# INTRODUCTION TO SET THEORY AND TOPOLOGY

*by*

KAZIMIERZ KURATOWSKI

*Professor of Mathematics*
*Member of the Polish Academy of Sciences*

*Containing a Supplement*
*on*
ELEMENTS OF ALGEBRAIC TOPOLOGY
*by*
*Professor* RYSZARD ENGELKING

COMPLETELY REVISED SECOND ENGLISH EDITION
FIRST EDITION TRANSLATED FROM POLISH BY LEO F. BOROŃ

PERGAMON PRESS
OXFORD • NEW YORK • TORONTO
SYDNEY • BRAUNSCHWEIG

PWN — POLISH SCIENTIFIC PUBLISHERS
WARSZAWA

Pergamon Press Ltd., Headington Hill Hall, Oxford
Pergamon Press Inc., Maxwell House, Fairview Park, Elmsford,
New York 10523
Pergamon of Canada Ltd., 207 Queen's Quay West, Toronto 1
Pergamon Press (Aust.) Pty. Ltd., 19a Boundary Street,
Rushcutters Bay, N.S.W. 2011, Australia
Vieweg & Sohn GmbH, Burgplatz 1, Braunschweig



Państwowe Wydawnictwo Naukowe
(PWN—Polish Scientific Publishers)
Warszawa

First Edition 1962
Second Edition 1972

Library of Congress Catalog Card Number 72-149123

Printed in Poland (D.R.P.)

08 016160 X

# FOREWORD TO THE FIRST ENGLISH EDITION

The ideas and methods of set theory and topology penetrate modern mathematics. It is no wonder then that the elements of these two mathematical disciplines are now an indispensable part of basic mathematical training. Concepts such as the union and intersection of sets, countability, closed set, metric space, and homeomorphic mapping are now classical notions in the whole framework of mathematics.

The purpose of the present volume is to give an accessible presentation of the fundamental concepts of set theory and topology; special emphasis being placed on presenting the material from the viewpoint of its applicability to analysis, geometry, and other branches of mathematics such as probability theory and algebra. Consequently, results important for set theory and topology but not having close connections with other branches of mathematics, are given a minor role or are omitted entirely. Such topics are, for instance, investigations on foundations, the theory of alephs, and the theory of curves.

The main body of the book is an introduction to set theory and topology, intended for the beginner. Sections marked with an asterisk cover either more complicated topics or points which are frequently omitted in a first course; this holds also for some exercises which allow the reader to get acquainted with many applications and some important results which could not be included in the text without unduly expanding it. Many new exercises not contained in the Polish edition have been included here.

I take great pleasure in thanking Professor J. Jaworowski and Dr. A. Granas for their cooperation in preparing the Polish edition and to thank also Professors A. Mostowski and R. Sikorski, Dr. S. Mrówka, Mr. R. Engelking and Dr. A. Schinzel for numerous comments which helped me to improve the original manuscript. Also, my thanks go to Mr. Leo F. Boroń and to Mr. A.H. Robinson for preparing the present text for English-speaking students of mathematics.

KAZIMIERZ KURATOWSKI

*Warsaw*
*September 1960*

# FOREWORD TO THE SECOND ENGLISH EDITION

Since the first English edition appeared, set theory and point-set topology have developed to such an extent that the author found it necessary to modify in many points the previous edition. This was done partially in the Polish edition (1965) and in the French edition (1966).

The most essential changes concern the second part of this book (devoted to topology). However, there are also changes worth noticing in the first part (on set theory).

The concepts of inverse limit, of lattice, of ideal, of filter, of a commutative diagram, and of a cartesian product of an arbitrary number of factors are considered. A slightly deeper insight into the axioms of set theory was needed; in particular, the notion of a class (in the sense of Bernays) is mentioned (and later applied, mainly in connection with the concept of category used in the Supplement).

In the theory of ordering relations, more emphasis was put on what was previously called partial ordering. This is now called, more concisely, ordering, and this change of terminology seems to be more appropriate to common use.

For the same reason, some notations have been changed. In particular, the Lebesgue notation $E_x\varphi(x)$ has been replaced by $\{x\colon \varphi(x)\}$; the union of members of a family $A$ of sets is denoted by $\bigcup A$, and the intersection by $\bigcap A$.

The changes in the second part of the book are more essential. In the first edition, this part of the book was chiefly devoted to the study of metric spaces. In this second edition, the general topological spaces form its main subject. Consequently, more than a half of the second part had to be written anew. It contains new topics which were not considered in the first edition, such as cartesian products of topological spaces, the Čech–Stone compactification, quotient-spaces, completely regular spaces, quasi-components, and a large number of exercises have been added.

In Chapter XX, on simplexes, more material will be found on simplicial mappings, on the nerve of a cover and related problems.

Finally, the rather short Chapter XXI, on complexes, chains and homologies, has been replaced by a much more extensive Supplement on the elements of algebraic topology. This Supplement, written by Professor Engelking, will certainly be a very valuable complement to my book.

I have received considerable help from the persons mentioned in the Foreword to the first edition and also from the young ladies Dr. Karłowicz and Dr. Vuilleumier. To all of them go my heartiest thanks.

KAZIMIERZ KURATOWSKI

*Warsaw*
*October 1968*

# INTRODUCTION TO PART I

The concept of a set is one of the most fundamental and most frequently used mathematical concepts. In every domain of mathematics we have to deal with sets such as the set of positive integers, the set of complex numbers, the set of points of a circle, the set of continuous functions, the set of integrable functions, and so forth.

The object of set theory is to investigate the properties of sets from the most general point of view; generality is an essential aspect of the theory of sets. In geometry we consider sets whose elements are points, in arithmetic we consider sets whose elements are numbers, in the calculus of variations we deal with sets of functions or curves; on the other hand, in the theory of sets we are concerned with the general properties of sets independently of the nature of the elements which comprise these sets. This will be made clear by several examples which we shall give here and by a brief overall view of the contents of the first part of this volume.

In Chapter II we shall consider operations on sets which are analogous to arithmetic operations: for every pair of sets $A$ and $B$ we shall form their *union* $A \cup B$, i.e. the set composed of all elements of the set $A$ and all elements of the set $B$; we shall also form the *intersection* $A \cap B$ of the sets $A$ and $B$, i.e. the set of all elements common to the sets $A$ and $B$. These operations have, in a certain sense, an algebraic character, e.g. they have the properties of commutativity, associativity and distributivity. It is clear that these properties do not depend on whether these sets consist of numbers, points or other mathematical objects; they are general properties of sets and therefore the investigation of these properties belongs to the realm of set theory.

In Chapter III we consider another operation, called *cartesian multiplication*. For two given sets $X$ and $Y$ we denote by $X \times Y$ the set of all pairs of elements $\langle x, y \rangle$ in which the first belongs to the set $X$ and the second to the set $Y$. Thus, e.g. if $X$ and $Y$ denote the set of real numbers then $X \times Y$ is the plane (whence the name "cartesian" product in honour of the great French mathematician Descartes (1596–1650), who, treating the plane as a set of pairs of real numbers, initiated a new branch of mathematics, called analytic geometry). The computational properties of cartesian multiplication in connection with the operations on sets mentioned above are given in Chapter III.

The concept of cartesian product allows us to define the concept of a *function* (or a *mapping*) in a completely general way. We shall concern ourselves with the concept of function in Chapter IV with emphasis on set-valued mappings. In the same chapter families of sets (in particular, Borel families) are considered.

An especially important role in the theory of sets is played by the one-to-one functions. These are functions which map the set $X$ onto the set $Y$ so that to every two distinct elements of the set $X$ there correspond two distinct elements of the set $Y$ (and then the inverse function with respect to the given function, which maps the set $Y$ onto the set $X$, is also one-to-one). If there exists such a one-to-one mapping of the set $X$ onto the set $Y$ we say that these sets are *of equal power*. The equality of powers is the generalization of the idea of equal number of elements; the significance of this generalization depends first of all on the fact that it can be applied to infinite as well as to finite sets. For example, it is easy to see that the set of all even numbers has the same power as the set of all odd numbers; on the other hand, the set of all real numbers does not have the same power as the set of all natural numbers—a fact which is not immediately obvious. Hence, we can—in some sense—classify infinite sets with respect to their power. We can also, thanks to this, extend the sequence of natural numbers, introducing numbers which characterize the power of infinite sets (called the *cardinal numbers*); in particular, to sets having the same power as the set of all natural numbers (or the *countably infinite* sets) we assign the cardinal number $\mathfrak{a}$, to the set of all real numbers we assign the number $\mathfrak{c}$ (the power of the continuum). It turns out that there is an infinite number of infinite cardinal numbers. However, in the applications of set theory to other branches of mathematics an essential role is played by only two of them: $\mathfrak{a}$ and $\mathfrak{c}$. So we also limit ourselves above all to the investigation of these two numbers. This forms the content of Chapters V and VI.

Chapter VII is devoted to *ordered sets* such as the set of all subsets of a given set (ordered by the relation of inclusion). Among the ordered sets, of particular importance are the *linearly ordered* sets such as the set of all natural numbers, the set of all rational numbers, the set of all real numbers. For each of these sets the relation $x \leqslant y$ determines the ordering; here the order types of these three sets differ in an essential manner: in the first of them there exist elements which are immediately adjacent to one another ($n$ and $n+1$), in the second there are no such elements (so we say, the ordering is dense), however, there exist gaps (in the Dedekind sense), but in the set of all real numbers there are no gaps.

An especially important kind of linearly ordered sets are the *well ordered* sets, i.e. those in which every non-empty subset has a least element. An example of a well ordered set is the set of all natural numbers (but the set of all integers is not well ordered since this set does not have a least element). Also well ordered—although of a different order type—is the set consisting of numbers of the form $1-1/n$ and numbers of the form $2-1/n$, $n = 1, 2, 3, \ldots$ In Chapter VIII we give the fundamental theorems concerning well ordering. So, we prove that given two distinct order types of well ordered sets one is always an extension of the other (in a sense which we shall make more precise). From this follows the important corollary that given two different well ordered sets one has power equal to that of a subset of the other; in the terminology of cardinal numbers this means that for two distinct cardinal numbers corresponding to well ordered sets, one is always smaller than the other. In connection with this theorem, there arises the fundamental problem: does there exist a relation for any set which establishes its well ordering? We shall prove that this is in fact so, if we assume the axiom of choice.

The discussion of set theory given here is based on a system of *axioms*. Even though in the introductory part of set theory, e.g. in the algebra of sets, the concept of set, with which we usually have to deal in applications to other branches of mathematics (and hence the concept of a set of numbers, points or curves, and so on), does not touch upon logical difficulties, a subsequent construction of set theory which is not based on a system of axioms turns out to be impossible; for there exist questions to which the so-called "naive" intuitive idea of a set does not give a unique answer. The lack of the necessary foundations of set theory in its initial period of development led to antinomies, i.e. contradictions, which arose from the "naive" intuitive idea of set. Only the axiomatic concept of the theory of sets allowed the removal of these antinomies (cf. Chapter VI, § 2, Remark 2).

In this book we do not analyse more closely the axiomatics of set theory or the logical foundation of the subject. Although these subjects form at the present time an important part of mathematics and are being actively developed, the discussion of them lies outside the principal goal of this book, which is: the presentation of the most important branches of set theory and topology from the point of view of their applications to other branches of mathematics. Therefore we limit ourselves to the formulation of some particularly important problems concerning axiomatics, such as the independence of the axiom of choice and of the continuum hypothesis. We mention also the existence of sets of power $\aleph_\omega$ and of inaccessible numbers, and we call attention to the necessity of introducing new axioms which imply their existence.

In the first part of this book the reader will find a certain amount of information on *mathematical logic*. The notation of mathematical logic is an indispensable tool of set theory and can be applied with great profit far beyond set theory. In Chapters I and III we give the main facts from this subject concerning the calculus of propositions, propositional functions and quantifiers. The notation of mathematical logic is not deprived of general didactic values; by examples for concepts such as uniform convergence or uniform continuity it is possible to observe how much the definition of these concepts gains in precision and lucidity, when they are written in the symbolism of mathematical logic.

In the first period of its existence, set theory was practically exclusively the creation of one scholar, G. Cantor (1845–1918). In the period preceding the appearance of the works of Cantor, there were published works containing concepts which are now included in the theory of sets (by authors such as Dedekind, Du

Bois-Reymond, Bolzano), but none the less the systematic investigation of the general properties of sets, the establishment of fundamental definitions and theorems and the creation on their foundation of a new mathematical discipline is the work of G. Cantor (during the years 1871–1883).

The stimulus to the investigations from which the theory of sets grew was given by problems of analysis, the establishing of the foundations of the theory of irrational numbers, the theory of trigonometric series, etc. However, the further development of set theory went initially in an abstract direction, little connected with other branches of mathematics. This fact, together with a certain strangeness of the methods of set theory which were entirely different from those applied up to that time, caused many mathematicians to regard this new branch of mathematics initially with a certain degree of distrust and reluctance. In the course of years, however, when set theory showed its usefulness in many branches of mathematics such as the theory of analytic functions or theory of measure, and when it became an indispensable basis for new mathematical disciplines (such as topology, the theory of functions of a real variable, the foundations of mathematics), it became an especially important branch and tool of modern mathematics.

The following list contains books which may be of aid in the study of set theory (also the books on topology mentioned in Part II, Introduction, contain the basic notions of set theory):

A. Abian, *The Theory of Sets and Transfinite Arithmetic*, Philadelphia, 1965.

P. S. Alexandrov, *Einführung in die Mengenlehre und die Theorie der reellen Funktionen*, Berlin, 1956.

P. Bernays, *Axiomatic Set Theory*, N.-Holland Publ. Comp., Amsterdam, 1958.

N. Bourbaki, *Théorie des Ensembles*, Hermann, Paris, 1970.

J. Breuer, *Introduction to the Theory of Sets*, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1958.

P. J. Cohen, *Set Theory and the Continuum Hypothesis*, Benjamin, New York, 1966.

A. Fraenkel, *Abstract Set Theory*, N.-Holland Publ. Comp., Amsterdam, 1953.

A. Fraenkel and Y. Bar-Hillel, *Foundations of Set Theory*, N.-Holland Publ. Comp., Amsterdam, 1958.
K. Gödel, *The Consistency of the Continuum Hypothesis*, Princeton, 1940.
P. Halmos, *Naive Set Theory*, Van Nostrand Comp., Princeton, 1960.
F. Hausdorff, *Set Theory*, Chelsea, New York, 1957.
D. Hilbert and P. Bernays, *Grundlagen der Mathematik*, 2 vols., Berlin, 1934–1939.
E. Kamke, *Theory of Sets*, Dover Publications, New York, 1950.
D. Klaua, *Allgemeine Mengenlehre*, Berlin, 1964.
A. H. Kruse, *Localization and Iteration of Axiomatic Set Theory*, Detroit, 1969.
K. Kuratowski and A. Mostowski, *Set Theory*, N.-Holland Publ. Comp. and PWN, Amsterdam–Warszawa, 1968.
A. P. Morse, *A Theory of Sets*, Academic Press, 1965.
A. Mostowski, *Constructible Sets with Applications*, N.-Holland Publ. Comp. and PWN, Amsterdam–Warszawa, 1969.
I. P. Natanson, *Theory of Functions of a Real Variable*, Ungar, New York, 1955, Chapters I and II.
J. E. Rubin, *Set Theory for the Mathematician*, Holden-Day, 1967.
Jürgen Schmidt, *Mengenlehre*, Mannheim, 1964.
W. Sierpiński, *Algèbre des ensembles*, Monografie Matematyczne, Warszawa–Wrocław, 1951.
—— *Cardinal and Ordinal Numbers*, Monografie Matematyczne, Warszawa–Wrocław, 1965.
L. E. Sigler, *Exercises in Set Theory*, Princeton, 1966.
P. Suppes, *Axiomatic Set Theory*, Van Nostrand Comp., Princeton, 1960.
A. Tarski, *Cardinal Algebras*, Oxford Univ. Press, New York, 1949.

CHAPTER I

# PROPOSITIONAL CALCULUS

We apply the propositional calculus to propositions each of which has one of two logical values, 0 and 1 (denoted also by $F$ and by $T$), where we assign the value 0 to a false proposition and the value 1 to a true proposition (in particular, all the propositions in mathematics are of this type, i.e. they take values either 0 or 1).†

## § 1. The disjunction and conjunction of propositions

If $\alpha$ and $\beta$ are two propositions, then we write "$\alpha$ or $\beta$" in the form of the *disjunction* $\alpha \vee \beta$ (called also the *sum*) and we write the proposition "$\alpha$ and $\beta$" in the form of the *conjunction* $\alpha \wedge \beta$ or $\alpha\beta$ (called also the *product*).

Clearly, the proposition $\alpha \vee \beta$ is true if at least one of the components is a true proposition and the proposition $\alpha \wedge \beta$ is true if both factors are true propositions. The above can be put in the form of the following table:

(1) $\quad 0 \vee 0 \equiv 0, \quad 0 \vee 1 \equiv 1, \quad 1 \vee 0 \equiv 1, \quad 1 \vee 1 \equiv 1,$

(2) $\quad 0 \wedge 0 \equiv 0, \quad 0 \wedge 1 \equiv 0, \quad 1 \wedge 0 \equiv 0, \quad 1 \wedge 1 \equiv 1.$

The equivalence sign used in the above formulas occurs between propositions; namely, the equivalence $\alpha \equiv \beta$ holds if and only if $\alpha$ and $\beta$ have the same logical value.

The disjunction and conjunction of propositions are *commutative* and *associative*, i.e.

(3)
$$\alpha \vee \beta \equiv \beta \vee \alpha, \quad \alpha \wedge \beta \equiv \beta \wedge \alpha,$$
$$\alpha \vee (\beta \vee \gamma) \equiv (\alpha \vee \beta) \vee \gamma, \quad \alpha \wedge (\beta \wedge \gamma) \equiv (\alpha \wedge \beta) \wedge \gamma.$$

† The propositional calculus in its actual form (called also algebra of logic) had its beginning in the papers of G. Boole and A. De Morgan around 1850.

The *distributive law*

$$(4) \qquad \alpha \wedge (\beta \vee \gamma) \equiv (\alpha \wedge \beta) \vee (\alpha \wedge \gamma)$$

also holds and more generally we have

$$(5) \quad (\alpha \vee \beta) \wedge (\gamma \vee \delta) \equiv (\alpha \wedge \gamma) \vee (\beta \wedge \gamma) \vee (\alpha \wedge \delta) \vee (\beta \wedge \delta).$$

We can verify the above laws—as well as all laws of the propositional calculus—by substituting the values 0 or 1 for the variables and then applying formulas (1) and (2).

## § 2. Negation

Next we introduce the operation of *negation* of a proposition $\alpha$ which we shall denote by $\alpha'$ (or by $\sim \alpha$ or $\neg\alpha$). The negation of a true proposition is a false proposition and, conversely, the negation of a false proposition is a true proposition. We therefore have the following table:

$$(6) \qquad 1' \equiv 0, \quad 0' \equiv 1.$$

From this we obtain the so-called *law of double negation,*

$$(7) \qquad \alpha'' \equiv \alpha.$$

Two fundamental theorems of Aristotelian logic (which follow easily from formulas (1), (2) and (6)) hold:

$$(8) \qquad \alpha \vee \alpha' \equiv 1, \quad \alpha \wedge \alpha' \equiv 0;$$

they are the *law of the excluded middle (principium tertii exclusi)* and the *law of contradiction* (these are formulated in classical logic in the following manner: from two contradictory propositions, one is true; no proposition can be true simultaneously with its negation).

