Non-orientable surfaces in orientable 3-manifolds.* Invent. Math. **7** (1969), 83–110.

[22] J. R. Breitenbach, *A criterion for the planarity of a graph.* J. Graph Theory **10** (1986), 529–532.

[23] L. E. J. Brouwer, *Über Abbildung von Mannigfaltigkeiten.* Math. Ann. **71** (1912), 97 115.

[24] ———, *Über den natürlichen Dimensionsbegriff.* J. Reine Angew. Math. **142** (1913), 146–152.

[25] A. B. Brown and S. S. Cairns, *Strengthening of Sperner's lemma applied to homology theory.* Proc. Nat. Acad. Sci. U.S.A. **47** (1961), 113–114.

[26] S. S. Cairns, *A simple triangulation method for smooth manifolds,* Bull. Amer. Math. Soc. **67** (1961), 389–390.

[27] D. Cheniot, *Le théorème de van Kampen sur le groupe fondamental du complémentaire d'une courbe algèbrique projective plane.* Fonctions de plusieurs variables complexes (Sem. François Norguet, a la mémoire d'André Martineau), Lecture Notes in Math., vol. 409, Springer-Verlag, Berlin, 1974, pp. 394–417.

[28] D. I. A. Cohen, *On the Sperner lemma.* J. Combin. Theory **2** (1967), 585–587.

[29] J. H. Conway and C. McA. Gordon, *Knots and links in spatial graphs.* J. Graph Theory **7** (1983), 445–453.

[30] J. H. C. Creighton, *An elementary proof of the classification of surfaces in the projective 3-space.* Proc. Amer. Math. Soc. **72** (1978), 191 192.

[31] R. H. Crowell, *On the van Kampen theorem.* Pacific J. Math. **9** (1959) 43–50.

[32] J. Dieudonné, *Une généralisation des espace compact.* J. Math. Pures Appl. **23** (1944), 65–76.

[33] R. Engelking, *Dimension theory.* North-Holland, Amsterdam–Oxford New York, 1978.

[34] Fr. Fabricius-Bjerre, *On the double tangents of plane closed curves.* Math. Scand. **11** (1962), 113–116.

[35] ———, *A proof of a relation between the numbers of singularities of a closed polygon.* J. Geom. **13** (1979), 126–132.

[36] I. Fáry *On straight line representation of planar graph.* Acta Sci. Math. (Szeged) **11** (1948), 229–233.

[37] A. Fathi, *Partitions of unity for countable covers.* Amer. Math. Monthly. **104** (1997), 720–723.

[38] A. Flores, *Über die Existenz n-dimensionaler Komplexe, die nicht in den R_{2n} topologisch einbettbar sind.* Ergeb. Math. Kolloqu. **5** (1933), 17–24.

[39] ———, *Über n-dimensionale Komplexe, die im R_{2n+1} absolut selbstverschlungen sind.* Ergeb. Math. Kolloqu. **6** (1935), 4–6.

[40] A. T. Fomenko and D. B. Fuks, *A Course in homotopic topology.* Nauka, Moscow, 1989. (Russian)

[41] H. Freudenthal, *Die Fundamentalgruppe der Mannigfaltigkeit der Tangentialrichtungen einer geschlossenen Fläche.* Fund. Math. **50** (1962), 537 538.

[42] R. Fritsch and R. A. Piccini, *Cellular structures in topology.* Cambridge Univ. Press, Cambridge, 1990.

[43] K. Gęba and A. Granas, *A proof of the Borsuk antipodal theorem.* J. Math. Anal. Appl. **96** (1983), 203–208.

[44] A. Gramain, *Le théorème de van Kampen.* Cahiers Top. Geom. Diff. Categoriques **33** (1992), 237 250.

[45] J. L. Gross and W. Tucker Thomas, *Topological graph theory.* Wiley, New York, 1987.

[46] B. Grünbaum *Imbeddings of simplicial complexes.* Comment. Math. Helv. **44** (1969), 502 513.

[47] B. Halpern, *An inequality for double tangents.* Proc. Amer. Math. Soc. **76** (1979), 133–139.

[48] Th. Hangan, *A Morse function on Grassmann manifolds.* J. Diff. Geom. **2** (1968), 363–367.

[49] G. Harris and C. Martin, *The roots of a polynomial vary continuously as a function of the coefficients.* Proc. Amer. Math. Soc. **100** (1987), 390–392.

[50] A. Hatcher, *Algebraic topology.* Cambridge Univ. Press, Cambridge, 2002.

[51] F. Hausdorff, *Set theory.* Chelsea, New York, 1957.

[52] P. J. Heawood, *Map colour theorem.* Quart. J. Math. **24** (1980), 332–338.

[53] M. W. Hirsch, *A proof of the nonretractibility of a cell onto its boundary.* Proc. Amer. Math. Soc. **4** (1963), 364 365.

[54] ———, *Differential topology.* Springer-Werlag, New York Heidelberg, 1976.

[55] Chung-Wu Ho, *A note on proper maps.* Proc. Amer. Math. Soc. **51** (1975), 237 241.

[56] ———, *When are immersions diffeomorphisms?* Canad. Math. Bull. **24** (1981), 491–492.

[57] E. Hopf, *Über die Drehung der Tangenten und sehnen ebener Kurven.* Comp. Math. **2** (1935), 50–62.

[58] H. Hopf, *Abbildungsklassen n-dimensionaler Mannigfaltigkeiten.* Math. Ann. **96** (1927), 209–224.

[59] S.-T. Hu, *Elements of general topology.* Holden-Day, San Francisco, 1964.

[60] W. Hurewicz and H. Wallman, *Dimension theory.* Princeton Univ. Press, Princeton, NJ, 1941.

[61] K. Jänich, *Topology.* Springer-Verlag, New York, 1980.

[62] C. Jordan, *Cours d'analyse de l'École Polytechnique,* vol. 3 (Gauthier-Villars, Paris, 1887), pp. 587–594.

[63] S. Kakutani *A generalization of Brouwer's fixed point theorem.* Duke Math. J. **8** (1941), 457–459.

[64] ———, *A proof that there exists a circumscribing cube around any bounded closed convex set in $\mathbb{R}^n$.* Ann. Math. **43** (1942), 739–741.

[65] E. van Kampen, *Komplexe in euklidischen Räumen.* Abh. Math. Sem. Univ. Hamburg **9** (1932), 72 78, 152 153.

[66] ———, *On the fundamental group of an algebraic curve.* Amer. J. Math. **55** (1933), 255–260.

[67] ———, *On the connection between the fundamental groups of some related spaces.* Amer. J. Math. **55** (1933), 261 267.

[68] B. Kuratowski, C. Knaster, and C. Mazurkiewicz, *Ein Beweis des Fixpunktsatzes für n-dimensionale Simplexe.* Fund. Math. **14** (1929), 132 137.

[69] G. W. Knutson, *A note on the universal covering space of a surface.* Amer. Math. Monthly **78** (1971), 505–509.

[70] R. Koch, *Matrix invariants.* Amer. Math. Monthly **91** (1984), 573–575.

[71] D. König, *Theorie der endlichen und unendlichen Graphen.* Leipzig, 1936.

[72] K. Kuratowski, *Sur le problème des courbes gauches en topologie.* Fund. Math. **15** (1930), 271 283.

[73] S.-N. Lee, *A combinatorial Lefschetz fixed-point theorem.* J. Combin. Theory, Ser. A. **61** (1992), 123–129.

[74] F. T. Leighton, *Finite common coverings of graphs.* J. Combin. Theory. Ser. B **33** (1982), 231–238.

[75] L. Lovász and A. Schrijver, *A Borsuk theorem for antipodal links and spectral characterization of linklessly embeddable graphs.* Proc. Amer. Math. Soc. **126** (1998), 1275–1285.

[76] L. A. Lyusternik and L. G. Shnirel'man, *Topological methods in variational problems.* Gosizdat, Moscow, 1930. (Russian)

[77] S. Mac Lane, *A combinatorial condition for planar graphs.* Fund. Math. **28** (1937), 22–32.

[78] H. Maehara, *Why is P^2 not embeddable in $\mathbb{R}^3$?* Amer. Math. Monthly **100** (1993), 862–864.

[79] R. Maehara, *The Jordan curve theorem via the Brouwer fixed point theorem.* Amer. Math. Monthly **91** (1984), 641–643.

[80] Yu. Makarychev, *A short proof of Kuratowski's graph planarity criterion.* J. Graph Theory **25** (1997), 129–131.

[81] M. L. Marx, *The Gauss realizability problem.* Proc. Amer. Math. Soc. **22** (1969), 610–613.