Further, the important *De Morgan laws* hold:

$$(9) \qquad (\alpha \vee \beta)' \equiv (\alpha' \wedge \beta'),$$

$$(10) \qquad (\alpha \wedge \beta)' \equiv (\alpha' \vee \beta').$$

The first of these laws asserts that, if it is not true that one of the propositions $\alpha$ and $\beta$ is true, then both of these propositions are false (and conversely); i.e. the negation of the first as well as the negation of the second are true propositions.

Similarly, if it is not true that both propositions $\alpha$ and $\beta$ are true, then this means that the negation of one of them is a true proposition, and conversely.

Taking the negation of both members of identity (10) we obtain, by virtue of formula (7), the identity

(11) $$\alpha \wedge \beta \equiv (\alpha' \vee \beta')'.$$

From this it is clear that conjunction can be defined with the aid of disjunction and negation (and, in a similar manner, one could define disjunction with the aid of conjunction and negation). This allows the reduction of the number of fundamental operations to two; however, from the computational-technical viewpoint it is more convenient to make use of three operations: disjunction, conjunction and negation.

## § 3. Implication

We write $\alpha \Rightarrow \beta$ if the proposition $\alpha' \vee \beta$ is true, i.e.

(12) $$(\alpha \Rightarrow \beta) \equiv (\alpha' \vee \beta);$$

$\alpha \Rightarrow \beta$ is read: the proposition $\alpha$ *implies* the proposition $\beta$, or: *if* $\alpha$ *then* $\beta$.

Tables (1) and (6) yield the following table:

(13) $$(0 \Rightarrow 0) \equiv 1, \quad (0 \Rightarrow 1) \equiv 1, \quad (1 \Rightarrow 0) \equiv 0, \quad (1 \Rightarrow 1) \equiv 1.$$

We also deduce from this that

(14) *if* $\alpha \Rightarrow \beta$ *and* $\beta \Rightarrow \alpha$ *then* $\alpha \equiv \beta$.

Clearly, implication has properties analogous to deduction. However, the current meaning of the expression "deduction" is different from the expression "implication". To say that a proposition $\beta$ is deducible from a proposition $\alpha$ (e.g. from a given theorem) usually means the possibility of proving proposition $\beta$ on the basis of proposition $\alpha$; but the implication $\alpha \Rightarrow \beta$ always holds, provided that the proposition $\beta$ is true (even if the proposition $\alpha$ were false).

Let us note further two laws: the *syllogism law* (or the law of transitivity of implication) and the *law of contraposition* (on which the proof by "*reductio ad absurdum*", or the indirect method of proof, depends):

(15) $$\text{if } \alpha \Rightarrow \beta \text{ and } \beta \Rightarrow \gamma \text{ then } \alpha \Rightarrow \gamma;$$

(16) $$\text{if } \beta' \Rightarrow \alpha' \text{ then } \alpha \Rightarrow \beta.$$

**Exercises**

**1.** Prove that if $\alpha$ is a true proposition, then $\beta \Rightarrow \alpha$ is also a true proposition. *Hint*: In this and the following exercises apply the "zero-unit" tables (1), (2), (6), (13).

**2.** If $\alpha' \Rightarrow \beta$ for each $\beta$, then $\alpha$ is a true proposition. *Law of Clausius.*

**3.** If $\alpha$ is a false proposition, then $\alpha \Rightarrow \beta$. *Law of Duns Scotus.*

**4.** Prove that $\alpha \wedge \beta \Rightarrow \alpha \Rightarrow \alpha \vee \beta$.

**5.** If $\alpha \Rightarrow \beta$ and $\gamma \Rightarrow \delta$, then $\alpha \wedge \gamma \Rightarrow \beta \wedge \delta$ and $\alpha \vee \gamma \Rightarrow \beta \vee \delta$.

**6.** If $\alpha \Rightarrow \beta$, then $\alpha \wedge \beta \equiv \alpha$ and $\alpha \vee \beta \equiv \beta$.

**7.** Prove that $\alpha \vee (\alpha \wedge \beta) \equiv \alpha \equiv \alpha \wedge (\alpha \vee \beta)$. *Law of absorption.* More generally: $\alpha \vee (\beta \wedge \gamma) \equiv (\alpha \vee \beta) \wedge (\alpha \vee \gamma)$.

**8.** Let $(\alpha \div \beta) \equiv [(\alpha \wedge \beta') \vee (\alpha' \wedge \beta)]$. Prove that $(\alpha \vee \beta) \equiv [(\alpha \div \beta) \div (\alpha \wedge \beta)]$. We call $\alpha \div \beta$ the *symmetric difference* of the propositions $\alpha$ and $\beta$; what is its logical meaning?

**9.** Prove the *laws of tautology*: $\alpha \vee \alpha \equiv \alpha$, $\alpha \wedge \alpha \equiv \alpha$.

**10.** Prove that $\alpha \vee 1 \equiv 1$, $\alpha \vee 0 \equiv \alpha$, $\alpha \wedge 1 \equiv \alpha$, $\alpha \wedge 0 \equiv 0$.

CHAPTER II

# ALGEBRA OF SETS. FINITE OPERATIONS

## § 1. Operations on sets

The *union* of two sets $A$ and $B$ is understood to be the set whose elements are all the elements of the set $A$ and all the elements of the set $B$ and which does not contain any other elements. We denote the union of the sets $A$ and $B$ by the symbol $A \cup B$.

The *intersection* of two sets $A$ and $B$ is understood to be the common part of these sets, i.e. the set containing those and only those elements which belong simultaneously to $A$ and to $B$. We denote the intersection of the sets $A$ and $B$ by the symbol $A \cap B$.

Finally, the *difference* of two sets $A$ and $B$, i.e. the set $A-B$, is the set consisting of those and only those elements which belong to $A$ but which do not belong to $B$ (instead of $A-B$ the symbols $A \setminus B$ and $A \sim B$ are also used).

The following examples illustrate the operations on sets: the union of the set of rational numbers and the set of irrational numbers is the set of all real numbers; the intersection of the set of numbers which are divisible by 2 and the set of numbers divisible by 3 is the set of numbers which are divisible by 6; the difference of the set of natural numbers and the set of even natural numbers is the set of odd natural numbers.

Other examples are given in Figs. 1–3,† where the sets $A$ and $B$ are circular disks. From Fig. 2 we see that there exists no point which belongs to both the sets $A$ and $B$; but despite this fact, we can consider forming the intersection to be possible in all cases by adopting the following definition.

† Called *Euler circles*, which are particular cases of *Venn diagrams* (John Venn, 1834–1923, an English logician).

The *null set* (or the *empty set* or the *void set*) is the set which contains no elements; we denote it by the symbol $\emptyset$.

Thus, in Fig. 2 we have $A \cap B = \emptyset$ and in Fig. 3 we have $B - A = \emptyset$.

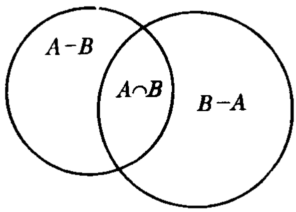



FIG. 1

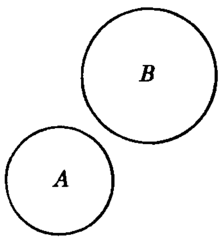



FIG. 2

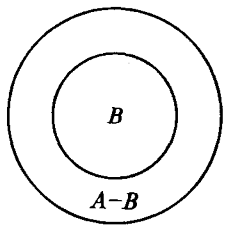



FIG. 3

The equality $A \cap B = \emptyset$ therefore denotes that the sets $A$ and $B$ do not have common elements. We then say that these sets are *disjoint*.

The role of the null set in set theory is analogous to the role of the number 0 in arithmetic; these concepts are necessary in order that it be possible to carry out all operations with no exception.

### § 2. Inter-relationship with the propositional calculus

Operations on sets are closely related to operations on propositions. Let us write $x \in A$,† to denote that $x$ is an element of the set $A$ (as a rule we shall denote elements with lower-case letters and sets with upper-case letters).

† The sign $\in$, introduced by G. Peano, is an abbreviation of the Greek word ἐστί (to be).

We assume also that

$$(x \notin A) \equiv (x \in A)'.$$

The following equivalences hold for all $x$:

$$[x \in (A \cup B)] \equiv (x \in A) \vee (x \in B), \tag{1}$$

$$[x \in (A \cap B)] \equiv (x \in A) \wedge (x \in B), \tag{2}$$

$$[x \in (A - B)] \equiv (x \in A) \wedge (x \in B)'. \tag{3}$$

By virtue of formulas (1)–(3), we can easily deduce theorems on the calculus of sets from analogous theorems in propositional calculus.

In this connection, let us note that

(4) *if the equivalence* $x \in A \equiv x \in B$ *holds for all* $x$, *then* $A = B$;

and therefore the proof of the equality $A = B$ reduces to showing that $x$ belongs to $A$ if and only if it belongs to $B$.

The operations of union and intersection of sets are commutative, i.e.

$$A \cup B = B \cup A, \qquad A \cap B = B \cap A. \tag{5}$$

These operations also satisfy the associative law:

$$\begin{aligned} A \cup (B \cup C) &= (A \cup B) \cup C, \\ A \cap (B \cap C) &= (A \cap B) \cap C. \end{aligned} \tag{6}$$

The distributive law

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \tag{7}$$

also holds, as can easily be verified.

It follows from this that

$$(A \cup B) \cap (C \cup D) = (A \cap C) \cup (B \cap C) \cup (A \cap D) \cup (B \cap D), \tag{8}$$

for, by virtue of formula (7), we have

$$\begin{aligned} (A \cup B) \cap (C \cup D) &= [(A \cup B) \cap C] \cup [(A \cup B) \cap D] \\ &= (A \cap C) \cup (B \cap C) \cup (A \cap D) \cup (B \cap D). \end{aligned}$$

Therefore, in general, as in arithmetic, in order to expand the intersection of two unions one must take the intersection of each term of the first union with each term of the second union and then form the union of the intersections obtained in this manner.

The analogy between arithmetic and the theory of sets is not, however, complete. For example, the following identities (compare the laws of tautology stated in Chapter I, Exercise 9) hold in set theory:

$$A \cup A = A, \tag{9}$$

$$A \cap A = A, \tag{10}$$

which point out that, in contrast to arithmetic, neither multiples nor exponents arise in set theory.

## § 3. Inclusion

We shall now introduce the important relation of *inclusion* between sets. We shall say that the set $A$ is a *subset* of the set $B$ (or also that the set $A$ is *contained* in $B$) if every element of the set $A$ is an element of the set $B$. We then write $A \subset B$ (or $B \supset A$).

We therefore have the following equivalence:

(11) $(A \subset B) \equiv$ [*the implication* $(x \in A) \Rightarrow (x \in B)$ *holds for all* $x$].

In particular, it follows from this that

$$A \subset A, \tag{12}$$

i.e. that every set is a subset of itself. Because of this inclusion, we also use the term *proper* subset for subsets of a given set which are different from the given set.

Obviously

(13) *if* $A \subset B$ *and* $B \subset A$, *then* $A = B$,

for the sets $A$ and $B$ consist of the same elements in this case.

Hence, in order to prove that $A = B$ it suffices to prove that $A \subset B$ and $B \subset A$; in other words, instead of the equivalence

$$(x \in A) \equiv (x \in B)$$

we prove the two implications

$$x \in A \Rightarrow x \in B \quad \text{and} \quad x \in B \Rightarrow x \in A$$

(cf. Chapter I, (15)).

It can easily be proved that

(14) *if* $A \subset B$ *and* $B \subset C$, *then* $A \subset C$,

(15) $(A \cap B) \subset A \subset (A \cup B), \quad A - B \subset A,$

(16) *if* $A \subset B$ *and* $C \subset D$, *then* $(A \cup C) \subset (B \cup D)$ *and* $(A \cap C) \subset (B \cap D)$.

The following equivalences hold:

(17) $$(A \subset B) \equiv (A \cup B = B) \equiv (A \cap B = A).$$

For, let $A \subset B$. Combining this inclusion with the inclusion $B \subset B$ (cf. (12)), we obtain, by virtue of (16) and (9),

$$(A \cup B) \subset (B \cup B) = B,$$

but since (cf. (15)) $B \subset (A \cup B)$, we have $A \cup B = B$ (cf. (13)).

Conversely, it follows from the relation $A \cup B = B$ that $A \subset B$ (by virtue of (15)); hence these relations are equivalent.

Similarly, combining the inclusion $A \subset B$ with the inclusion $A \subset A$ we obtain $A \subset (A \cap B)$, whence $A = A \cap B$, because $A \cap B \subset A$ by virtue of (15).

Conversely, from the relation $A \cap B = A$ we obtain the relation $A \subset B$ because $(A \cap B) \subset B$.

From this we deduce the following formula which is important in applications:

(18) $$(A \cup B) \cap (A \cup C) = A \cup (B \cap C).$$

In fact, by virtue of (8) and (10) we have

$$(A \cup B) \cap (A \cup C) = (A \cap A) \cup (B \cap A) \cup (A \cap C) \cup (B \cap C)$$
$$= A \cup (B \cap A) \cup (A \cap C) \cup (B \cap C),$$

and $(B \cap A) \subset A$ by (15), and hence by (17) $A \cup (B \cap A) = A$ and similarly $A \cup (A \cap C) = A$. Formula (18) follows.

Let us note further the following formulas, the proofs of which do not present any difficulties:

(19) $$A \cap B = A - (A - B),$$

(20) $$A \cup (B - A) = A \cup B,$$

(21) $$A - (A \cap B) = A - B,$$

(22) $$A \cap (B - C) = (A \cap B) - C.$$

## § 4. Space. Complement of a set

In the applications of the theory of sets, we assume, as a rule, that all the sets under consideration are subsets of some fixed set, called the *space*. For example, in analysis, the set of real numbers or the set of complex numbers forms a space, and in geometry we have to deal with the Euclidean space.

Under this assumption, the theorems of the algebra of sets assume a still simpler form which is closer to the calculus of propositional functions.

Hence, let 1 denote a given space (this notation is expedient from the calculational point of view). We therefore have $A \subset 1$ for each of the sets $A$ considered. We denote by $A^c$ (or by $\sim A$ or $CA$) the set of elements of the space which do not belong to $A$, i.e.

$$A^c = 1 - A.$$

$A^c$ is called the *complement* of the set $A$ (with respect to the given space 1). We therefore have

$$(23) \qquad x \in A^c \equiv (x \in A)' \equiv x \notin A.$$

Formulas (6)–(8) (Chapter I, § 2) yield immediately the (almost obvious) formulas:

$$(24) \qquad 1^c = \emptyset, \quad \emptyset^c = 1,$$

$$(25) \qquad A^{cc} = A,$$

$$(26) \qquad A \cup A^c = 1, \quad A \cap A^c = \emptyset.$$

Formulas (3), (23) and (2) imply the formula

$$(27) \qquad A - B = A \cap B^c,$$

which allows us to define subtraction in terms of intersection and complementation.

In fact,

$$(x \in A - B) \equiv (x \in A)(x \in B)'$$
$$\equiv (x \in A)(x \in B^c) \equiv (x \in A \cap B^c).$$

Formula (16) (Chapter I, § 3) implies that:

(28) *if* $B^c \subset A^c$, *then* $A \subset B$.

Finally, formulas (9) and (10) (Chapter I, § 2) yield the *De Morgan laws* for sets:

(29) $$(A \cup B)^c = A^c \cap B^c,$$

(30) $$(A \cap B)^c = A^c \cup B^c.$$

For, we have

$$x \in (A \cup B)^c \equiv [x \in (A \cup B)]' \equiv [(x \in A) \vee (x \in B)]'$$
$$\equiv (x \in A^c) \wedge (x \in B^c) \equiv x \in (A^c \cap B^c).$$

The proof of formula (30) is analogous.

The obvious formula

(31) $$A \cap 1 = A$$

yields by (26), that

(32) $$A = (A \cap B) \cup (A \cap B^c),$$

inasmuch as

$$A = A \cap 1 = A \cap (B \cup B^c) = (A \cap B) \cup (A \cap B^c).$$

Formula (17) can be supplemented by the following equivalence which is frequently applied in practice:

(33) $$(A \subset B) \equiv (A \cap B^c = \emptyset).$$

For, forming the intersection of both sides of the inclusion $A \subset B$ with $B^c$ we obtain $(A \cap B^c) \subset (B \cap B^c) = \emptyset$ (by virtue of (26)). But from formula (32) we deduce, assuming the equality $A \cap B^c = \emptyset$, that $A = A \cap B$, whence, by (17) it follows that $A \subset B$.

## § 5. The axiomatics of the algebra of sets

In the considerations up to this point we used only some properties of sets. The properties can be taken as a system of axioms, from which all the theorems of set theory, given above, follow.

We take, namely, as primitive concepts the concept of element of set and the relation of an element belonging to a set, i.e. the relation $x \in A$. We assume the following four axioms.

I. UNIQUENESS AXIOM (called also *Axiom of extensionality*). *If the sets A and B have the same elements then A and B are identical.*

II. UNION AXIOM. *For arbitrary sets A and B there exists a set whose elements are all the elements of the set A and all the elements of the set B, and which does not contain any other element.*

III. DIFFERENCE AXIOM. *For arbitrary sets A and B there exists a set whose elements are those and only those elements of the set A which are not elements of the set B.*

IV. EXISTENCE AXIOM. *There exists at least one set.*

It is not necessary to assume an axiom on the existence of an intersection because, as we saw (formula (19)), the intersection can be defined in terms of the difference. Likewise, the existence of the void set is a consequence of our system of axioms, for we can define the void set by means of the formula $\emptyset = A - A$, where $A$ is an arbitrary set (the existence of at least one set is guaranteed by axiom IV).

An important consequence of axiom I is the uniqueness of the operations, i.e. for given sets $A$ and $B$ there exists only one set satisfying axiom II (which justifies the use of the symbol $A \cup B$ to denote this set); the same applies to the intersection and difference.

As we have already stated, it is possible from the above axioms to deduce all the theorems of the theory of sets considered till now, without referring back to the intuitive concept of set.

## § 6. Boolean algebra.† Lattices

Examining the theorems of §§ 2–4, we see that the symbol $\in$ does not occur in the majority of them (though it does appear in their proofs). This suggests stating a system of axioms which will enable us to prove those theorems without referring to the relation $\in$.

We take as primitive concepts the set $\emptyset$ and the operations $\cup$, $\cap$, $-$, and we assume the following axioms:

(1°) $$A \cup B = B \cup A,$$

(2°) $$A \cap B = B \cap A,$$

(3°) $$A \cup (B \cup C) = (A \cup B) \cup C,$$

† For more details on Boolean algebra, see R. Sikorski, *Boolean Algebras*, 2nd edition, Berlin, 1964, and R. Halmos, *Lectures on Boolean Algebras*, Princeton, 1963.