[82] M. Mather, *Paracompactness and partitions of unity.* Thesis, Cambridge Univ., 1965.

[83] J. Mayer *Le problème des régions voisines sur les surfaces closes orientables.* J. Comb. Theory **5** (1969), 177–195.

[84] M. D. Meyerson and A. H. Wright, *A new and constructive proof of the Borsuk–Ulam theorem.* Proc. Amer. Math. Soc. **73** (1979), 134–136.

[85] J. Milnor, *On manifolds homeomorphic to the 7-sphere.* Ann. of Math. (2) **64** (1959), 399–405.

[86] ———, *Morse theory.* Princeton Univ. Press, Princeton, NJ, 1963.

[87] J. Milnor and A. Wallace, *Topology from the differentiable viewpoint. Differential topology.* Princeton Univ. Press, Princeton, NJ, 1968.

[88] J. Milnor *Analytic proof of the "hairy ball theorem" and the Brouwer fixed point theorem.* Amer. Math. Monthly. **85** (1978), 521–524.

[89] E. M. Moise, *Geometric topology in dimension 2 and 3.* Springer-Verlag, New York, 1977.

[90] H. R. Morton, *Symmetric products of the circle.* Proc. Cambridge Phil. Soc. **63** (1967), 349–352.

[91] G. L. Naber, *Topological methods in Euclidean spaces.* Cambridge Univ. Press, Cambridge, 1980.

[92] C. St. J. A. Nash-Williams and W. T. Tutte, *More proofs of Menger's theorem.* J. Graph Theory **1** (1977), 13–17.

[93] S. Negami, *Polynomial invariants of graphs.* Trans. Amer. Math. Soc. **299** (1987), 601–622.

[94] M. Oka, *Some plane curves whose complements have non-abelian fundamental groups.* Math. Ann. **218** (1975), 55–65.

[95] T. Ozawa, *On Halpern's conjecture for closed plane curves.* Proc. Amer. Math. Soc. **92** (1984), 554–560.

[96] T. D. Parsons, *On planar graphs.* Amer. Math. Monthly. **78** (1971), 176–178.

[97] J. Petro, *Real division algebras of dimension* > 1 *contains* **C**. Amer. Math. Monthly **94** (1987), 445–449.

[98] H. Poincaré, *Sur les courbes définies par les équations différentielles.* IV. J. Math. Pures Appl. **2** (1886), 151 217.

[99] M. M. Postnikov, *Lectures on algebraic topology: Homotopy theory of CW complexes.* Nauka, Moscow, 1985.

[100] V. V. Prasolov, *Problems and theorems in linear algebra.* Amer. Math. Soc., Providence, RI, 1994.

[101] ———, *Intuitive topology.* Amer. Math. Soc., Providence, RI, 1995.

[102] V. V. Prasolov and A. B. Sossinsky, *Knots, links, braids and 3-manifolds.* Amer. Math. Soc., Providence, RI, 1997.

[103] V. V. Prasolov and V. M. Tikhomirov, *Geometry.* Amer. Math. Soc., Providence, RI, 2001.

[104] V. V. Prasolov, *Polynomials.* Springer-Verlag, Berlin–Heidelberg–New York, 2004.

[105] M. Prüfer, *Complementary pivoting and the Hopf degree theorem.* J. Math. Anal. Appl. **84** (1981), 133–149.

[106] G. Ringel, *Das Geschlecht des vollständiger paaren Graphen.* Abh. Math. Semin. Univ. Hamburg. **28** (1965), 139–150.

[107] G. Ringel and J. W. T. Youngs, *Solution of the Heawood map-coloring problem.* Proc. Nat. Acad. Sci. USA **60** (1968), 438–445.

[108] ———, *Remarks on the Heawood conjecture.* Proof Techniques in Graph Theory. Academic Press, New York, 1969, pp. 133–138.

[109] H. Robbins. *Some complements to Brouwer's fixed point theorem.* Israel J. Math. **5** (1967), 225–226.

[110] N. Seymour, D. D. Robertson, P. D. Sanders, and R. Thomas, *The four-colour theorem.* J. Comb. Theory, Ser. B **70** (1997), 2–44.

[111] C. A. Rogers, *A less strange version of Milnor's proof of Brouwer's fixed-point theorem.* Amer. Math. Monthly **87** (1980), 525–527.