(4°) $$A \cap (B \cap C) = (A \cap B) \cap C,$$

(5°) $$A \cup \emptyset = A,$$

(6°) $$A \cup (A \cap B) = A,$$

(7°) $$A \cap (A \cup B) = A,$$

(8°) $$A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$$

(9°) $$(A - B) \cup B = A \cup B,$$

(10°) $$(A - B) \cap B = \emptyset.$$

From these axioms we are able to deduce all the theorems of the algebra of sets in which the relation $\in$ does not appear. Also, if we desire to restrict the domain of the variables to subsets of a fixed set 1, we assume, in addition, the axiom

(11°) $$A \cap 1 = A.$$

We add that we define inclusion with the aid of the formula (cf. (17)):

$$(A \subset B) \equiv (A \cup B = B).$$

The theory based on the above axioms is called *Boolean algebra*. The applications of Boolean algebra extend far beyond the theory of sets; we need not interpret the variables $A, B, \ldots$ as sets. Interpreting them, e.g. as propositions, we obtain the propositional calculus.

This explains the duality between the propositional calculus and the algebra of sets: to the disjunction (or sum) $\vee$ of propositions corresponds the union $\cup$ of sets, to the conjuction (product) $\wedge$ of propositions—the intersection $\cap$ of sets, to the negation $\alpha'$ of a proposition $\alpha$—the complement $A^c$ of a set $A$, etc. (see also Chapter IV, § 3).

Other interpretations of Boolean algebra in recent times permit us to apply it in various branches of mathematics, and even outside mathematics (for example, in the theory of electrical networks).

R e m a r k. If we omit axioms (8°)–(10°), we obtain the notion of *lattice* (with 0 and 1).† Of course, every Boolean algebra

† For a detailed study of lattices, see G. Birkhoff, *Lattice Theory*, New York, 1961.

(with 1) is a lattice. However, the converse is not true. There are important examples of distributive lattices which are not Boolean algebras; such is the family of closed subsets of a topological space which is a lattice (since the union and intersection of two closed sets is closed) but is not a Boolean algebra (since the difference of two closed sets does not need to be closed, see Chapter X, § 4).

## § 7. Ideals and filters

A non-empty family† $\boldsymbol{R}$ of subsets of 1 is called *ideal* if the two following conditions are fulfilled:

$$(A \in \boldsymbol{R})(B \subset A) \Rightarrow B \in \boldsymbol{R},$$
$$(A \in \boldsymbol{R})(B \in \boldsymbol{R}) \Rightarrow (A \cup B) \in \boldsymbol{R}.$$

A non-empty family $\boldsymbol{S}$ is called a *filter* if

$$(A \in \boldsymbol{S})(A \subset B) \Rightarrow B \in \boldsymbol{S},$$
$$(A \in \boldsymbol{S})(B \in \boldsymbol{S}) \Rightarrow (A \cap B) \in \boldsymbol{S}.$$

It is easy to show that a family of sets is a filter iff the family of the complements of these sets is an ideal.

Obviously the family of all subsets of a given set $E$ is an ideal and the family of all sets $F$ such that $E \subset F \subset 1$ is a filter.

A *proper ideal* (i.e. an ideal which does not contain 1) is called *maximal* if it is not a subset of any other proper ideal. The definition of a maximal filter (called also *ultrafilter*) is analogous (we mean by *proper filter* a filter which does not contain $\emptyset$).

One shows with the aid of the axiom of choice (see Chapter VIII, Exercise 12) that each proper ideal (filter) is contained in a maximal ideal (filter).

### Exercises

1. Prove the following formulas:

(a) $$A \cup (A \cap B) = A = A \cap (A \cup B),$$

(b) $$(A \cup B) - C = (A - C) \cup (B - C),$$

(c) $$A - (B - C) = (A - B) \cup (A \cap C),$$

(d) $$A - (B \cup C) = (A - B) - C.$$

---

† By *family* we mean a set of sets. We write then $\boldsymbol{R}$ instead of $R$.

**2.** The set

$$A \dot{-} B = (A-B) \cup (B-A)$$

is called the *symmetric difference* of the sets $A$ and $B$.

Prove the following formulas:

(a) $$A \dot{-} (B \dot{-} C) = (A \dot{-} B) \dot{-} C \text{ (associativity)},$$

(b) $$A \cap (B \dot{-} C) = A \cap B \dot{-} A \cap C \text{ (distributivity)},$$

(c) $$A \cup B = A \dot{-} B \dot{-} A \cap B,$$

(d) $$A-B = A \dot{-} A \cap B,$$

(e) define $\cap$ by means of (i) $\dot{-}$, $\cup$, (ii) $-$, $\dot{-}$,

(f) $$(A \dot{-} B = C) \Rightarrow (B = A \dot{-} C),$$

(g) $$(A_1 \cup \ldots \cup A_n) \dot{-} (B_1 \cup \ldots \cup B_n) \subset (A_1 \dot{-} B_1) \cup \ldots \cup (A_n \dot{-} B_n),$$

$$(A_1 \cap \ldots \cap A_n) \dot{-} (B_1 \cap \ldots \cap B_n) \subset (A_1 \dot{-} B_1) \cup \ldots \cup (A_n \dot{-} B_n)$$

(Hausdorff).

**3.** Let $A = A_1 \dot{-} A_2 \dot{-} \ldots \dot{-} A_n$. Prove that $A$ is the set of elements belonging to an odd number of sets $A_1, \ldots, A_n$. (Thus the set $A$ is not affected by changing order in which operations are performed.)

**4.** We define *division* by means of the formula $A{:}B = A \cup B^c$. Compute

$$A{:}(B \cap C), \quad A{:}(B \cup C), \quad A \cap (B{:}A).$$

**5.** Let $A_1, A_2, \ldots, A_n$ be fixed subsets of the space 1. Let us assume that $A_i^1 = 1-A_i$, $A_i^0 = A_i$. Every intersection of the form

$$A_1^{i_1} \cap A_2^{i_2} \cap \ldots \cap A_n^{i_n}, \quad \text{where} \quad i_j = 0 \text{ or } 1,$$

is called a *constituant* of the space (with respect to the sets $A_1, A_2, \ldots, A_n$).

Prove that the constituants are disjoint and that their union is equal to 1 (therefore the decomposition into constituants effects a classification of the elements of the space with respect to their belonging to sets $A_1, A_2, \ldots, A_n$).

**6.** Represent the set $A-(B-C)$ as the union of constituants of the space with respect to the sets $A$, $B$ and $C$.

**7.** The sets $A_1, \ldots, A_n$ are called *independent*† if all the constituants are non-empty. Prove that in this case the number of constituants is $2^n$.

**8.** Let $\mathscr{I}^n$ denote the $n$-dimensional cube composed of points $(x_1, \ldots, x_n)$ such that $0 \leqslant x_i \leqslant 1$ for $i = 1, 2, \ldots, n$. Denote by $I_m$ the set of points $(x_1, \ldots, x_n)$ where $1/2 \leqslant x_m \leqslant 1$. Show that the sets $I_1, \ldots, I_n$ are independent. What is the geometrical interpretation for $n = 3$?

**9.** We say that the operations $x+y$ and $x \cdot y$ form a (commutative) *ring* if they satisfy the following conditions:

(i) $$x+y = y+x,$$

(ii) $$x+(y+z) = (x+y)+z,$$

† The notion of independent sets has important applications in probability theory. See E. Marczewski, Indépendence d'ensembles et prolongement de mesure, *Coll. Math.* **1** (1948), pp. 122–132.

(iii) there exists an element 0 such that $x+0 = x$,

(iv) for every pair $x, y$ there exists an element $z$ $(z = x-y)$ such that $y+z = x$,

(v) $$x \cdot y = y \cdot x,$$

(vi) $$x \cdot (y \cdot z) = (x \cdot y) \cdot z,$$

(vii) $$x \cdot (y+z) = (x \cdot y)+(x \cdot z).$$

Prove that sets form a ring with respect to the operations $A \dot{-} B$ and $A \cap B$, but they do not form a ring with respect to the operations $A \cup B$ and $A \cap B$.

**10.** Show that a proper ideal (filter) $\boldsymbol{R}$ is maximal iff

$$\text{for each } X \subset 1, \text{ either } X \in \boldsymbol{R}, \text{ or } (1-X) \in \boldsymbol{R}.$$

**11.** Let $\boldsymbol{A}$ be a family of subsets of the space 1. Consider the family $\boldsymbol{R}$ of sets of the form

$$(A_1 \cap X_1) \cup \ldots \cup (A_n \cap X_n), \quad \text{where} \quad A_1, \ldots, A_n \in \boldsymbol{A}.$$

Show that $\boldsymbol{R}$ is an ideal containing $\boldsymbol{A}$ (it is the ideal "generated by $\boldsymbol{A}$").

Establish the dual theorem for filters.

**12.** Show that the family of all linear subsets of $\mathscr{E}^n$ (i.e. of the origin, straight lines, planes, etc.) passing through the origin is a (non-distributive) lattice (with 0 and 1) relative to the operations $\cup$ and $\cap$ defined as follows: $A \cap B$ is the intersection defined in the usual sense and $A \cup B$ is the least linear set containing $A$ and $B$.

CHAPTER III

# PROPOSITIONAL FUNCTIONS. CARTESIAN PRODUCTS

Let a fixed set $A$ be given, which in the sequel we shall consider to be the space. Let $\varphi(x)$ be an expression which becomes a proposition when one substitutes for $x$ an arbitrary value of $x$ belonging to $A$. We call this expression a *propositional function*† (with the argument bounded by $A$).

For example, if the space is the set of all real numbers, then the expression "$x > 0$" is a propositional function; it becomes a true proposition if we substitute, say, 1 for $x$; it becomes a false proposition if we substitute $-1$ for $x$.

## § 1. The operation $\{x: \varphi(x)\}$

The set of all those values of the variable $x$ for which $\varphi(x)$ is a true proposition (or, as we say, the set of $x$'s which satisfy the propositional function $\varphi(x)$) is denoted by the symbol

$$\{x: \varphi(x)\}.$$

For example, in the space of real numbers $\{x: (x > 0)\}$ is the set of all positive numbers, $\{x: (x = x)\}$ is the set of all real numbers, and $\{x: (x+1 = x)\}$ is the null set.

It follows from the definition of the operation $\{x: \varphi(x)\}$ that a necessary and sufficient condition, that the element $a$ should belong to the set $\{x: \varphi(x)\}$, is that the proposition $\varphi(a)$ be true. Hence, the following equivalence holds:

(1) $$\textit{for every } a\text{: } [a \in \{x: \varphi(x)\}] \equiv \varphi(a).$$

The following four formulas hold:

(2) $$\{x: \varphi(x) \vee \psi(x)\} = \{x: \varphi(x)\} \cup \{x: \psi(x)\},$$

† Following Bertrand Russell the term *predicate* is also used in the same sense (see Hilbert and Ackermann, *Grundlagen der Mathematik*, vol. I, 1928).

$$(3) \qquad \{x\colon \varphi(x) \wedge \psi(x)\} = \{x\colon \varphi(x)\} \cap \{x\colon \psi(x)\},$$

$$(4) \qquad \{x\colon \varphi(x) \wedge [\psi(x)]'\} = \{x\colon \varphi(x)\} - \{x\colon \psi(x)\},$$

$$(5) \qquad \{x\colon [\varphi(x)]'\} = \{x\colon \varphi(x)\}^c.$$

We obtain the proof of the formula (2) from the formula (1) above and formula (1) of Chapter II, § 2:

$$\begin{aligned} a \in \{x\colon \varphi(x) \vee \psi(x)\} &\equiv \varphi(a) \vee \psi(a) \\ &\equiv [a \in \{x\colon \varphi(x)\}] \vee [a \in \{x\colon \psi(x)\}] \\ &\equiv a \in \{x\colon \varphi(x)\} \cup \{x\colon \psi(x)\}, \end{aligned}$$

whence equality (2) follows (cf. Chapter II, § 2, (4)).

Formulas (3)–(5) are proved similarly.

## § 2. Quantifiers

Let us now consider the following two operations on propositional functions:

$$\bigvee_x \varphi(x) \quad \text{and} \quad \bigwedge_x \varphi(x).$$

We read the formula $\bigvee_x \varphi(x)$ as follows: there exists *some* $x$ which satisfies the function $\varphi(x)$; $\bigwedge_x \varphi(x)$ denotes that *every* $x$ satisfies this function. (The symbols $\exists_x, \Sigma_x$ and $\forall_x, \Pi_x$, respectively, are used in the same sense.)

Clearly, the above operations transform propositional functions into propositions. The symbols of these operations $\bigvee$ and $\bigwedge$ are called the *existential* and the *universal quantifiers*, respectively.†

For example, in the space of real numbers the proposition $\bigvee_x (x > 0)$ is true but the proposition $\bigwedge_x (x > 0)$ is false.

The variable $x$ which appears as the *free* variable in the propositional function $\varphi(x)$ becomes a *bound* variable in the proposition $\bigvee_x \varphi(x)$ (like $x$ in $\int_0^1 f(x)dx$).

It may be noted that

$$\bigvee_x \varphi(x) \equiv \bigvee_y \varphi(y).$$

† The concept of a quantifier was introduced by G. Frege (*Begriffsschrift*, 1879) and studied by C. S. Peirce (1885).

Analogous remarks can be made about the universal quantifier.

The operations $\bigvee$ and $\bigwedge$ may be considered as generalizations of the operations of disjunction and conjunction. For, if the domain of variation of $x$ is finite, consisting of the elements $a_1$, $a_2$, ..., $a_n$, then

$$\text{(6)}\qquad \begin{aligned} \bigvee_x \varphi(x) &\equiv [\varphi(a_1) \vee \varphi(a_2) \vee \ldots \vee \varphi(a_n)], \\ \bigwedge_x \varphi(x) &\equiv [\varphi(a_1) \wedge \varphi(a_2) \wedge \ldots \wedge \varphi(a_n)]. \end{aligned}$$

We now set down the following easily proved formulas:

(7) *for every* $x_0$ *we have* $[\bigwedge_x \varphi(x)] \Rightarrow \varphi(x_0) \Rightarrow [\bigvee_x \varphi(x)]$,

$$\text{(8)}\qquad [\bigvee_x \varphi(x) \vee \bigvee_x \psi(x)] \equiv \bigvee_x [\varphi(x) \vee \psi(x)],$$

$$\text{(9)}\qquad [\bigvee_x \varphi(x) \wedge \psi(x)] \Rightarrow \bigvee_x \varphi(x) \wedge \bigvee_x \psi(x).$$

Let us note that in formula (9) we cannot replace the implication sign by the equivalence sign; in other words, implication in the opposite direction may not hold. For example, both of the propositions

$\bigvee_x$ ($x$ is a positive number) and $\bigvee_x$ ($x$ is a negative number)

are true, and hence a true proposition appears in the right member of formula (9); but on the left side there appears, in this example, a false proposition (inasmuch as there is no number which is simultaneously positive and negative).

The duals of formulas (8) and (9) are the following formulas:

$$\text{(10)}\qquad [\bigwedge_x \varphi(x) \wedge \bigwedge_x \psi(x)] \equiv \bigwedge_x [\varphi(x) \wedge \psi(x)],$$

$$\text{(11)}\qquad [\bigwedge_x \varphi(x) \vee \bigwedge_x \psi(x)] \Rightarrow \bigwedge_x [\varphi(x) \vee \psi(x)].$$

This duality is expressed by the *generalized De Morgan formulas* (which appear very frequently in applications):

$$\text{(12)}\qquad [\bigwedge_x \varphi(x)]' \equiv \bigvee_x \varphi'(x),$$

$$\text{(13)}\qquad [\bigvee_x \varphi(x)]' \equiv \bigwedge_x \varphi'(x).$$

As in the case of finite operations, the De Morgan formulas permit the definition of the universal quantifier in terms of the existential quantifier and negation (and the existential quantifier in terms of the universal quantifier and negation):

$$\text{(14)}\qquad \bigwedge_x \varphi(x) \equiv \big(\bigvee_x \varphi'(x)\big)', \quad \bigvee_x \varphi(x) \equiv \big(\bigwedge_x \varphi'(x)\big)'.$$

Remark. Instead of the symbols $\bigvee_x$ and $\bigwedge_x$ we often use the more complicated symbols $\bigvee_{\psi(x)}$ and $\bigwedge_{\psi(x)}$ where $\psi(x)$ is a given propositional function. We assume that

$$\bigvee_{\psi(x)} \varphi(x) \equiv \bigvee_x [\psi(x) \wedge \varphi(x)],$$

$$\bigwedge_{\psi(x)} \varphi(x) \equiv \bigwedge_x [\psi(x) \Rightarrow \varphi(x)].$$

## § 3. Ordered pairs

We denote a set consisting of only one element $a$ by the symbol $\{a\}$ (let us note that $\{a\} \neq a$). We denote the set consisting of the two elements $a$ and $b$ by $\{a, b\}$; similarly, $\{a, b, c\}$ denotes the set consisting of the elements $a, b$ and $c$, and so on.

Obviously the symbols $\{a, b\}$ and $\{b, a\}$ denote the same set. In the sequel we shall need the concept of an *ordered pair* with antecedent $a$ and successor $b$ which we shall denote by the symbol $\langle a, b\rangle$. We consider the pair $\langle a, b\rangle$ as distinct from the pair $\langle b, a\rangle$ (unless $a = b$); more generally, the pairs $\langle a, b\rangle$ and $\langle c, d\rangle$ are equal only when $a = c$ and $b = d$, i.e. when they have identical antecedents and identical successors:

(15) $$[\langle a, b\rangle = \langle c, d\rangle] \Rightarrow (a = c) \wedge (b = d).$$

An ordered pair can be defined in the following way:

(16) $$\langle a, b\rangle = \{\{a\}, \{a, b\}\}.$$

It is easy to verify that condition (15) is satisfied by this definition.

## § 4. Cartesian product

The *cartesian product* of the sets $X$ and $Y$ is the set of all ordered pairs $\langle x, y\rangle$ where $x \in X$ and $y \in Y$. We denote this set by $X \times Y$ and therefore

(17) $$[\langle x, y\rangle \in (X \times Y)] \equiv (x \in X) \wedge (y \in Y).$$

Cartesian products appear very frequently in mathematics. For example, the complex number plane is $\mathscr{E} \times \mathscr{E}$, where $\mathscr{E}$ is the set of all real numbers (since a complex number is an ordered pair made up of two real numbers). A cylinder can be considered as the cartesian product of the circumference of a circle (base) by a closed interval (height); the surface of a torus can be treated as the cartesian product of two circles.

Let us set down several easily proved formulas concerning the distributivity of cartesian multiplication with respect to the operations of the algebra of sets:

$$(X_1 \cup X_2)\times Y = X_1\times Y \cup X_2\times Y, \tag{18}$$

whence

$$(X_1 \cup X_2)\times(Y_1 \cup Y_2) = X_1\times Y_1 \cup X_1\times Y_2 \cup X_2\times Y_1 \cup X_2\times Y_2, \tag{19}$$

$$(X_1 - X_2)\times Y = X_1\times Y - X_2\times Y, \tag{20}$$

$$(X_1 \cap X_2)\times(Y_1 \cap Y_2) = (X_1\times Y_1) \cap (X_2\times Y_2). \tag{21}$$

If the sets $X_1, X_2, Y_1$ and $Y_2$ are nonvoid then

$$[(X_1\times Y_1) = (X_2\times Y_2)] \Rightarrow (X_1 = X_2)(Y_1 = Y_2). \tag{22}$$

All the above formulas can easily be interpreted geometrically, if we assume that $X\times Y$ is the plane with axes $X$ and $Y$ and that $X_1 \subset X, X_2 \subset X, Y_1 \subset Y, Y_2 \subset Y$.

Similarly, the following two formulas have a clear geometric interpretation:

$$A\times B = (A\times Y) \cap (X\times B), \tag{23}$$

$$(A\times B)^c = (A^c\times Y) \cup (X\times B^c), \tag{24}$$

where $A \subset X$, $B \subset Y$, $A^c$ and $B^c$ denote the complements with respect to $X$ and $Y$, respectively, and $(A\times B)^c$ denotes the complement with respect to $X\times Y$.

Formula (23) follows from (21), and (24) follows from (23) by virtue of the De Morgan rules since

$$(A\times Y) \cap (X\times B) = (A \cap X)\times(Y \cap B) = A\times B,$$

$$(A\times Y)^c = (A^c\times Y) \quad \text{and} \quad (X\times B)^c = (X\times B^c). \tag{24a}$$

## § 5. Propositional functions of two variables. Relations

Let $Z = X\times Y$. Let $\varphi(z)$ be a propositional function of the variable $z$ which ranges over the set $Z$. Since $z = \langle x, y\rangle$, the propositional function $\varphi(z)$ can be considered as a *function of two variables x and y*; we write $\varphi(x, y)$ instead of $\varphi(\langle x, y\rangle)$. We also call a propositional function of two variables a *relation*. Denoting this relation by $\varrho$ we write sometimes $x\varrho y$ instead of $\varphi(x, y)$.

Thus a propositional function of two variables ranging over the space $X$ and $Y$ is the same as a propositional function of one variable ranging over the cartesian product of these spaces. Consequently instead of $\{z: \varphi(z)\}$ we also write $\{\langle x, y\rangle: \varphi(x, y)\}$ or $\{\langle x, y\rangle: x\varrho y\}$. For example $\{\langle x, y\rangle: (x < y)\}$ is the half-plane situated above the line $x = y$, and $\{\langle x, y\rangle: (y = x^2)\}$ is the parabola $y = x^2$.

Quite often in practice we identify the relation $\varrho$ with the set $\{\langle x, y\rangle: x\varrho y\}$. This set is a subset of $X\times Y$. Conversely, every set $R \subset X\times Y$ can obviously be considered as a relation; namely

$$R = \{\langle x, y\rangle: \varphi(x, y)\} \quad \text{where} \quad \varphi(x, y) \equiv \langle x, y\rangle \in R.$$

Let us add definitions of various kinds of relations which will be frequently used later.

Definitions. A relation $\varrho$ is called *reflexive* if

$$x\varrho x \quad \text{for each} \quad x;$$

$\varrho$ is *symmetric* if

$$x\varrho y \Rightarrow y\varrho x;$$

$\varrho$ is *transitive* if

$$(x\varrho y)\,(y\varrho z) \Rightarrow (x\varrho z);$$

$\varrho$ is an *equivalence relation* if $\varrho$ is reflexive, symmetric and transitive.

Let $\varphi(x, y)$ be a given propositional function of two variables. Hence $\bigwedge_y \varphi(x, y)$ and $\bigvee_y \varphi(x, y)$ are propositional functions of one variable, namely of the variable $x$.

We set down the following easily proved formulas:

$$\bigvee_x \bigvee_y \varphi(x, y) \equiv \bigvee_y \bigvee_x \varphi(x, y), \tag{25}$$

$$\bigwedge_x \bigwedge_y \varphi(x, y) \equiv \bigwedge_y \bigwedge_x \varphi(x, y). \tag{26}$$

In both of these formulas we may alternatively write

$$\bigvee_{x,y} \varphi(x, y) \text{ or } \bigvee_z \varphi(z), \quad \text{and} \quad \bigwedge_{x,y} \varphi(x, y) \text{ or } \bigwedge_z \varphi(z).$$

These formulas express the commutativity of the operation $\bigvee$ with respect to $\bigvee$ and similarly of the operation $\bigwedge$ with respect to $\bigwedge$. On the other hand, the sequence of the quantifiers $\bigvee$ and $\bigwedge$ is significant. The following important formula holds:

$$\bigvee_x \bigwedge_y \varphi(x, y) \Rightarrow \bigwedge_y \bigvee_x \varphi(x, y). \tag{27}$$

The left-hand member denotes that there exists an $x_0$ such that, for every value of the variable $y$, $\varphi(x_0, y)$ is true; and therefore to every $y$ we can assign an $x$ (namely, $x = x_0$) such that $\varphi(x, y)$ is true; and this is exactly what the right-hand member states.

On the other hand, the implication in the opposite direction does not hold (compare formula (9)). For example, in the domain of real numbers it is true that

$$\bigwedge_y \bigvee_x (y < x),$$

but it is not true that

$$\bigvee_x \bigwedge_y (y < x).$$

Another example is: the assumption that the real valued function $f$ is bounded can be written in the following form:

$$\bigvee_y \bigwedge_x (|f(x)| < y).$$

On the other hand, the proposition $\bigwedge_x \bigvee_y (|f(x)| < y)$ is true in general (for all real valued functions), for it suffices to set $y = |f(x)|+1$.

The obvious formula

$$\bigwedge_x \varphi(x) \Rightarrow \bigvee_x \varphi(x) \tag{28}$$

(under the assumption that $X \neq \emptyset$) can be replaced, for functions of two variables, with the additional assumption that $X = Y$, by the more general formula

$$\bigwedge_{x,y} \varphi(x, y) \Rightarrow \bigwedge_x \varphi(x, x) \Rightarrow \bigvee_x \varphi(x, x) \Rightarrow \bigvee_{x,y} \varphi(x, y). \tag{29}$$

With this same assumption we can replace formula (9) by the following formula

$$\bigvee_x [\varphi(x) \wedge \psi(x)] \Rightarrow \bigvee_{x,y} [\varphi(x) \wedge \psi(y)] \tag{30}$$
$$\equiv \bigvee_x \varphi(x) \wedge \bigvee_y \psi(y) \equiv \bigvee_x \varphi(x) \wedge \bigvee_x \psi(x).$$

Analogously, (11) can be replaced by the formula

$$(31)\qquad \bigwedge_x \varphi(x) \vee \bigwedge_x \psi(x) \equiv \bigwedge_{x,y}[\varphi(x) \vee \psi(y)] \Rightarrow \bigwedge_x [\varphi(x) \vee \psi(x)].$$

## § 6. Cartesian products of $n$ sets. Propositional functions of $n$ variables

The above reasoning can easily be generalized to a larger number of variables than two. For example, Euclidean three-dimensional space is the set of ordered triples of real numbers, i.e. $\mathscr{E}\times\mathscr{E}\times\mathscr{E}$, which we write more briefly as $\mathscr{E}^3$. More generally, $\mathscr{E}^n$ denotes $n$-dimensional Euclidean space; denoting by $\mathscr{I}$ the closed interval $0 \leqslant t \leqslant 1$ we denote the $n$-dimensional unit cube by $\mathscr{I}^n$.

Similarly, we may speak about a propositional function of $n$ variables which run over the same or distinct spaces. The following examples illustrate the role of the quantifiers and the meaning of some formulas which are related to them:

EXAMPLE 1. The continuity of a function $f$ at a given point $x_0$ is expressed by the following condition (in the Cauchy formulation):

$$(32)\qquad \bigwedge_\varepsilon \bigvee_\delta \bigwedge_h (|h| < \delta) \Rightarrow (|f(x_0+h)-f(x_0)| < \varepsilon),$$

where the domain of variation of the variables $\varepsilon$ and $\delta$ is the set of positive real numbers.

Therefore the continuity of a function in the interval under consideration $a < x < b$ is expressed by prefacing formula (32) with the quantifier $\bigwedge_x$ and replacing the constant $x_0$ by the variable $x$. Since we can interchange the order of the quantifiers $\bigwedge_x$ and $\bigwedge_\varepsilon$, this condition takes on the following form:

$$(33)\qquad \bigwedge_\varepsilon \bigwedge_x \bigvee_\delta \bigwedge_h (|h| < \delta) \Rightarrow (|f(x+h)-f(x)| < \varepsilon).$$

If we interchange the order of the quantifiers $\bigwedge_x$ and $\bigvee_\delta$, we obtain a *stronger* condition, namely the condition for *uniform* continuity. Since, after this interchange, the quantifier $\bigvee_\delta$ follows $\bigwedge_\varepsilon$, but is still before $\bigwedge_x$, it is immediately clear that $\delta$ depends on $\varepsilon$ but that it does not depend on $x$ (which is exactly what "uniform" continuity means).

EXAMPLE 2. The condition that the sequence $a_1, a_2, \ldots$ be convergent to the limit $b$ can be written in the form

$$(34) \qquad \bigwedge_\varepsilon \bigvee_m \bigwedge_n |a_{m+n}-b| < \varepsilon.$$

Therefore the condition that the sequence of functions $f_1, f_2, \ldots$ be convergent to the limit $f$ is

$$(35) \qquad \bigwedge_x \bigwedge_\varepsilon \bigvee_m \bigwedge_n |f_{m+n}(x)-f(x)| < \varepsilon.$$

By interchanging the order of $\bigwedge_x$ and $\bigwedge_\varepsilon$, we obtain an equivalent condition. Let us now interchange $\bigwedge_x$ and $\bigvee_m$. We then obtain the stronger condition from which condition (35) follows, namely

$$(36) \qquad \bigwedge_\varepsilon \bigvee_m \bigwedge_x \bigwedge_n |f_{m+n}(x)-f(x)| < \varepsilon.$$

This is the condition for *uniform* convergence.

EXAMPLE 3. The rules (12) and (13) of De Morgan lead to the following rule (which can be extended to $n$ variables):

$$(37) \quad [\bigwedge_x \bigvee_y \bigwedge_z \varphi(x, y, z)]' \equiv \bigvee_x \bigwedge_y \bigvee_z [\varphi(x, y, z)]'.$$

## § 7. On the axiomatics of set theory

The four axioms given in Chapter II, § 5, are not sufficient for the discussions of Chapter III. Adding three further axioms, we obtain a system of axioms which expresses all those properties of the set concept with which we shall deal in this volume, and which—generally speaking—suffice for the applications of the theory of sets to other branches of mathematics. These are the new axioms.

V. *For every propositional function $\varphi(x)$ and for every set $A$ there exists a set consisting of those and only those elements of the set $A$ which satisfy this propositional function.*

As seen in § 1, we denote this set by the symbol

$$\{x: \varphi(x)(x \in A)\}, \text{ or, more briefly, by } \{x: \varphi(x)\},$$

where the domain of variation of $x$ is restricted to $A$.

We had examples of the applications of axiom V in § 3. The existence of the sets $\{a\}$, $\{a, b\}$, and so on (where $a \in A$, $b \in A$) follows from axiom V, since

$$\{a\} = \{x\colon\ (x = a)(x \in A)\},$$
$$\{a, b\} = \{x\colon\ [(x = a) \vee (x = b)](x \in A)\}.$$

On the other hand, the existence of an ordered pair requires the use of a further axiom.

VI. *For every set $A$ there exists a set, denoted by $2^A$, whose elements are all the subsets of the set $A$.*

VII. Axiom of choice. *For every family $\boldsymbol{R}$ of non-empty disjoint sets there exists a set which has one and only one element in common with each of the sets of the family $\boldsymbol{R}$.*

We have not applied the axiom of choice yet but we shall use it in the later chapters.

We point out that if we complete the system of axioms I–IV by means of the axioms V–VII we can at the same time omit some of the earlier axioms. In particular, axiom III follows from the rest, for the set

$$A - B = \{x\colon\ (x \in A)(x \in B)'\},$$

exists by virtue of axiom V.

Similarly, we can do without axiom II in the formation of the union of the sets $A$ and $B$ provided that we assume that both $A$ and $B$ are subsets of a fixed "space" $C$ (which is usually the case). For the existence of the set

$$A \cup B = \{x\colon\ [(x \in A) \vee (x \in B)](x \in C)\}$$

follows from axiom V.

Axiom IV is also superfluous in applications; for in its place appears the axiom which asserts that the space under consideration is a set.

Remark. Besides propositional functions $\varphi(x)$ where $x$ is supposed to be bounded by a given set, we consider sometimes propositional functions without this restriction on $x$. We suppose only—what is sufficient for our aims†—that the free variable $x$ (as well as bound variables, if they appear in $\varphi$), ranges over arbitrary sets (and, of course, individuals belonging to a given set). Then we assume a new axiom asserting the existence of

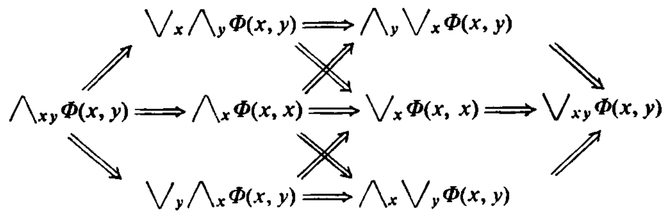
† For a more general approach, see A. P. Morse, *A Theory of Sets.*

$K_\varphi = \{x\colon \varphi(x)\}$ called a *class*† and we add the term class to the list of primitive terms of set theory.‡ A class does not need to be a set (satisfying axioms I–VII); for example, the class of all sets (compare Chapter VI, § 2, Remark 2).

Similarly, we consider propositional functions of two variables $\varphi(x, y)$, i.e. *relations (in the generalized sense)*, where $x, y$ are variables ranging over sets.

**Exercises**

**1.** Prove that none of the implications in the table below can be inverted.




**2.** Prove that

$$\bigvee_x \varphi(x) \vee \bigwedge_x \psi(x) \equiv \bigvee_x \bigwedge_y [\varphi(x) \vee \psi(y)] \equiv \bigwedge_y \bigvee_x [\varphi(x) \vee \psi(y)],$$
$$\bigvee_x \varphi(x) \wedge \bigwedge_x \psi(x) \equiv \bigvee_x \bigwedge_y [\varphi(x) \wedge \psi(y)] \equiv \bigwedge_y \bigvee_x [\varphi(x) \wedge \psi(y)].$$

**3.** Prove the following equivalences:

$$\bigwedge_x \{[\bigwedge_y \varphi(x,y)] \Rightarrow \psi(x)\} \equiv \bigwedge_x \bigvee_y [\varphi(x,y) \Rightarrow \psi(x)],$$
$$\bigwedge_x \{[\bigvee_y \varphi(x,y)] \Rightarrow \psi(x)\} \equiv \bigwedge_{x,y} [\varphi(x,y) \Rightarrow \psi(x)].$$

**4.** Write down the definition of the uniform convergence of the improper integral $\int_a^\infty f(x,y)\,dy$ making use of quantifiers.

**5.** Show that each formula (8)–(11) is a particular case of the corresponding formula (25)–(27).

*Hint.* Put $Y = \{a, b\}$, $\varphi(x, a) \equiv \alpha(x)$ and $\varphi(x, b) \equiv \beta(x)$.

**6.** Sentences $p$ can be considered as propositional functions, assuming that

$$(\bigwedge_x p) \equiv p \quad \text{and} \quad (\bigvee_x p) \equiv p.$$

Show that

$$\bigwedge_x [p \vee \varphi(x)] \equiv p \vee \bigwedge_x \varphi(x), \qquad \bigvee_x [p \wedge \varphi(x)] \equiv p \wedge \bigvee_x \varphi(x),$$
$$\bigwedge_x [p \Rightarrow \varphi(x)] \equiv [p \Rightarrow \bigwedge_x \varphi(x)], \qquad \bigwedge_x [\varphi(x) \Rightarrow p] \equiv [(\bigwedge_x \varphi(x)) \Rightarrow p].$$

† Following P. Bernays (*Axiomatic Set Theory*).

‡ See K. Gödel, *The Consistency of the Axiom of Choice and of the Generalized Continuum-hypothesis with the Axioms of Set Theory*, Princeton, 1940.

CHAPTER IV

# THE MAPPING CONCEPT. INFINITE OPERATIONS. FAMILIES OF SETS

## § 1. The mapping concept

D e f i n i t i o n. Every relation $\varrho$ such that the conditions $x\varrho y$ and $x\varrho y'$ imply $y' = y$ is called a *mapping* (or a *function*).

Under the above assumption we write $y = f(x)$ instead of $x\varrho y$ (since $y$ is uniquely determined by $x$) and denote this mapping by $f$. If $x$ ranges over the set $X$ and $y$ belongs to $Y$, we call $X$ the *domain* of $f$ and $Y$ its *range*, and we say that $f$ is a *mapping of $X$ into $Y$*. The elements of $X$ are also called the *arguments* of $f$ and the elements $f(x)$ its *values*.

Clearly, the set $G = \{\langle x, y\rangle: [y = f(x)]\}$, called the *graph* of $f$, is a subset of $X \times Y$. Its characteristic property is that for every $x \in X$ there exists one and only one $y$ such that $\langle x, y\rangle \in G$. As in the case of arbitrary relations, we may identify $f$ and $G$ (when no confusion can occur).

The set of all mappings $f$ of $X$ into $Y$ is denoted $Y^X$. Instead of $f \in Y^X$ we write also

$$f: X \to Y \quad \text{or} \quad X \xrightarrow{f} Y.$$

Clearly, in the case where $X$ and $Y$ denote sets of real numbers $G$ denotes the graph of the function $f$ in the usual sense of the word. An analogous remark applies to a function of two real variables (or a function of a complex variable).

We do not assume that the values of $f$ fill the entire set $Y$. But if this condition is fulfilled, then we say that $f$ is a *mapping of the set $X$ onto the set $Y$*.

If $X$ is the set of natural numbers, then we call $f$ an *infinite sequence*. Instead of $f(n)$ we then write $f_n$ (or more frequently $a_n$) and we call the values of $f$ *terms* of the sequence.

Remark 1. The same notation and terminology can be applied to the more general case, where $X$ and $Y$ are classes (see Chapter III, § 7, Remark). Then the graph $G$ of $f$ is a class.

Remark 2. Let $A \subset X$, $B \subset Y$ and $f\colon X \to Y$. If $f$ transforms $A$ onto $B$, we write $f\colon (X, A) \to (Y, B)$.

Definition 1. Let $f\colon X \to Y$ and $g\colon Y \to Z$. The mapping $h$ defined by the condition

$$h(x) = g[f(x)] \tag{0}$$

is called *composed* and is denoted by $h = g \circ f$ (or briefly $gf$). In other words, if $X \xrightarrow{f} Y \xrightarrow{g} Z$, then $X \xrightarrow{g \circ f} Z$.

Obviously

$$[z = (g \circ f)(x)] \equiv \bigvee_y [y = f(x)]\,[z = g(y)].$$

One can easily show that the composition of mappings is *associative*, i.e.

$$(h \circ g) \circ f = h \circ (g \circ f) \quad \text{if} \quad X \xrightarrow{f} Y \xrightarrow{g} Z \xrightarrow{h} W$$

(this means that $X^X$ is a semi-group relative to the operation $g \circ f$).

The composition of mappings can be represented by the commutative diagram

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ & {}_{h}\searrow & \downarrow{\scriptstyle g} \\ & & Z \end{array}$$

Generally speaking, the diagram

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle g} \\ T & \xrightarrow[k]{} & Z \end{array}$$

is said to be *commutative* if $g \circ f = k \circ h$, i.e. if

$$g[f(x)] = k[h(x)] \quad \text{whatever } x \in X \text{ is.}$$

Definition 2. Let $f\colon X \to Y$, $f_1\colon X_1 \to Y$ and $X \subset X_1$. If $f(x) = f_1(x)$ for each $x \in X$, we call $f_1$ an *extension* of $f$ and $f$ a *restriction* of $f_1$. We write then

$$f \subset f_1 \quad \text{and} \quad f = f_1 | X.$$

## § 2. Set-valued mappings

We shall now consider the case where the values of the mappings are sets. Thus, let $F$ be a mapping whose domain is a non-empty set $T$ and whose range is $2^X$, i.e.

$$F\colon T \to 2^X, \quad \text{hence} \quad F(t) \subset X.$$

We shall write $F_t$ instead of $F(t)$.