[112] V. A. Rokhlin and D. B. Fuks, *Beginner's course in topology: Geometric chapters.* Nauka, Moscow, 1977; English transl., Springer-Verlag, Berlin, 1984.

[113] J. J. Rotman, *An introduction to algebraic topology.* Springer-Verlag, New York, 1988.

[114] M. E. Rudin, *A new proof that metric spaces are paracompact.* Proc. Amer. Math. Soc. **20** (1969), 603.

[115] H. Sachs, *On a spatial analogue of Kuratowski's theorem on planar graphs—an open problem.* Graph theory (Lagów, 1981). Lecture Notes in Math., vol. 1018, Springer-Verlag, Berlin New York, pp. 231 240.

[116] H. Samelson, *Orientability of hypersurfaces in* $\mathbb{R}^n$. Proc. Amer. Math. Soc. **22** (1969), 301 302.

[117] A. Sard, *The measure of the critical points of differentiable maps.* Bull. Amer. Math. Soc. **48** (1942), 883–890.

[118] K. S. Sarkaria, *A generalized Kneser conjecture.* J. Comb. Theory. Ser. B **49** (1990), 236–240.

[119] ———, *A one-dimensional Whitney trick and Kuratowski's graph planarity criterion.* Israel J. Math. **73** (1991), 79–89.

[120] H. Seifert, *Konstruktion dreidimensionaler geschlossener Räume.* Ber. Sächs. Akad. Wiss. **83** (1931), 26–66.

[121] E. Sperner, *Neuer Beweis für die Invarianz der Dimensionzahl und des Gebietes.* Abh. Math. Semin. Hamburg. Univ. **6** (1928), 265–272.

[122] E. Steinitz, *Polyeder und Raumeinteilungen.* Enzyklopadie der mathemaischen Wissenschaften mit Einschluss ihrer Anwendungen, Band 3: Geometrie, Teil 3, Heft 12. Teubner, Leipzig, 1922, pp. 1–139.

[123] A. H. Stone, *Paracompactness and product spaces.* Bull. Amer. Math. Soc. **54** (1948), 977 982.

[124] C. Thomassen, *Kuratowski's theorem.* J. Graph Theory. **5** (1981), 225–241.

[125] ———, *The Jordan Schönflies theorem and the classification of surfaces.* Amer. Math. Monthly **99** (1992), 116–130.

[126] A. W. Tucker, *Some topological properties of disk and sphere.* Proc. First Canadian Math. Congress, Montreal, 1945. University of Toronto Press, Toronto, 1946, pp. 285–309.

[127] W. T. Tutte, *A contribution to the theory of chromatic polynomials.* Canad. J. Math. **6** (1954), 80–91.

[128] H. Tverberg, *A proof of the Jordan curve theorem.* Bull. London Math. Soc. **12** (1980), 34–38.

[129] P. S. Urysohn, *Über die Mächtigkeit der zusammenhängenden Mengen.* Math. Ann. **94** (1925), 262–295.

[130] C. L. Van, *Topological degree and the Sperner lemma.* J. Optimiz. Theory Appl. **37** (1982), 371 377.

[131] V. A. Vassiliev, *Introduction to topology.* Amer. Math. Soc., Providence, RI, 2001.

[132] O. Veblen, *Theory of plane curves in nonmetrical analysis situs.* Trans. Amer. Math. Soc. **6** (1905), 83–98.

[133] A. P. Veselov and I. A. Dynnikov, *Integrable gradient flows and Morse theory.* Algebra i Analiz. **8** (1996), no. 3, 78–103; English transl., St. Petersburg Math. J. **8** (1997), no. 3, 429–446.

[134] B. L. van der Waerden, *Algebra.* Ungar, New York, 1970.

[135] E. B. Vinberg, *A course in algebra.* Amer. Math. Soc., Providence, RI, 2003.

[136] K. Wagner, *Bemerkungen zum Vierfarbenproblem.* Jahresber. Deutsch. Math. Verein. **46** (1936), 26–32.

[137] Zh. Wang, *On Bott polynomials.* J. Knot Theory Ramifications **3** (1994), 537 546.

[138] G. N. Watson, *A problem in analysis situs.* Proc. London Math. Soc. **15** (1916), 227 242.

[139] A. Weil, *Sur le théorèmes de de Rham.* Comment. Math. Helv. **26** (1952), 119–145.