We introduce the following two operations on set-valued mappings, called *generalized union* and *generalized intersection* (which are analogues of the quantifiers $\bigvee_t$ and $\bigwedge_t$).

$\bigcup_t F_t$ is the set of all $x$ which belong to at least one of the sets $F_t$.

$\bigcap_t F_t$ is the set of all $x$ which belong to all the sets $F_t$.

In the notation of logic this means that

$$(1) \qquad (x \in \bigcup_t F_t) \equiv \bigvee_t (x \in F_t),$$

$$(2) \qquad (x \in \bigcap_t F_t) \equiv \bigwedge_t (x \in F_t).$$

It is to be noted that the existence of the sets $\bigcup_t F_t$ and $\bigcap_t F_t$ can be deduced from our axioms (see also § 5, Remark).

These operations are indeed generalizations of known operations of union and intersection of sets (see Chapter II, § 1). For, if the set $T$ consists of the numbers $1, 2, \ldots, n$, then

$$\bigcup_t F_t = F_1 \cup F_2 \cup \ldots \cup F_n, \qquad \bigcap_t F_t = F_1 \cap F_2 \cap \ldots \cap F_n.$$

Let us add that if $F$ is an infinite sequence of sets, i.e. if $T$ is the set of natural numbers, we use the notation $\bigcup_{n=1}^{\infty} F_n$ instead of $\bigcup_t F_t$, $\bigcap_{n=1}^{\infty} F_n$ instead of $\bigcap_t F_t$.

We now set down several formulas which can easily be proved ((4) is the generalized De Morgan formula):

$$(3) \qquad \bigcap_t F_t \subset F_t \subset \bigcup_t F_t,$$

$$(4) \qquad (\bigcup_t F_t)^c = \bigcap_t F_t^c, \quad (\bigcap_t F_t)^c = \bigcup_t F_t^c,$$

$$(5) \qquad \textit{if } F_t \subset A \textit{ for every } t, \textit{ then } \bigcup_t F_t \subset A,$$

$$(6) \qquad \textit{if } A \subset F_t \textit{ for every } t, \textit{ then } A \subset \bigcap_t F_t.$$

As an example, we shall prove formula (5). Hence, let $x \in \bigcup_t F_t$. By virtue of (1) there exists a $t_0$ such that $x \in F_{t_0}$; but by assumption $F_{t_0} \subset A$. Therefore $x \in A$. This means that $\bigcup_t F_t \subset A$.

Remark. As in Chapter III (cf. Remark, § 2) we also make use of the operations $\bigcup_{\psi(t)} F_t$ and $\bigcap_{\psi(t)} F_t$, where $\psi(t)$ is a given propositional function. The meaning of these operations is defined by formulas (1) and (2), replacing $\bigvee_t$ by $\bigvee_{\psi(t)}$ and $\bigcup_t$ by $\bigcup_{\psi(t)}$ and so on.

## § 3. The mapping $F_x = \{y: \varphi(x, y)\}$

Let $\varphi(x, y)$ be a given propositional function of two variables $x \in X$ and $y \in Y$. For fixed $x_0$, $\{y: \varphi(x_0, y)\}$ is some subset of the space $Y$. Hence, if we put

$$F_x = \{y: \varphi(x, y)\}, \tag{7}$$

we define a set-valued mapping $F$ which assigns to every element $x \in X$ a subset of the space $Y$. Let us apply the operations $\bigcup_x$ and $\bigcap_x$ to this mapping. We obtain the following formulas which display the duality between these operations and quantifiers (compare Chapter III, § 1, (2) and (3)):

$$\bigcup_x \{y: \varphi(x, y)\} = \{y: \bigvee_x \varphi(x, y)\}, \tag{8}$$

$$\bigcap_x \{y: \varphi(x, y)\} = \{y: \bigwedge_x \varphi(x, y)\}. \tag{9}$$

In fact, by formulas (1), § 2, and (1), Chapter III, § 1, we have

$$y_0 \in \bigcup_x \{y: \varphi(x, y)\} \equiv \bigvee_x [y_0 \in \{y: \varphi(x, y)\}]$$
$$\equiv \bigvee_x \varphi(x, y_0) \equiv y_0 \in \{y: \bigvee_x \varphi(x, y)\}.$$

Formula (9) is proved analogously.

The set $\{y: \bigvee_x \varphi(x, y)\}$ has the following interesting geometric interpretation.

Noting the analogy to analytic geometry, we shall say that the element $\langle x, y\rangle$ of the cartesian product $X \times Y$ has the *abscissa* $x$ and the *ordinate* $y$, and that $X$ is the axis of abscissas and $Y$ is the axis of ordinates of the space $X \times Y$. Similarly, if $A \subset X \times Y$, then the set of abscissas of the elements of the set $A$ will be called the *X-projection* of the set $A$ and the set of ordinates will be called the *Y-projection* of $A$. Now:

(10) *the set* $\{y: \bigvee_x \varphi(x, y)\}$ *is the Y-projection of the set* $\{\langle x, y\rangle: \varphi(x, y)\}$.

In fact, $y_0$ is an element of the $Y$-projection of the set $A = \{\langle x, y\rangle: \varphi(x, y)\}$ if and only if there exists an $x_0$ such that $\langle x_0, y_0\rangle \in A$, i.e. if $\varphi(x_0, y_0)$ holds; in other words, if $\bigvee_x \varphi(x, y_0)$, i.e. if $y_0 \in \{y: \bigvee_x \varphi(x, y)\}$.

The universal quantifier does not lead to such a simple geometric interpretation.

EXAMPLE. By the parametric definition of the circle $S$ with centre $\langle 0, 0\rangle$ and radius $r$ the point $\langle x, y\rangle$ belongs to this circle if there exists a $t$ such that

$$x = r\cos t, \quad y = r\sin t, \tag{11}$$

that is

$$S = \{\langle x, y\rangle: \bigvee_t (x = r\cos t)(y = r\sin t)\}.$$

This means that the formulas (11) which give the parametric definition of the circle $S$ define this circle as the projection onto a subset of the plane $X\times Y$ (i.e. the $XY$-projection) of the helix lying in the three-dimensional space $X\times Y\times T$ and defined (in an explicit manner) by the same system of equations (11).

## § 4. Images and inverse images determined by a mapping

Let $f$: $X \to Y$. Suppose $A \subset X$. We denote the *image* of the set $A$ with respect to $f$ by $f(A)$, i.e. $f(A)$ is the set of values which $f$ assumes when $x$ ranges over the set $A$; in other words,

$$[y \in f(A)] \equiv \bigvee_x (x \in A)\,(y = f(x)), \tag{12}$$

i.e.

$$f(A) = \{y: \bigvee_x (x \in A)\,(y = f(x))\}.$$

Thus $f(A)$ is the projection of the set $f|A$ into the $Y$-axis.

The *inverse image* of the set $B$ contained in $Y$ is the set $f^{-1}(B)$ consisting of all $x$ such that $f(x) \in B$; thus

$$[x \in f^{-1}(B)] \equiv [f(x) \in B], \quad \text{i.e.} \quad f^{-1}(B) = \{x: f(x) \in B\}. \tag{13}$$

(In order to avoid misunderstanding we assume that $A \notin X$ and $B \notin Y$.)

For example, for the function given by the equation $y = x^2$, the set $f^{-1}(\{1\})$ consists of two numbers: 1 and $-1$.

Let us note the following formulas:

(14) $f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$

and more generally $f(\bigcup_t F_t) = \bigcup_t f(F_t)$,

(15) $f(A_1 \cap A_2) \subset f_1(A) \cap f(A_2)$

and more generally $f(\bigcap_t F_t) \subset \bigcap_t f(F_t)$,

(15a) $f(A_1) - f(A_2) \subset f(A_1 - A_2)$,

(16) $f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$, $f^{-1}(\bigcup_t G_t) = \bigcup_t f^{-1}(G_t)$,

(17) $f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$, $f^{-1}(\bigcap_t G_t) = \bigcap_t f^{-1}(G_t)$,

(17a) $f^{-1}(B_1 - B_2) = f^{-1}(B_1) - f^{-1}(B_2)$,

(18) $ff^{-1}(B) = B$ if $B \subset f(X)$,

(19) $A \subset f^{-1}f(A)$;

if $f\colon X \to Y$, $g\colon Y \to Z$ and $h = g \circ f$, then

(20) $h^{-1}(C) = f^{-1}(g^{-1}(C))$ for each $C \subset Z$.

We shall prove, say, formula (15). Since $(\bigcap_t F_t) \subset F_t$, we have $f(\bigcap_t F_t) \subset f(F_t)$ and (15) follows by (6).

## § 5. The operations $\bigcup \boldsymbol{R}$ and $\bigcap \boldsymbol{R}$. Covers

Besides the operations $\bigcup_x$ and $\bigcap_x$ on mappings we consider the operations $\bigcup \boldsymbol{R}$ and $\bigcap \boldsymbol{R}$ on families of sets. Namely assuming that $\boldsymbol{R}$ is a family of subsets of some fixed set $A$, we denote by $\bigcup \boldsymbol{R}$ the union and by $\bigcap \boldsymbol{R}$ the intersection of all sets belonging to the family $\boldsymbol{R}$, that is

(21a) $x \in \bigcup \boldsymbol{R} \equiv \bigvee_X (x \in X \in \boldsymbol{R})$,

(21b) $x \in \bigcap \boldsymbol{R} \equiv \bigwedge_X [(X \in R) \Rightarrow (x \in X)]$.

We use the same terminology ("union" and "intersection") here as in the case where $\boldsymbol{R}$ is a family consisting of a finite number of sets: $\boldsymbol{R} = \{A_1, \ldots, A_n\}$; for

$$\bigcup \boldsymbol{R} = A_1 \cup A_2 \cup \ldots \cup A_n,$$

$$\bigcap \boldsymbol{R} = A_1 \cap A_2 \cap \ldots \cap A_n.$$

It is easy to see that

$$\bigcap \boldsymbol{R} \subset X \subset \bigcup \boldsymbol{R} \quad \text{for each} \quad X \in \boldsymbol{R},$$
$$\bigcup \emptyset = \emptyset, \quad \bigcap \emptyset = A, \quad \bigcup 2^A = A.$$

Remark. We have assumed that all members of the family $\boldsymbol{R}$ are subsets of a fixed set $A$. This assumption allows us to deduce the existence of the set $\bigcup \boldsymbol{R}$ from axiom V of Chapter III, § 7. Without this assumption the existence of the set $\bigcup \boldsymbol{R}$ would require a new axiom (called the *generalized Axiom of union*).

Definitions. If $\bigcup \boldsymbol{R} = A$, $\boldsymbol{R}$ is called a *cover* of $A$. If, moreover, the members of $\boldsymbol{R}$ are disjoint, $\boldsymbol{R}$ is called a *partition* (or a *decomposition*) of $A$.

If $\boldsymbol{R}$ and $\boldsymbol{S}$ are two covers of $A$ and to each $X \in \boldsymbol{S}$ there exists $Y \in \boldsymbol{R}$ such that $X \subset Y$, then $\boldsymbol{S}$ is called a *refinement* of $\boldsymbol{R}$.

The notions defined above have important applications in topology.

## § 6. Additive and multiplicative families of sets

We say that the family $\boldsymbol{R}$ of sets is *additive* if

$$(22) \qquad (X \in \boldsymbol{R}, Y \in \boldsymbol{R}) \Rightarrow (X \cup Y \in \boldsymbol{R}),$$

*multiplicative* if

$$(23) \qquad (X \in \boldsymbol{R}, Y \in \boldsymbol{R}) \Rightarrow (X \cap Y \in \boldsymbol{R}),$$

*subtractive* if

$$(24) \qquad (X \in \boldsymbol{R}, Y \in \boldsymbol{R}) \Rightarrow (X - Y \in \boldsymbol{R}).$$

An additive and subtractive family of sets is multiplicative since $X \cap Y = X - (X - Y)$. Clearly, the operations of union, intersection and subtraction performed on sets belonging to that family do not take us outside it (we say that this family is closed relative to these operations).

EXAMPLES. The family of finite subsets of a fixed set $A$ satisfies (22)–(24). Sets which are the unions of a finite number of closed intervals form an additive family, but they do not form a subtractive family.

THEOREM. *For every family $\boldsymbol{Z}$ of subsets of a set $A$ there exists*

(1) *a smallest additive family $\boldsymbol{R}_s$ such that $\boldsymbol{Z} \subset \boldsymbol{R}_s$,*

(2) *a smallest multiplicative family* $\boldsymbol{R}_p$ *such that* $\boldsymbol{Z} \subset \boldsymbol{R}_p$,

(3) *a smallest additive and subtractive family of sets* $\boldsymbol{R}_c$ *such that* $\boldsymbol{Z} \subset \boldsymbol{R}_c$.

P r o o f. Let us denote by $\mathscr{M}$ the totality of all additive families $\boldsymbol{R}$ which satisfy the condition $\boldsymbol{Z} \subset \boldsymbol{R}$ (consisting of subsets of the set $A$). Obviously $\mathscr{M} \neq \emptyset$, for the family of all subsets of the set $A$ is an element of the totality $\mathscr{M}$. Let

$$\boldsymbol{R}_s = \bigcap \mathscr{M}. \tag{25}$$

We shall show that the family $\boldsymbol{R}_s$ is additive and that $\boldsymbol{Z} \subset \boldsymbol{R}_s$.

Let $X \in \boldsymbol{R}_s$ and let $Y \in \boldsymbol{R}_s$. Therefore (cf. 21b)) $X \in \boldsymbol{R}$ and $Y \in \boldsymbol{R}$ for every $\boldsymbol{R} \in \mathscr{M}$. Since the families $\boldsymbol{R}$ belonging to $\mathscr{M}$ are additive, we therefore have $X \cup Y \in \boldsymbol{R}$; but since this last formula holds for every $\boldsymbol{R} \in \mathscr{M}$, hence (cf. (21b)) $X \cup Y \in \boldsymbol{R}_s$.

We shall next prove that $\boldsymbol{Z} \subset \boldsymbol{R}_s$. By assumption we have $\boldsymbol{Z} \subset \boldsymbol{R}$ for every $\boldsymbol{R} \in \mathscr{M}$. In other words, if $X \in \boldsymbol{Z}$, then $X \in \boldsymbol{R}$; and therefore $X \in \boldsymbol{R}_s$. This means that $X \in \boldsymbol{Z} \Rightarrow X \in \boldsymbol{R}_s$, i.e. that $\boldsymbol{Z} \subset \boldsymbol{R}_s$.

Finally, the family $\boldsymbol{R}_s$ is the smallest additive family containing the family $\boldsymbol{R}$, being the intersection of all the families with this property.

In order to define the family $\boldsymbol{R}_p$, we denote by $\mathscr{N}$ the totality of all multiplicative families $\boldsymbol{R}$ which satisfy the condition $\boldsymbol{Z} \subset \boldsymbol{R}$ and we set

$$\boldsymbol{R}_p = \bigcap \mathscr{N}. \tag{26}$$

The proof of the fact that the family $\boldsymbol{R}_p$ satisfies condition (2) is entirely analogous to the preceding proof.

We define the family $\boldsymbol{R}_c$ in a similar manner.

R e m a r k. Denoting by $\boldsymbol{Z}$ the family of all the one-element subsets of the set $A$, we obtain as $\boldsymbol{R}_s$ the family of all finite subsets of the set $A$.

It follows that a necessary and sufficient condition that the set $A$ be finite is that the family of all its non-empty subsets be identical with $\boldsymbol{R}_s$. This equivalence can serve as the definition of a finite set (which does not refer to the concept of natural number).

## § 7. Borel families of sets

We say that the family $\boldsymbol{R}$ of sets is *countably additive* or *countably multiplicative* if the conditions $X_n \in \boldsymbol{R}$ for $n = 1, 2, \ldots$ imply that

(27) $\bigcup_{n=1}^{\infty} X_n \in \boldsymbol{R}$, or $\bigcap_{n=1}^{\infty} X_n \in \boldsymbol{R}$, respectively.

(These concepts play an important role in the theory of probability.)

We shall encounter a rather large number of examples of families of this sort in the second part of this book; e.g. the family of closed subsets of the space of real numbers is countably multiplicative (a closed set is a set which contains all its accumulation points); the family of its complements is countably additive. Note that the family of closed sets is not only countably multiplicative, it is *absolutely* multiplicative, i.e. the intersection of an arbitrary family of closed sets is closed (see Part II, Chapter X, § 5).

A family of sets is said to be a *Borel family* if it is simultaneously countably additive and countably multiplicative, i.e. if it is closed relative to the operations $\bigcup_{n=1}^{\infty}$ and $\bigcap_{n=1}^{\infty}$.

The following theorem, analogous to the theorem of § 6, holds:

THEOREM. *To every family $\boldsymbol{Z}$ of subsets of the set $A$ there exists*

(1) *a smallest countably additive family $\boldsymbol{R}_\sigma$ such that $\boldsymbol{Z} \subset \boldsymbol{R}_\sigma$,*

(2) *a smallest countably multiplicative family $\boldsymbol{R}_\delta$ such that $\boldsymbol{Z} \subset \boldsymbol{R}_\delta$,*

(3) *a smallest Borel family $\boldsymbol{R}_\beta$ such that $\boldsymbol{Z} \subset \boldsymbol{R}_\beta$.*

In order to prove (1) let us consider the totality $\mathfrak{A}$ of all countably additive families $\boldsymbol{R}$ which satisfy the condition $\boldsymbol{Z} \subset \boldsymbol{R}$ (and consist of subsets of the set $A$) and let us set $\boldsymbol{R}_\sigma = \bigcap \mathfrak{A}$. In exactly the same way that we proved the theorem of § 6, we show that the family $\boldsymbol{R}_\sigma$ satisfies condition (1).

The families $\boldsymbol{R}_\delta$ and $\boldsymbol{R}_\beta$ are defined analogously.

Remarks. We also say that the family $\boldsymbol{R}_\beta$ is the Borel family *generated* by the family $\boldsymbol{Z}$. If $\boldsymbol{Z}$ is the family of all closed intervals then the sets belonging to $\boldsymbol{R}_\beta$ are called briefly the Borel subsets of the space of real numbers. It is worth remarking that all the sets (contained in the space of real numbers) with which we have to deal in practice are Borel sets (cf. also Chapter XI, § 1).

## § 8. Generalized cartesian products

Let $A_1, A_2, \ldots, A_n, \ldots$ be a given infinite sequence of sets. By the *cartesian product* of these sets we understand the set of all infinite sequences of the form

$$(28) \qquad a_1, a_2, \ldots, a_n, \ldots, \qquad \text{where} \qquad a_n \in A_n \text{ for every } n.$$

We denote this set by the symbol

$$(29) \qquad \prod_{n=1}^{\infty} A_n.$$

The product (29) when $A_n = \mathscr{E}$, i.e. when $A_n$ is the set of real numbers for all $n$, is especially important in applications. We denote this product by the symbol $\mathscr{E}^{\aleph_0}$; this is the natural extension of the concept of the $n$-dimensional Euclidean space $\mathscr{E}^n$ to an infinite number of dimensions.

Similarly, if $\mathscr{I}$ denotes the interval $0 \leqslant t \leqslant 1$, then $\mathscr{I}^{\aleph_0}$, called the infinite dimensional cube, is the set of all infinite sequences with terms belonging to the interval $\mathscr{I}$.

We obtain further generalizations of the concept of cartesian product by considering, instead of sequences, sets of arbitrary set-valued mappings. Thus let $F: T \to 2^X$. Then the cartesian product

$$(30) \qquad \prod_t F_t$$

is the set of all mappings $f: T \to X$ such that $f(t) \in F_t$ (where $t \in T$). Thus we have

$$(31) \qquad (f \in \prod_t F_t) \equiv \bigwedge_t [f(t) \in F_t].$$

As can be seen, when $T$ is the set of all natural numbers, then the sets (30) and (29) are identical. It can also be easily shown that if $F_t = X$ for each $t \in T$, then $\prod_t F_t = X^T$.

As in the case of a finite number of factors, we call the values $f(t)$ of $f$ the *coordinates* of the point $f \in \prod_t F_t$. Hence $f(t)$ is the $t$th coordinate of $f$, and the mapping $\pi_t$ defined by the condition

$$(32) \qquad \pi_t(f) = f(t)$$

is the *projection* of $\prod_t F_t$ into $F_t$.

If $X_t = X$ for each $t \in T$, then the mapping $\pi_t$: $X^T \to X$ defined by the formula (32) is called the *evaluation of* $X^T$ *at* $t$. The mappings $\pi_t$ define a mapping $\pi$: $T \to X^{(X^T)}$ called the *evaluation of* $X^T$.

Remark. More generally, if $\Phi \subset X^T$ we call the mapping $e_t$: $\Phi \to X$ such that

(33) $$e_t(f) = f(t) \quad \text{for each} \quad f \in \Phi$$

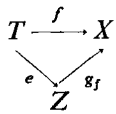
the *evaluation of* $\Phi$ *at* $t$ (we write sometimes $e(t)$ instead of $e_t$).

As before, the mappings $e_t$ define a mapping $e$: $T \to X^{\Phi}$, called the *evaluation of* $\Phi$.

Let $g_f$ be the evaluation of $X^{\Phi}$ at $f \in \Phi$. In other words

(34) $g_f(h) = h(f)$ *for* $h \in X^{\Phi}$, *and hence* $g_f$: $X^{\Phi} \to X$.

The following diagram is commutative (where $Z = X^{\Phi}$):




In other words

(35) $$f = g_f \circ e.$$

Because we have for each $t \in T$ (according to (0), (34) and (33))

$$(g_f \circ e)(t) = g_f[e(t)] = g_f(e_t) = e_t(f) = f(t).$$

**Exercises**

**1.** Prove the following formulas:

(a) $\bigcap_t (F_t \cap G_t) = \bigcap_t F_t \cap \bigcap_t G_t, \quad \bigcup_t (F_t \cup G_t) = \bigcup_t F_t \cup \bigcup_t G_t,$

(b) $\bigcap_t F_t \cup \bigcap_t G_t = \bigcap_{t,s} (F_t \cup G_s) \subset \bigcap_t (F_t \cup G_t),$

(c) $\bigcup_t (F_t \cap G_t) \subset \bigcup_{t,s} (F_t \cap G_s) = \bigcup_t F_t \cap \bigcup_t G_t,$

(d) $\bigcap_t (A \cup F_t) = A \cup \bigcap_t F_t, \quad \bigcup_t (A \cap F_t) = A \cap \bigcup_t F_t.$

Prove that the inclusion sign cannot be replaced by the identity sign in formulas (b) and (c).

**2.** Prove that if

$$A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots \quad \text{and} \quad B_1 \supset B_2 \supset \ldots \supset B_n \supset \ldots$$

then

$$\bigcap_{n=1}^{\infty} (A_n \cup B_n) = \bigcap_{n=1}^{\infty} A_n \cup \bigcap_{n=1}^{\infty} B_n.$$

**3.** Prove that

$$(\bigcup_t F_t)\times(\bigcup_t G_t) = \bigcup_{t,s}(F_t\times G_s),$$

$$(\bigcap_t F_t)\times(\bigcap_t G_t) = \bigcap_{t,s}(F_t\times G_s).$$

**4.** If $F_n \subset F_0$ for $n = 1, 2, \ldots$, then

$$F_0 = (F_0 - F_1) \cup (F_1 - F_2) \cup (F_2 - F_3) \cup \ldots \cup \bigcap_{n=0}^{\infty} F_n.$$

If $F_0 \supset F_1 \supset F_2 \supset \ldots$, then

$$(F_1 - F_2) \cup (F_3 - F_4) \cup \ldots \cup \bigcap_{n=0}^{\infty} F_n = F_0 - [(F_0 - F_1) \cup (F_2 - F_3) \cup \ldots].$$

**5.** If $\bigcap_{n=1}^{\infty} A_n \cap \bigcap_{n=1}^{\infty} B_n = \emptyset$ and $B_0 = 1$, then

$$\bigcap_{n=1}^{\infty} A_n \subset \bigcup_{n=1}^{\infty} A_n \cap (B_{n-1} - B_n).$$

**6.** We define the *least upper bound* and the *greatest lower bound* of an infinite sequence of sets $F_1, F_2, \ldots, F_n, \ldots$ as follows:

$$\operatorname{Lim\,sup} F_n = \bigcap_{n=0}^{\infty}\bigcup_{k=0}^{\infty} F_{n+k}, \qquad \operatorname{Lim\,inf} F_n = \bigcup_{n=0}^{\infty}\bigcap_{k=0}^{\infty} F_{n+k}.$$

Prove the following formulas:

(a) $$\operatorname{Lim\,inf} A_n^c = (\operatorname{Lim\,sup} A_n)^c,$$

(b) $$\operatorname{Lim\,inf}(A_n \cap B_n) = \operatorname{Lim\,inf} A_n \cap \operatorname{Lim\,inf} B_n,$$

(c) $$\operatorname{Lim\,sup}(A_n \cup B_n) = \operatorname{Lim\,sup} A_n \cup \operatorname{Lim\,sup} B_n,$$

(d) $$\bigcap_{n=1}^{\infty} A_n \subset \operatorname{Lim\,inf} A_n \subset \operatorname{Lim\,sup} A_n \subset \bigcup_{n=1}^{\infty} A_n,$$

(e) $$\operatorname{Lim\,inf} A_n \cup \operatorname{Lim\,inf} B_n \subset \operatorname{Lim\,inf}(A_n \cup B_n),$$

(f) $$\operatorname{Lim\,sup}(A_n \cap B_n) \subset \operatorname{Lim\,sup} A_n \cap \operatorname{Lim\,sup} B_n,$$

(g) $$A \dot{-} \operatorname{Lim\,inf} A_n \subset \operatorname{Lim\,sup}(A \dot{-} A_n),$$

(h) $$A \dot{-} \operatorname{Lim\,sup} A_n \subset \operatorname{Lim\,sup}(A \dot{-} A_n).$$

Show that the inclusion cannot be replaced by identity in the above formulas.

**7.** If $\operatorname{Lim\,sup} F_n = \operatorname{Lim\,inf} F_n$, then we say that the sequence $F_1, F_2, \ldots$ *converges* to the limit

$$\operatorname{Lim} F_n = \operatorname{Lim\,sup} F_n = \operatorname{Lim\,inf} F_n.$$

Prove that

(a) $$\text{if } F_1 \subset F_2 \subset \ldots, \text{ then } \bigcup_{n=1}^{\infty} F_n = \operatorname{Lim} F_n,$$

(b) $$\text{if } F_1 \supset F_2 \supset \ldots, \text{ then } \bigcap_{n=1}^{\infty} F_n = \operatorname{Lim} F_n.$$

**8.** Define the *characteristic function* $f_A$ of the set $A$ by the conditions

$$f_A(x) = \begin{cases} 1 & \text{if} \quad x \in A, \\ 0 & \text{if} \quad x \in A^c, \end{cases}$$

and prove the equivalence

$$(F = \operatorname{Lim} F_n) \equiv \left(f_F(x) = \lim f_{F_n}(x)\right).$$

**9.** Prove the formulas:

(a) $$f[A \cap f^{-1}(B)] = f(A) \cap B,$$

(b) $$\text{if } A_1 \subset A_2, \text{ then } f(A_1) \subset f(A_2),$$

(c) $$\text{if } B_1 \subset B_2, \text{ then } f^{-1}(B_1) \subset f^{-1}(B_2).$$

**10.** Let $g = f|A$ (cf. Chapter IV, § 1). Prove that

$$g^{-1}(B) = A \cap f^{-1}(B).$$

**11.** Use the axiom of choice to prove that

(a) $$\bigwedge_x \bigvee_y \varphi(x, y) \equiv \bigvee_f \bigwedge_x \varphi(x, f(x)),$$

(b) $$\bigcup_y \bigcap_x F_{x,y} \subset \bigcap_x \bigcup_y F_{x,y},$$

(c) if the conditions $x \neq x_1$ and $y \neq y_1$ imply that

$$F_{x,y} \cap F_{x_1,y_1} = \emptyset,$$

then

$$\bigcup_y \bigcap_x F_{x,y} = \bigcap_x \bigcup_y F_{x,y}.$$

**12.** If $\boldsymbol{R}$ is a family of sets, we denote by $\boldsymbol{R}_\varrho$ the family of all sets of the form $Z = X - Y$, where $X, Y \in \boldsymbol{R}$. Prove that $\boldsymbol{R}_\varrho \subset \boldsymbol{R}_{\varrho\varrho}$, and show by an example that the inverse inclusion can be false.

**13.** Prove that

$$\bigcup(\boldsymbol{R}_1 \cup \boldsymbol{R}_2) = \bigcup \boldsymbol{R}_1 \cup \bigcup \boldsymbol{R}_2, \qquad \bigcup(\boldsymbol{R}_1 \cap \boldsymbol{R}_2) \subset \bigcup \boldsymbol{R}_1 \cap \bigcup \boldsymbol{R}_2.$$

Prove that if the elements of $\boldsymbol{R}_1 \cup \boldsymbol{R}_2$ are disjoint sets, then

$$\bigcup(\boldsymbol{R}_1 \cap \boldsymbol{R}_2) = \bigcup \boldsymbol{R}_1 \cap \bigcup \boldsymbol{R}_2.$$

**14.** Prove that

$$\bigcap \boldsymbol{R}_1 \cap \bigcap \boldsymbol{R}_2 \subset \bigcap(\boldsymbol{R}_1 \cap \boldsymbol{R}_2).$$

**15.** Prove that

$$2^{A \cap B} = 2^A \cap 2^B, \qquad 2^{\bigcap_t A_t} = \bigcap_t 2^{A_t}.$$

**16.** The mapping $f$: $X \to Y$ *induces* the mapping $d$: $Z^Y \to Z^X$ defined as follows:

$$d(\varphi) = \varphi \circ f \quad \text{for} \quad \varphi \in Z^Y.$$

Let $h$: $Z^{(Z^X)} \to Z^{(Z^Y)}$ be induced by $d$. Show that if $e$ is the evaluation of $Z^X$ and $g$ the evaluation of $Z^Y$, the following diagram is commutative (where $U = Z^{(Z^X)}$ and $V = Z^{(Z^Y)}$):

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ {\scriptstyle e}\downarrow & & \downarrow{\scriptstyle g} \\ U & \xrightarrow[h]{} & V \end{array}$$

**17.** Let $f$: $X \to V$ and $g$: $Y \to W$. We call the mapping $h = (f \times g)$: $(X \times Y) \to (V \times W)$, such that

$$h(x, y) = \langle f(x), g(x) \rangle,$$

the *product-mapping*.

Show that, if $M \subset V$ and $N \subset W$, then

$$h^{-1}(M \times N) = f^{-1}(M) \times g^{-1}(N).$$

More generally, if $f_t: X_t \to V_t$ for $t \in T$, we put

$$h = \Pi_t f_t: \Pi_t X_t \to \Pi_t V_t,$$

where $[v = h(z)] \equiv \bigwedge_t [v(t) = f_t(z(t))]$ and $z \in \Pi_t X_t$.

Show that, if $M_t \subset V_t$, then

$$h^{-1}(\Pi_t M_t) = \Pi_t f_t^{-1}(M_t),$$

and if $A_t \subset X_t$, then

$$h(\Pi A_t) = \Pi_t f_t(A_t).$$

**18.** Let $f: X \to V$ and $g: X \to W$. We call the mapping $h = \langle f, g \rangle: X \to V \times W$ such that

$$h(x) = \langle f(x), g(x) \rangle$$

the *complex-mapping*.

Show that, if $M \subset V$ and $N \subset W$, then

$$h^{-1}(M \times N) = f^{-1}(M) \cap g^{-1}(N).$$

More generally, if $f_t: X \to Y_t$ for $t \in T$, let $h: X \to \Pi_t Y_t$ where $\pi_t \circ h = f_t$. Show that, if $B_t \subset Y_t$, then

$$h^{-1}(\Pi_t B_t) = \bigcap_t f_t^{-1}(B_t),$$

and if $A \subset X$, then

$$h(A) \subset \Pi_t f_t(A).$$

**19.** Let $X$ be a given set and $\varrho$ an equivalence relation defined on $X$. Thus

$$x \varrho x, \quad (x \varrho y) \Rightarrow (y \varrho x), \quad (x \varrho y)(y \varrho z) \Rightarrow (x \varrho z). \tag{36}$$

Given an element $x_0$ of $X$, the set

$$P(x_0) = \{x: x_0 \varrho x\} \tag{37}$$

is called an *equivalence set*; the family of equivalence sets is called the *quotient* $X/\varrho$. Show that the elements of $X/\varrho$ are disjoint and that $X$ is their union.

**20.** Establish the inverse theorem: if $\boldsymbol{R}$ is a partition of $X$ in non-empty (disjoint) sets, then there is an equivalence relation $\varrho$ such that

$$\boldsymbol{R} = X/\varrho. \tag{38}$$

**21.** Let $X_j$ $(j = 1, 2)$ be two sets, $\varrho_j$ two equivalence relations (defined on $X_j$) and $P_j: X_j \to X_j/\varrho_j$ two mappings defined by the formula (37) (called projections). Let $f: X_1 \to X_2$ and suppose that

$$x' \varrho_1 x'' \Rightarrow f(x') \varrho_2 f(x''). \tag{39}$$

Show that there is a mapping $F\colon X_1/\varrho_1 \to X_2/\varrho_2$ such that the following diagram is commutative

$$\begin{array}{ccc} X_1 & \xrightarrow{f} & X_2 \\ {\scriptstyle p_1}\downarrow & & \downarrow{\scriptstyle p_2} \\ X_1/\varrho_1 & \xrightarrow[F]{} & X_2/\varrho_2 \end{array}$$

Moreover if $f$ is onto, so is $F$.

**22.** As in Exercise 16, let

(40) $$d_f(h) = h \circ f \quad \text{for} \quad f\colon X \to Y \text{ and } h\colon Y \to Z.$$

Show that

(41) $$d_{g \circ f} = d_f \circ d_g \quad \text{if} \quad f\colon X \to Y \text{ and } g\colon Y \to T.$$

**23.** Let $f_j\colon X_j \to Y_j$ and $g_j\colon Y_j \to Z_j$ for $j = 0, 1$. Then the following *distributivity* holds (for the definition of a product mapping, see Exercise 17):

(42) $$(g_0 \times g_1) \circ (f_0 \times f_1) = (g_0 \circ f_0) \times (g_1 \circ f_1).$$

More generally, if $f_t\colon X_t \to Y_t$ and $g_t\colon Y_t \to Z_t$ for $t \in T$, then

(43) $$(\prod_t g_t) \circ (\prod_t f_t) = \prod (g_t \circ f_t) \quad \text{(Bourbaki)}.$$

CHAPTER V

# THE CONCEPT OF THE POWER OF A SET. COUNTABLE SETS

## § 1. One-to-one mappings

The mapping $f$: $X \to Y$ is said to be *one-to-one* (more concisely: 1-1) if

(1) $$(x_1 \neq x_2) \Rightarrow [f(x_1) \neq f(x_2)]$$

or, equivalently, if

(2) $$[f(x_1) = f(x_2)] \Rightarrow (x_1 = x_2).$$

For example, the function $x^3$ is one-to-one (in the domain of reals) but the function $x^2$ is not.

If the mapping $f$ is onto, then $f$ is one-to-one if it forms a set of pairs $\langle x, y\rangle$ such that every element $x \in X$ is the predecessor and every $y \in Y$ is the successor of one and only one of these pairs.

Still another way of stating this is: $f$ is one-to-one if for every $y \in f(X)$ the set $f^{-1}(y)$ reduces to one element $x$ (such that $y = f(x)$). In this case we usually use the symbol $f^{-1}(y)$ to denote $x$ (and not the set $\{x\}$) and we call $f^{-1}$ the *inverse mapping of* $f$; $Y$ is its domain and $X$ is its range.

Obviously

(3) $$[y = f(x)] \equiv [x = f^{-1}(y)].$$

THEOREM 1. *The inverse of a one-to-one mapping is one-to-one.*

For

(4) $$(f^{-1})^{-1} = f.$$

Geometrically, the transition to the inverse function can be interpreted (in the case where $X$ and $Y$ each denote the set of real numbers) as the reflection of the graph of the function with respect to the line $y = x$.

THEOREM 2. *The composition of two one-to-one mappings is a one-to-one mapping.*

In other words, if $f$ is a one-to-one mapping of the set $X$ onto the set $Y$ and $g$ is a one-to-one mapping of the set $Y$ onto the set $Z$, then the mapping $h = g \circ f$ is a one-to-one mapping of $X$ onto $Z$.

For, if $h(x_1) = h(x_2)$, then $g(f(x_1)) = g(f(x_2))$, whence $f(x_1) = f(x_2)$, and consequently $x_1 = x_2$.

Under the assumption that $f$ is one-to-one, formulas (15) and (19) (Chapter IV, § 4) may be strengthened: they can be replaced by the formulas

$$(5) \qquad f(A_1 \cap A_2) = f(A_1) \cap (A_2)$$

and more generally $f(\bigcap_t F_t = \bigcap_t f(F_t)$,

$$(6) \qquad A = f^{-1}f(A).$$

If $f$ is one-to-one, we have besides the equivalence (13) of Chapter IV, § 4, the symmetric equivalence

$$(7) \qquad [x \in A] \equiv [f(x) \in f(A)].$$

First, we shall establish formula (7). If $f(x) \in f(A)$, there exists $x_1 \in A$ such that $f(x) = f(x_1)$, which implies by (2) that $x_1 = x$ and hence $x \in A$.

Conversely, $x \in A$ implies that $f(x) \in f(A)$.