[140] B. Weiss, *A combinatorial proof of the Borsuk Ulam antipodal point theorem.* Israel J. Math. **66** (1989), 364–368.

[141] J. H. C. Whitehead, *Combinatorial homotopy.* I. Bull. Amer. Math. Soc. **55** (1949), 213–245.

[142] H. Whitney, *Nonseparable and planar graphs.* Trans. Amer. Math. Soc. **34** (1932), 339–362.

[143] ———, *The coloring of graphs.* Ann. Math. **33** (1932), 687 718.

[144] ———, *A set of topological invariants for graphs.* Amer. J. Math. **55** (1933), 231 235.

[145] ———, *Differentiable manifolds.* Ann. Math. **45** (1936), 645–680.

[146] ———, *On regular closed curves in the plane.* Comp. Math. **4** (1937), 276–284.

[147] W. T. Wu, *On critical sections of convex bodies.* Sci. Sinica **14** (1965), 1721 1728.

[148] O. Zariski, *On the problem of existence of algebraic functions of two variables possessing a given branch curve.* Amer. J. Math. **51** (1929), 305–328.

[149] ———, *Algebraic surfaces.* Springer-Verlag, Berlin, 1935.

[150] ———, *On the Poincaré group of rational plane curves.* Amer. J. Math. **58** (1936), 607–619.

Index

Titles in This Series

74 **V. V. Prasolov,** Elements of combinatorial and differential topology, 2006
73 **Louis Halle Rowen,** Graduate algebra: Commutative view, 2006
72 **R. J. Williams,** Introduction the the mathematics of finance, 2006
70 **Seán Dineen,** Probability theory in finance, 2005
69 **Sebastián Montiel and Antonio Ros,** Curves and surfaces, 2005
68 **Luis Caffarelli and Sandro Salsa,** A geometric approach to free boundary problems, 2005
67 **T.Y. Lam,** Introduction to quadratic forms over fields, 2004
66 **Yuli Eidelman, Vitali Milman, and Antonis Tsolomitis,** Functional analysis, An introduction, 2004
65 **S. Ramanan,** Global calculus, 2004
64 **A. A. Kirillov,** Lectures on the orbit method, 2004
63 **Steven Dale Cutkosky,** Resolution of singularities, 2004
62 **T. W. Körner,** A companion to analysis: A second first and first second course in analysis, 2004
61 **Thomas A. Ivey and J. M. Landsberg,** Cartan for beginners: Differential geometry via moving frames and exterior differential systems, 2003
60 **Alberto Candel and Lawrence Conlon,** Foliations II, 2003
59 **Steven H. Weintraub,** Representation theory of finite groups: algebra and arithmetic, 2003
58 **Cédric Villani,** Topics in optimal transportation, 2003
57 **Robert Plato,** Concise numerical mathematics, 2003
56 **E. B. Vinberg,** A course in algebra, 2003
55 **C. Herbert Clemens,** A scrapbook of complex curve theory, second edition, 2003
54 **Alexander Barvinok,** A course in convexity, 2002
53 **Henryk Iwaniec,** Spectral methods of automorphic forms, 2002
52 **Ilka Agricola and Thomas Friedrich,** Global analysis: Differential forms in analysis, geometry and physics, 2002
51 **Y. A. Abramovich and C. D. Aliprantis,** Problems in operator theory, 2002
50 **Y. A. Abramovich and C. D. Aliprantis,** An invitation to operator theory, 2002
49 **John R. Harper,** Secondary cohomology operations, 2002
48 **Y. Eliashberg and N. Mishachev,** Introduction to the h-principle, 2002
47 **A. Yu. Kitaev, A. H. Shen, and M. N. Vyalyi,** Classical and quantum computation, 2002
46 **Joseph L. Taylor,** Several complex variables with connections to algebraic geometry and Lie groups, 2002
45 **Inder K. Rana,** An introduction to measure and integration, second edition, 2002
44 **Jim Agler and John E. McCarthy,** Pick interpolation and Hilbert function spaces, 2002
43 **N. V. Krylov,** Introduction to the theory of random processes, 2002
42 **Jin Hong and Seok-Jin Kang,** Introduction to quantum groups and crystal bases, 2002
41 **Georgi V. Smirnov,** Introduction to the theory of differential inclusions, 2002
40 **Robert E. Greene and Steven G. Krantz,** Function theory of one complex variable, third edition, 2006
39 **Larry C. Grove,** Classical groups and geometric algebra, 2002
38 **Elton P. Hsu,** Stochastic analysis on manifolds, 2002
37 **Hershel M. Farkas and Irwin Kra,** Theta constants, Riemann surfaces and the modular group, 2001
36 **Martin Schechter,** Principles of functional analysis, second edition, 2002