Formulas (5) and (6) can be established as follows (referring to formulas (2) and (13) of Chapter IV and to the equivalence (7)):

$$f(x) \in \bigcap_t f(A_t) \equiv \bigwedge_t f(x) \in f(A_t) \equiv \bigwedge_t x \in A_t$$

$$\equiv x \in \bigcap_t A_t = f(x) \in f(\bigcap_t A_t),$$

$$x \in f^{-1}f(A) \equiv f(x) \in f(A) \equiv x \in A.$$

Remark. As in the Remark of Chapter IV, § 8, let $\Phi \subset X^T$, let $e: T \to X^\Phi$ be the evaluation of $\Phi$, and $g_f: X^\Phi \to X$ the evaluation of $X^\Phi$ at $f$. Let $e$ be one-to-one. Then by formula (35) of Chapter IV, we have

$$f \circ e^{-1} \subset g_f, \quad \text{i.e.} \quad f(e^{-1}(h)) = g_f(h) \quad \text{for} \quad h \in e(T).$$

## § 2. Power of a set

Definition. Two sets $X$ and $Y$ are said to be *equipollent*, or to *have the same power*, symbolically

$$X \sim Y,$$

if there exists a one-to-one mapping of $X$ onto $Y$.

If the set $X$ is finite: $X = (a_1, \ldots, a_n)$, then the set $Y$ has the same power as $X$ if and only if it has the same number $n$ of elements. The concept of equipollent sets therefore coincides, in the case of finite sets, with the elementary concept of having the same number of elements; this concept can however be applied also to infinite sets.

For example, the set of all odd natural numbers has the same power as the set of all even natural numbers; in fact, the function $f(n) = n+1$ establishes a one-to-one mapping of the set $(1, 3, 5, \ldots)$ onto the set $(2, 4, 6, \ldots)$.

Similarly, the set of all natural numbers is of the same power as the set of all even numbers (which shows that an infinite set can have the same power as a proper subset of itself!). Here the corresponding function is $f(n) = 2n$.

Two intervals $a < x < b$ and $c < x < d$ are of equal power, as is easily shown using a linear mapping. The open interval $-\pi/2 < x < +\pi/2$ has the same power as the set of all real numbers; the corresponding mapping is $y = \tan x$.

Next, we shall show that the set of all natural numbers does not have the same power as the set of all real numbers; it will follow from this that, in the domain of infinite sets, there exist sets of different powers, and—as we shall show—there even exists an infinite number of infinite sets of which no two have the same power.

THEOREM 3. *The relation $X \sim Y$ is an equivalence relation, i.e.*

(8) $$X \sim X,$$

(9) $$(X \sim Y) \Rightarrow (Y \sim X),$$

(10) $$(X \sim Y)(Y \sim Z) \Rightarrow (X \sim Z).$$

Proof. Formula (8) follows from the fact that the identity, i.e. the function $f(x) = x$, is a one-to-one mapping of the set

$X$ onto itself. Formulas (9) and (10) follow from Theorems 1 and 2, respectively.

Theorem 3 permits the classification of sets with respect to their "power". This leads to the extension to infinite sets of the elementary concept of the number of elements in a set. Namely, to each set $X$ we assign a *cardinal number*, or its power, which we denote by the symbol $\overline{\overline{X}}$, in such a way that the same cardinal number is assigned to two distinct sets if and only if these sets have the same power.

The cardinal number of a finite set is the number of its elements.

Remark. Cardinal numbers play an auxiliary role in the theory of sets, inasmuch as all the theorems of set theory can be formulated without using them. However, many theorems gain in lucidity when expressed in terms of cardinal numbers.

From the axiomatic point of view the introduction of cardinal numbers requires a new axiom, namely the axiom of their existence.

One can also—using the term "class"—*define* $\overline{\overline{X}}$ to be the *class* of all sets equipollent to $X$ (cf. Chapter III, § 7).

## § 3. Countable sets

A set $A$ is said to be *infinitely countable* if it has the same power as the set of all natural numbers; in other words, if its elements can be arranged in an infinite sequence of distinct terms.

Finite sets are called countable sets as well.

Hence *a nonvoid set is countable if its elements can be arranged in an infinite sequence* (which may have repetitions). For, if the infinite sequence contains an infinite number of distinct terms, then there exists a subsequence which contains each of these terms precisely once.

As we saw above, the set of even natural numbers (and similarly the set of odd natural numbers) is countable.

THEOREM 1. *The set of all real numbers is noncountable.*

To prove this theorem it obviously suffices to show that for every sequence of real numbers $a_1, a_2, \ldots, a_n, \ldots$ we can define a real number $c$ which does not belong to this sequence.

To this end, we define a sequence of closed intervals $p_1q_1$, $p_2q_2, \ldots, p_nq_n, \ldots$ which are such that

$$q_n - p_n = 1/3^n, \quad p_nq_n \subset p_{n-1}q_{n-1}, \quad a_n \notin p_nq_n.$$

Thus, in the closed interval (0, 1) we determine a closed interval $p_1q_1$ which does not contain the point $a_1$ [this will be one of the three intervals (0, 1/3) or (1/3, 2/3) or (2/3, 1)]. Similarly, in the interval $p_1q_1$ we determine a closed interval $p_2q_2$ of length 1/9 which does not contain the point $a_2$. In general, in the closed interval $p_{n-1}q_{n-1}$ we determine a closed interval $p_nq_n$ of length $1/3^n$ which does not contain the point $a_n$.

Let $c$ be the common point of all the closed intervals $p_nq_n$:

$$\{c\} = \bigcap_{n=1}^{\infty} p_nq_n, \qquad \text{i.e.} \qquad c = \lim_{n\to\infty} p_n = \lim_{n\to\infty} q_n.$$

Obviously, $c \neq a_n$ for every $n$ since $a_n \notin p_nq_n$ whereas $c \in p_nq_n$.

We shall now list several important properties of countable sets.

THEOREM 2. *The union $A \cup B$ of two countable sets $A$ and $B$ is countable.*

In fact, under the assumption that the elements of the set $A$ can be written in the form of an infinite sequence $a_1, a_2, \ldots, a_n, \ldots$, and the elements of the set $B$ in the form of a sequence $b_1, b_2, \ldots, b_n, \ldots$, we consider the sequence

(11) $$a_1, b_1, a_2, b_2, \ldots, a_n, b_n, \ldots$$

The terms of this sequence obviously form the set $A \cup B$.

It follows from this that *the set of all integers is countable*. For the set of all positive integers as well as the set of all nonpositive integers is countable.

THEOREM 3. *The cartesian product of two (or, more generally, of a finite number) of countable sets is a countable set.*

P r o o f. We shall prove that the set of pairs $\langle m, n\rangle$, where $m$ and $n$ are natural numbers, is countable. Hence we have to represent this set as a sequence. To this end, we adopt the following rule: of two pairs $\langle m, n\rangle$ and $\langle m', n'\rangle$ we consider that one to be the earlier whose sum of elements is smaller; but if $m+n = m'+n'$, then the earlier pair is the one with the smaller antecedent. And therefore this sequence can be represented as follows:

(12) $$\langle 1, 1\rangle, \langle 1, 2\rangle, \langle 2, 1\rangle, \langle 1, 3\rangle, \langle 2, 2\rangle, \langle 3, 1\rangle, \ldots$$

From this we easily deduce, that given two arbitrary infinite sequences $a_1, a_2, \ldots, a_m, \ldots$ and $b_1, b_2, \ldots, b_n, \ldots$, we can write the sequence of all pairs $\langle a_m, b_n \rangle$ in the form of an infinite sequence.

The generalization from two to an arbitrary finite number of countable sets presents no trouble.

If follows from Theorem 3 that *the set* $\mathfrak{R}$ *of all rational numbers is countable.*

For, every positive rational number can be represented as a pair of numbers $p/q$ (in the irreducible form), i.e. the set of positive rational numbers can be represented as a subsequence of the sequence (12). The set of positive rational numbers is therefore countable. The same is true of the set of negative rational numbers together with the number zero. Therefore, according to Theorem 2, the set of all rational numbers is countable.

From Theorem 3 it also follows that every double sequence $\{a_{mn}\}$ can be transformed into a simple sequence, i.e. it is possible to write down the elements of the array

$$\begin{array}{l} a_{11}, a_{12}, \ldots, a_{1n}, \ldots \\ a_{21}, a_{22}, \ldots, a_{2n}, \ldots \\ \cdots\cdots\cdots\cdots\cdots \\ a_{m1}, a_{m2}, \ldots, a_{mn}, \ldots \\ \cdots\cdots\cdots\cdots\cdots \end{array} \tag{13}$$

in the form of the infinite sequence

$$a_{11}, a_{12}, a_{21}, a_{13}, \ldots \tag{14}$$

From this we deduce the following theorem:

THEOREM 4. *The union* $S = A_1 \cup A_2 \cup \ldots \cup A_m \cup \ldots$ *of a countable sequence of countable sets is countable.*

P r o o f. We write each of the sets $A_m$ in the form of a sequence $a_{m1}, a_{m2}, \ldots, a_{mn}, \ldots$, and then we transform the double sequence $\{a_{mn}\}$ into the simple sequence (14) (perhaps with repetitions). [Here we apply the axiom of choice (Chapter III, § 7), for the set of sequences consisting of the elements of the set $A_m$ contains more than one element (and none of them in general can be distinguished).]

THEOREM 5. *The set of all finite sequences with terms belonging to a given countable set is countable.*

For this set can be represented in the form of a union $A_1 \cup A_2 \cup \dots \cup A_m \cup \dots$, where $A_m$ is the set of sequences with $m$ elements. And the countability of the set $A_m$ follows from Theorem 3.

From this we deduce that *the set of all polynomials with rational coefficients is countable.*

For every polynomial is determined by its coefficients, i.e. the polynomial $a_0+a_1x+ \dots +a_mx^m$ is determined by the sequence consisting of $m+1$ rational numbers $a_0, a_1, \dots, a_m$.

COROLLARY. *The set of all algebraic numbers is countable.*

In fact, the set of all polynomials with rational coefficients is countable and hence we can write it in the form of an infinite sequence $w_1, w_2, \dots, w_m, \dots$ Let $A_m$ denote the set of roots of the equation $w_m(x) = 0$; this set, as is known, is finite (the number of its elements does not exceed the degree of the polynomial $w_m$). By virtue of Theorem 4, the set $A_1 \cup A_2 \cup \dots \cup A_m \cup \dots$, i.e. the set of all algebraic numbers is therefore countable.

Remark. This last result together with Theorem 1 leads to the result that transcendental (i.e. nonalgebraic) numbers exist, and even that there is a noncountable number of them (for the union of two countable sets is countable). Making use of the methods given here, one could even define a transcendental number; namely, to this end we set down all real algebraic numbers in the form of an infinite sequence and then apply the method used in the proof of Theorem 1, which determines a real number not belonging to this sequence.

We recall that the numbers $e$ and $\pi$ are proved to be transcendental numbers—by entirely different means.

**Exercises**

1. Consider the transformation of the plane into itself given by the system of equations

$$x = au+bv, \quad y = cu+dv.$$

Give the conditions on the coefficients $a$, $b$, $c$, $d$ under which this transformation is one-to-one.

2. Is the homographic transformation of the Gaussian plane (i.e. the plane of complex numbers together with the point at infinity)

$$w = (az+b)/(cz+d)$$

one-to-one?

**3.** Suppose $u_1, u_2, \ldots, u_n, \ldots$ is a given sequence of real numbers. Let

$$u_n = c_{n0} \cdot c_{n1} c_{n2} c_{n3} \ldots$$

be the decimal expansion of the number $u_n$ containing an infinite number of digits different from 9.

We define the number $l = 0 \cdot e_1 e_2 e_3 \ldots$ in the following way: $e_n = 0$ if $c_{nn} \neq 0$, $e_n = 1$ if $c_{nn} = 0$. Prove that the number $l$ is not a term of the sequence $u_1, u_2, \ldots$, and deduce from this Theorem 1, § 3.

**4.** Prove that the set of all intervals (in the set of real numbers) with both endpoints rational is countable.

**5.** We say that a function $f$ (with real arguments and values) has a proper maximum at the point $a$ if there exists an interval $bc$ containing the point $a$ in its interior such that the conditions $b < x < c$ and $x \neq a$ imply the inequality $f(x) < f(a)$. Prove that the set of proper maxima of the function $f$ is countable.

*Hint*: Give the points $b$ and $c$ rational values.

**6.** Prove that every family of disjoint intervals is countable.

*Hint*: Make use of the countability of the set of rational numbers.

**7.** Prove that the set of points of discontinuity of a monotonic function is countable.

*Hint*: A monotonic function has at every point a left and right limit (which are different at points of discontinuity). Then make use of Exercise 6.

**8.** Prove that the set of spheres (in 3-dimensional space) which have both rational radii and rational coordinates of the centre is countable.

**9.** With the notations of Exercise 21 of Chapter IV, show that, if

$$x' \varrho_1 x'' = f(x') \varrho_2 f(x''),$$

then the mapping $F$ is one-to-one.

**10.** Establish the following formulas (used in the next chapter):

(a) $$(A \times B) \sim (B \times A),$$

(b) $$[A \times (B \times C)] \sim [(A \times B) \times C],$$

(c) $$(A_1 \sim B_1)(A_2 \sim B_2) \Rightarrow [(A_1 \times A_2) \sim (B_1 \times B_2)],$$

(d) $$(A \sim B) \Rightarrow (2^A \sim 2^B),$$

(e) $$(A \times \{a\}) \sim A,$$

(f) $$[(A_1 \sim B_1)(A_2 \sim B_2)(A_1 \cap A_2 = 0 = B_1 \cap B_2)] \Rightarrow [(A_1 \cup A_2) \sim (B_1 \cup B_2)].$$

CHAPTER VI

# OPERATIONS ON CARDINAL NUMBERS. THE NUMBERS $\mathfrak{a}$ AND $\mathfrak{c}$

We denote the power of the set of natural numbers by $\mathfrak{a}$ (or $\aleph_0$) and the power of the set of real numbers (the power of the "continuum") by $\mathfrak{c}$.

The numbers $\mathfrak{a}$ and $\mathfrak{c}$ are the most important of the infinite cardinal numbers which occur in analysis and geometry. So far, we know (Chapter V, § 3, Theorem 1) that

$$\mathfrak{a} \neq \mathfrak{c}. \tag{1}$$

The operations on arbitrary cardinal numbers which we shall now define will interest us primarily in relation to the numbers $\mathfrak{a}$ and $\mathfrak{c}$.

## § 1. Addition and multiplication

The *sum* $\mathfrak{m}+\mathfrak{n}$ of two cardinal numbers $\mathfrak{m}, \mathfrak{n}$ is defined to be the power of the union of two disjoint sets which have the powers $\mathfrak{m}$ and $\mathfrak{n}$ respectively.

We therefore have

$$\overline{\overline{X}}+\overline{\overline{Y}} = \overline{\overline{X \cup Y}}, \quad \text{if} \quad X \cap Y = \emptyset. \tag{2}$$

We note that for every pair of sets $X$ and $Y$ there exists a pair of disjoint sets $X_1$ and $Y_1$ such that $\overline{\overline{X_1}} = \overline{\overline{X}}$ and $\overline{\overline{Y_1}} = \overline{\overline{Y}}$. For, denoting any two distinct elements by $a$ and $b$, it suffices to set $X_1 = \{a\}\times X$ and $Y_1 = \{b\}\times Y$.

Keeping this remark in mind, we can assert that for every two cardinal numbers their sum is defined uniquely (i.e. independently of the choice of the sets $X$ and $Y$, compare Exercise 10(f) of Chapter V).

We define the *product* $\mathfrak{m}\cdot\mathfrak{n}$ of $\mathfrak{m}$ and $\mathfrak{n}$ to be the power of the cartesian product of two sets having powers $\mathfrak{m}$ and $\mathfrak{n}$ respectively, i.e.

$$\overline{\overline{X}}\cdot\overline{\overline{Y}} = \overline{\overline{X\times Y}}. \tag{3}$$

Thus, the product of cardinal numbers is uniquely defined (comp. Exercise 10(c) of Chapter V).

It can easily be verified that the above definitions, in the case where $\mathfrak{m}$ and $\mathfrak{n}$ denote natural numbers, are in agreement with the usual definitions of addition and multiplication in arithmetic. We deduce from Theorems 2 and 3 (Chapter V, § 3) that

$$\mathfrak{a}+\mathfrak{a} = \mathfrak{a}, \quad \mathfrak{a}\cdot\mathfrak{a} = \mathfrak{a}, \quad \mathfrak{a}+n = \mathfrak{a}, \quad \mathfrak{a}\cdot n = \mathfrak{a}, \tag{4}$$

where $n$ is a natural number.

Multiplication and addition satisfy the associative and commutative laws. The distributive law is also satisfied:

$$\mathfrak{m}\cdot(\mathfrak{n}+\mathfrak{p}) = \mathfrak{m}\cdot\mathfrak{n}+\mathfrak{m}\cdot\mathfrak{p}. \tag{5}$$

For, let $\mathfrak{m} = \overline{\overline{X}}$, $\mathfrak{n} = \overline{\overline{Y}}$ and $\mathfrak{p} = \overline{\overline{Z}}$ where $Y\cap Z = \emptyset$. Then (cf. Chapter III, § 4 (18) and (21)):

$$X\times(Y\cup Z) = X\times Y\cup X\times Z,$$

$$(X\times Y)\cap(X\times Z) = X\times(Y\cap Z) = \emptyset,$$

and therefore $\overline{\overline{X\times(Y\cup Z)}} = \overline{\overline{X\times Y}}+\overline{\overline{X\times Z}}$, which was to be proved.

It follows from this (by induction) that

$$\mathfrak{m}\cdot n = \mathfrak{m}+\mathfrak{m}+\ldots+\mathfrak{m}, \tag{6}$$

where the right member has $n$ terms.

For formula (6) is obvious for $n = 1$, and by virtue of (5):

$$\mathfrak{m}\cdot(n+1) = \mathfrak{m}\cdot n+\mathfrak{m}\cdot 1 = \mathfrak{m}\cdot n+\mathfrak{m}.$$

Equation (6) asserts that $\mathfrak{m}\cdot n$ is the power of the union of $n$ disjoint sets each of which is of power $\mathfrak{m}$. This theorem can be generalized to the sum of an infinite number of terms as follows.

*Let* $\overline{\overline{T}} = \mathfrak{n}$ *and let* $F\colon T\to 2^X$ *be a set-valued mapping such that*

$$\overline{\overline{F_t}} = \mathfrak{m}, \quad F_t\cap F_{t'} = \emptyset \quad \textit{for} \quad t\neq t', \tag{7}$$

*then*

$$\overline{\overline{\bigcup_t F_t}} = \mathfrak{m}\cdot\mathfrak{n}. \tag{8}$$

P r o o f. Let $t_0$ be a fixed element of $T$ and let $g_t$ be a one-to-one mapping of $F_{t_0}$ onto $F_t$ (we apply the axiom of choice here). Let us set

$$f(x, t) = g_t(x), \quad \text{where} \quad x \in F_{t_0} \text{ and } t \in T. \tag{9}$$

$f$ is a one-to-one mapping of the cartesian product $F_{t_0} \times T$ onto the union $\bigcup_t F_t$. For let

$$f(x, t) = f(x', t'), \quad \text{i.e.} \quad g_t(x) = g_{t'}(x'). \tag{10}$$

If $t \neq t'$, then $g_t(x) \neq g_{t'}(x')$, since $g_t(x) \in F_t$, $g_{t'}(x') \in F_{t'}$, and $F_t \cap F_{t'} = \emptyset$.

Thus $t = t'$. If $x \neq x'$, then $g_t(x) \neq g_t(x')$, because $g_t$ is one-to-one.

Therefore, (10) implies that $t = t'$ and $x = x'$.

We have thus proved that the sets $F_{t_0} \times T$ and $\bigcup_t F_t$ have the same power. This completes the proof of formula (8).