TITLES IN THIS SERIES

35 **James F. Davis and Paul Kirk,** Lecture notes in algebraic topology, 2001
34 **Sigurdur Helgason,** Differential geometry, Lie groups, and symmetric spaces, 2001
33 **Dmitri Burago, Yuri Burago, and Sergei Ivanov,** A course in metric geometry, 2001
32 **Robert G. Bartle,** A modern theory of integration, 2001
31 **Ralf Korn and Elke Korn,** Option pricing and portfolio optimization: Modern methods of financial mathematics, 2001
30 **J. C. McConnell and J. C. Robson,** Noncommutative Noetherian rings, 2001
29 **Javier Duoandikoetxea,** Fourier analysis, 2001
28 **Liviu I. Nicolaescu,** Notes on Seiberg-Witten theory, 2000
27 **Thierry Aubin,** A course in differential geometry, 2001
26 **Rolf Berndt,** An introduction to symplectic geometry, 2001
25 **Thomas Friedrich,** Dirac operators in Riemannian geometry, 2000
24 **Helmut Koch,** Number theory: Algebraic numbers and functions, 2000
23 **Alberto Candel and Lawrence Conlon,** Foliations I, 2000
22 **Günter R. Krause and Thomas H. Lenagan,** Growth of algebras and Gelfand-Kirillov dimension, 2000
21 **John B. Conway,** A course in operator theory, 2000
20 **Robert E. Gompf and András I. Stipsicz,** 4-manifolds and Kirby calculus, 1999
19 **Lawrence C. Evans,** Partial differential equations, 1998
18 **Winfried Just and Martin Weese,** Discovering modern set theory. II: Set-theoretic tools for every mathematician, 1997
17 **Henryk Iwaniec,** Topics in classical automorphic forms, 1997
16 **Richard V. Kadison and John R. Ringrose,** Fundamentals of the theory of operator algebras. Volume II. Advanced theory, 1997
15 **Richard V. Kadison and John R. Ringrose,** Fundamentals of the theory of operator algebras. Volume I: Elementary theory, 1997
14 **Elliott H. Lieb and Michael Loss,** Analysis, 1997
13 **Paul C. Shields,** The ergodic theory of discrete sample paths, 1996
12 **N. V. Krylov,** Lectures on elliptic and parabolic equations in Hölder spaces, 1996
11 **Jacques Dixmier,** Enveloping algebras, 1996 Printing
10 **Barry Simon,** Representations of finite and compact groups, 1996
9 **Dino Lorenzini,** An invitation to arithmetic geometry, 1996
8 **Winfried Just and Martin Weese,** Discovering modern set theory. I: The basics, 1996
7 **Gerald J. Janusz,** Algebraic number fields, second edition, 1996
6 **Jens Carsten Jantzen,** Lectures on quantum groups, 1996
5 **Rick Miranda,** Algebraic curves and Riemann surfaces, 1995
4 **Russell A. Gordon,** The integrals of Lebesgue, Denjoy, Perron, and Henstock, 1994
3 **William W. Adams and Philippe Loustaunau,** An introduction to Gröbner bases, 1994
2 **Jack Graver, Brigitte Servatius, and Herman Servatius,** Combinatorial rigidity, 1993
1 **Ethan Akin,** The general topology of dynamical systems, 1993

Modern topology uses very diverse methods. This book is devoted largely to methods of combinatorial topology, which reduce the study of topological spaces to investigations of their partitions into elementary sets, and to methods of differential topology, which deal with smooth manifolds and smooth maps. Many topological problems can be solved by using either of these two kinds of methods, combinatorial or differential. In such cases, both approaches are discussed.

One of the main goals of this book is to advance as far as possible in the study of the properties of topological spaces (especially manifolds) without employing complicated techniques. This distinguishes it from the majority of other books on topology.

The book contains many problems; almost all of them are supplied with hints or complete solutions.

ISBN 0-8218-3809-1

9 780821 838099

GSM/74

For additional information and updates on this book, visit
www.ams.org/bookpages/gsm-74