## § 2. Exponentiation

Let $\overline{\overline{X}} = \mathfrak{m}$ and $\overline{\overline{Y}} = \mathfrak{n}$. The cardinal number $\mathfrak{n}^{\mathfrak{m}}$ is defined to be the power of the set, denoted by $Y^X$, of all mappings $f$: $X \to Y$, i.e.

$$\overline{\overline{Y}}^{\overline{\overline{X}}} = \overline{\overline{Y^X}}.$$

The following formulas, known from the arithmetic of natural numbers, are valid:

$$\mathfrak{n}^{\mathfrak{m}+\mathfrak{p}} = \mathfrak{n}^{\mathfrak{m}} \cdot \mathfrak{n}^{\mathfrak{p}}, \tag{11}$$

$$(\mathfrak{m}\mathfrak{n})^{\mathfrak{p}} = \mathfrak{m}^{\mathfrak{p}} \cdot \mathfrak{n}^{\mathfrak{p}}, \tag{12}$$

$$(\mathfrak{n}^{\mathfrak{m}})^{\mathfrak{p}} = \mathfrak{n}^{\mathfrak{m}\mathfrak{p}}. \tag{13}$$

P r o o f. Let $\mathfrak{m} = \overline{\overline{X}}$, $\mathfrak{n} = \overline{\overline{Y}}$ and $\mathfrak{p} = \overline{\overline{T}}$.

In order to prove formula (11), we must prove that

$$Y^{X \cup T} \sim Y^X \times Y^T \quad \text{provided} \quad X \cap T = \emptyset. \tag{14}$$

Hence, let $f \in Y^{X \cup T}$. Assign to $f$ the pair $\langle f|X, f|T \rangle$. This correspondence, as can easily be verified, establishes a one-to-one correspondence between the elements of the sets $Y^{X \cup T}$ and $Y^X \times Y^T$. Thus formula (14) is proved.

Formula (12) means that

$$(X\times Y)^T \sim X^T\times Y^T. \tag{15}$$

Let $f\in(X\times Y)^T$. Hence $f$ is a complex mapping, i.e. its values are ordered pairs belonging to $X\times Y$; we can therefore write

$$f(t)=\langle g(t),h(t)\rangle, \quad \text{where} \quad g(t)\in X \text{ and } h(t)\in Y.$$

And therefore $g\in X^T$ and $h\in Y^T$. We have thus assigned to $f$ a pair $\langle g,h\rangle$, i.e. an element of the set $X^T\times Y^T$. It is easy to verify that this correspondence is one-to-one. This yields (15).

In order to prove (13) we have to show that

$$(Y^X)^T \sim Y^{X\times T}. \tag{16}$$

Hence let $f\in Y^{X\times T}$. $f$ assigns to every pair $\langle x,t\rangle$ the element $f(x,t)$ of the set $Y$. For a fixed $t$ we obtain a function $g_t$ of the variable $x$ defined by means of the formula

$$g_t(x)=f(x,t), \tag{17}$$

i.e. $g_t\in Y^X$, for every value of the variable $t$. We have thus defined a mapping—let us denote it by $g$—which assigns to elements of $T$ elements of $Y^X$, i.e. $g\in(Y^X)^T$.

To every $f$ belonging to $Y^{X\times T}$ we have therefore assigned some $g$ belonging to $(Y^X)^T$. It is easy to prove that this correspondence is one-to-one.

Let us now consider certain particular cases.

It is almost obvious that

$$\mathfrak{n}^1=\mathfrak{n}$$

(in this case the domain reduces to a single element).

Let $m$ be a natural number. By (11) we have

$$\mathfrak{n}^{m+1}=\mathfrak{n}^m\cdot\mathfrak{n}^1=\mathfrak{n}^m\cdot\mathfrak{n}.$$

And therefore (by induction)

$$\mathfrak{n}^m=\mathfrak{n}\cdot\mathfrak{n}\cdot\ldots\cdot\mathfrak{n}, \tag{18}$$

where the right-hand member has $m$ factors.

It also follows that the definition of exponentiation of cardinal numbers which we assumed coincides with the arithmetic definition when these numbers are finite ($\mathfrak{m}=m$, $\mathfrak{n}=n$).

Let us now assume that $\mathfrak{n} = 2$. Hence let $\overline{\overline{X}} = \mathfrak{m}$, and $Y = \{0, 1\}$ (i.e. $Y$ is the set consisting of two numbers: 0 and 1). Hence the set $Y^X$ is the set of functions defined on the set $X$ and assuming only two values 0 and 1 (or only one of them). We call such functions *characteristic* functions (see Chapter IV, Exercise 8); namely, the function satisfying the condition

$$(19) \qquad f(x) = \begin{cases} 1 & \text{for} \quad x \in A, \\ 0 & \text{for} \quad x \in X - A \end{cases}$$

is the characteristic function of the set $A$.

*The set* $\{0, 1\}^X$ *and the set of all subsets of the set* $X$ *are of equal power, namely of the power* $2^{\mathfrak{m}}$, *where* $\mathfrak{m} = \overline{\overline{X}}$.

Proof. Assign to the set $A \subset X$ its characteristic function $f_A$. This correspondence is one-to-one. For let $A \neq B$ and let $a \in A - B$. Hence we have $f_A(a) = 1$ but $f_B(a) = 0$ and therefore $f_A \neq f_B$. Here every characteristic function has been assigned to some subset of the set $X$.

CANTOR THEOREM. $2^{\mathfrak{m}} \neq \mathfrak{m}$; in other words, *no set* $X$ *has power equal to that of the family of all its subsets.*

Proof. It suffices to show that if $F\colon X \to 2^X$, $F$ is not an onto mapping, i.e. that there exists a set $Z \subset X$ which is not a value of $F$. (This is the so-called *diagonal theorem.*) The Cantor theorem will follow because if the set $X$ were of power equal to that of the family of all its subsets, then there would exist a (one-to-one) mapping $F\colon X \to 2^X$ onto.

Define the set $Z$ as follows:

$$(20) \qquad Z = \{x\colon\ x \notin F(x)\}.$$

We have to show that $Z \neq F(x)$ for every $x \in X$. Let us assume on the contrary that $Z = F(x_0)$. By virtue of (20) the following equivalence holds:

$$(x \in Z) \equiv [x \notin F(x)].$$

Setting $x = x_0$ in this equivalence, we obtain

$$(x_0 \in Z) \equiv [x_0 \notin F(x_0)],$$

and therefore $Z \neq F(x_0)$, which is a contradiction.

Remarks. 1. The diagonal theorem can be illustrated geometrically as follows. Let $X$ be the closed interval $0 \leqslant x \leqslant 1$. We place the set $F(x)$, which by assumption is a subset of this interval, on the vertical line passing through the point $x$. In this way we obtain a planar set $M = \{\langle x, y\rangle : y \in F(x)\}$ contained in the square $X \times X$. Let $P$ denote the diagonal of this square. Thus, the set $Z$ is the projection of the set $P-M$ onto the $X$-axis.

2. The proof of the Cantor theorem given above permits us to verify easily that the family of all subsets of the set $X$ is not of the same power as that of any of the subsets of this set.

It follows immediately that *there does not exist a set of all sets* (for the family of its subsets would itself be one of its subsets).

This same conclusion follows, after all, immediately from the theorem on the diagonal. For, if there existed a set $X$ whose elements were all sets, then the mapping $F$ defined by the condition $F(x) = x$ (i.e. the identity) would obviously assume as values all subsets of the set $X$ (since these subsets would be elements of the set $X$).

Let us add that from the (false) assumption that there exists the set of all sets there follows the existence of

$$Z = \{x:\ (x \notin x)\}.$$

However, the existence of the set $Z$ leads immediately to a contradiction (called the *Russell antinomy*) because $x \in Z \equiv x \notin x$, and therefore $Z \in Z \equiv Z \notin Z$.

The theorem on the non-existence of the set of all sets was deduced by us from the axioms given in Chapter III, § 7. The assumption, that for a given set $A$, the propositional function $\varphi(x)$ (with unbounded domain of variation for $x$) determines the set $\{x:\ \varphi(x)\,(x \in A)\}$ plays an essential role in the formulation of axiom V. Omitting the expression $x \in A$ would lead to a contradiction. For, taking as $\varphi(x)$ the propositional function "$x$ is a set", we should obtain as an immediate consequence the existence of the set of all sets which — as we saw — leads to a contradiction. Thus *the class of all sets* (which does exist, comp. Chapter III, § 7) *is not a set.*

Let us note that in the period before the axiomatization of set theory, and hence in the period of "naive" set theory, it was common to assume as obvious the existence for every propositional function $\varphi(x)$ of the set $\{x\colon \varphi(x)\}$. This has led to the contradictions which we mentioned above (which were then called antinomies of set theory), and which have necessitated revising the foundations of set theory. The axiomatic theory of sets, which arose around 1904, eliminated these antinomies.

## § 3. Inequalities for cardinal numbers

Let $\overline{\overline{X}} = \mathfrak{m}$ and $\overline{\overline{Y}} = \mathfrak{n}$. Let us assume that $\mathfrak{m} \leqslant \mathfrak{n}$ if the set $X$ has the same power as some subset of the set $Y$. Therefore

$$(X \subset Y) \Rightarrow (\overline{\overline{X}} \leqslant \overline{\overline{Y}}).$$

If $\mathfrak{m} \leqslant \mathfrak{n}$ and $\mathfrak{m} \neq \mathfrak{n}$, then we write $\mathfrak{m} < \mathfrak{n}$.

By virtue of (1) we have

(21) $$\mathfrak{a} < \mathfrak{c}.$$

We can state the Cantor theorem (§ 2) in the form

(22) $$\mathfrak{m} < 2^{\mathfrak{m}}.$$

In fact, $\mathfrak{m} \neq 2^{\mathfrak{m}}$, and at the same time $\mathfrak{m} \leqslant 2^{\mathfrak{m}}$, since the set $X$ has the same power as the family of all its one-element subsets.

It is easy to prove the following formulas:

(23) if $\mathfrak{m} \leqslant \mathfrak{n}$ and $\mathfrak{n} \leqslant \mathfrak{p}$ then $\mathfrak{m} \leqslant \mathfrak{p}$,

(24) if $\mathfrak{m} \leqslant \mathfrak{n}$ then $\mathfrak{m}+\mathfrak{p} \leqslant \mathfrak{n}+\mathfrak{p}$,

(25) if $\mathfrak{m} \leqslant \mathfrak{n}$ then $\mathfrak{m}\mathfrak{p} \leqslant \mathfrak{n}\mathfrak{p}$,

(26) if $\mathfrak{m} \leqslant \mathfrak{n}$ then $\mathfrak{m}^{\mathfrak{p}} \leqslant \mathfrak{n}^{\mathfrak{p}}$,

(27) if $\mathfrak{m} \leqslant \mathfrak{n}$ then $\mathfrak{p}^{\mathfrak{m}} \leqslant \mathfrak{p}^{\mathfrak{n}}$.

We shall now prove the fundamental *Cantor–Bernstein theorem*:

(28) *if* $\mathfrak{m} \leqslant \mathfrak{n}$ *and* $\mathfrak{n} \leqslant \mathfrak{m}$ *then* $\mathfrak{m} = \mathfrak{n}$.

Proof. Let $\overline{\overline{X}} = \mathfrak{m}$. Since $\mathfrak{n} \leqslant \mathfrak{m}$, the set $X$ contains a subset $Y$ of power $\mathfrak{n}$. But since $\mathfrak{m} \leqslant \mathfrak{n}$, the set $X$ is of power equal to that of some subset of the set $Y$; i.e. there exists a one-to-one mapping $f$ defined on $X$ such that

(29) $$f(X) \subset Y \subset X.$$

We have to define a one-to-one mapping $g$ of $X$ onto $Y$.

Let us set

$$Z = Y - f(X), \qquad S = Z \cup f(Z) \cup ff(Z) \cup \ldots \tag{30}$$

(see Fig. 4 in which $X$ is the largest rectangle, $Y$ is the second in size, $f(X)$ is the third, and so on; $X-S$ is the shaded part).

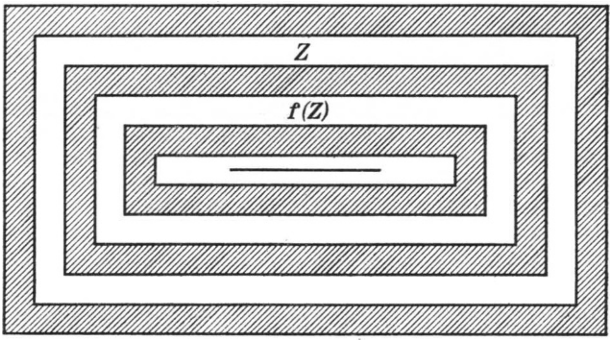



FIG. 4

We define $g$ as follows:

$$g(x) = \begin{cases} x & \text{for} \quad x \in S, \\ f(x) & \text{for} \quad x \in X-S. \end{cases} \tag{31}$$

We shall first prove that

$$g(X) = Y. \tag{32}$$

Since $S \subset X$,

$$X = S \cup (X-S). \tag{33}$$

And therefore

$$g(X) = g(S) \cup g(X-S) = S \cup f(X-S) \tag{34}$$

by virtue of (31). At the same time (because of (30) and Chapter IV, § 4, (14)):

$$f(S) = f(Z) \cup ff(Z) \cup fff(Z) \cup \ldots,$$

and hence applying (30):

$$S = Z \cup f(S). \tag{35}$$

From this and (34) and (33), we obtain

$$g(X) = S \cup f(X-S) = Z \cup f(S) \cup f(X-S) = Z \cup f(X),$$

but by (30) we have

$$Z \cup f(X) = [Y - f(X)] \cup f(X) = Y.$$

We have thus proved formula (32).

It remains to show that $g$ is one-to-one.

Since (according to (31)) $g$ is one-to-one on each of the sets $S$ and $X-S$ separately, we ought to prove that

$$g(S) \cap g(X-S) = \emptyset. \tag{36}$$

Now by (31) we have

$$g(S) = S \quad \text{and} \quad g(X-S) = f(X-S) = f(X) - f(S); \tag{37}$$

at the same time, $f(X) = f(X) - Z$ because $f(X) \cap Z = \emptyset$, and hence

$$f(X) - f(S) = f(X) - [Z \cup f(S)] = f(X) - S$$

because of (35).

Hence, we have $S \cap [f(X) - f(S)] = \emptyset$, whence formula (36) follows by virtue of (37).

This completes the proof of the Cantor–Bernstein theorem.

Another form of this theorem, which is frequently used, is the following:

$$\text{if} \quad A \subset B \subset C \quad \text{and} \quad \overline{\overline{A}} = \overline{\overline{C}}, \quad \text{then} \quad \overline{\overline{A}} = \overline{\overline{B}} = \overline{\overline{C}}. \tag{38}$$

The following theorem holds for an arbitrary mapping $f$:

*If $X$ is the domain of $f$, then*

$$\overline{\overline{f(X)}} \leqslant \overline{\overline{X}}. \tag{39}$$

For, let $y \in f(X)$ and let $g(y)$ be an arbitrary element of the set $f^{-1}(y)$ [we make use of the axiom of choice (Chapter III, § 7) here]. Since the sets $f^{-1}(y)$ for various $y$'s are disjoint, $g$ determines a one-to-one mapping of the set $f(X)$ onto a subset of the set $X$. From this follows formula (39).

## § 4. Properties of the number $\mathfrak{c}$

We have defined the number $\mathfrak{c}$ as the power of the set $\mathscr{E}$ of all real numbers. Let us note that, as stated in Chapter V, § 2, *every open interval $a < x < b$ is of power $\mathfrak{c}$.*

The interval $a \leqslant x \leqslant b$ (where $a < b$) is also of power $\mathfrak{c}$. This follows immediately from formula (38) since

$$\{x\colon a < x < b\} \subset \{x\colon a \leqslant x \leqslant b\} \subset \mathscr{E}.$$

Further, we also deduce from formula (38) that

(40) $\mathfrak{c} = \mathfrak{c}+n = \mathfrak{c}+\mathfrak{a} = \mathfrak{c}+\mathfrak{c} = n\cdot\mathfrak{c}$

($n$ being a natural number);

for (cf. (24)) $\mathfrak{c} \leqslant \mathfrak{c}+n \leqslant \mathfrak{c}+\mathfrak{a} \leqslant \mathfrak{c}+\mathfrak{c}$ and $\mathfrak{c}+\mathfrak{c} \leqslant \mathfrak{c}$, since $\mathfrak{c}+\mathfrak{c}$ is the power of the set

$$\{x\colon 0 < x < 1\} \cup \{x\colon 1 < x < 2\},$$

which is a subset of $\mathscr{E}$.

The generalization to $n$ terms is obtained immediately by induction.

(41) $$2^{\mathfrak{a}} = \mathfrak{c}.$$

For, let $A$ denote the set of all infinite sequences consisting of the numbers 0 and 1. Therefore $\overline{\overline{A}} = 2^{\mathfrak{a}}$. Let $B$ denote the subset of the set $A$ consisting of sequences with an infinite number of zeros. To the sequence $t = (t_1, t_2, \ldots)$ belonging to $B$ we assign the number

$$f(t) = t_1/2+t_2/4+\ldots+t_n/2^n+\ldots, \quad \text{i.e.} \quad f(t) = (0.t_1t_2\ldots)_2,$$

and if $t \in A-B$ we let

$$f(t) = 1+t_1/2+t_2/4+\ldots+t_n/2^n+\ldots, \quad \text{i.e.} \quad f(t) = (1.t_1t_2\ldots)_2$$

(in the binary system of calculation).

It is easy to verify that $f$ is one-to-one. At the same time

$$\{x\colon 0 < x < 1\} \subset f(A) \subset \mathscr{E},$$

and therefore $\overline{\overline{A}} = \overline{\overline{f(A)}} = \mathfrak{c}$ by virtue of (38).

We deduce from this that

(42) $$\mathfrak{a}^{\mathfrak{a}} = \mathfrak{c} = \mathfrak{c}^{\mathfrak{a}},$$

because (cf. (26)) $2^{\mathfrak{a}} \leqslant \mathfrak{a}^{\mathfrak{a}} \leqslant \mathfrak{c}^{\mathfrak{a}} = (2^{\mathfrak{a}})^{\mathfrak{a}} = 2^{(\mathfrak{a}^2)} = 2^{\mathfrak{a}}$.

Similarly, we have

(43) $$n^{\mathfrak{a}} = \mathfrak{c} \quad \text{for} \quad n \geqslant 2.$$

Formula $\mathfrak{a}^{\mathfrak{a}} = \mathfrak{c} = n^{\mathfrak{a}}$ asserts that the set of all infinite sequences whose terms are natural numbers (or whose terms are $1, 2, \ldots, n$) is of power $\mathfrak{c}$.

We shall now deduce from (42) that

$$(44) \quad \mathfrak{c} = \mathfrak{c}\cdot\mathfrak{a} = \mathfrak{c}\cdot\mathfrak{c} = \mathfrak{c}^n = \mathfrak{c}^{\mathfrak{a}} \qquad (n \text{ is a natural number} > 1).$$

In fact

$$\mathfrak{c} \leqslant \mathfrak{c}\cdot\mathfrak{a} \leqslant \mathfrak{c}\cdot\mathfrak{c} \leqslant \mathfrak{c}^n \leqslant \mathfrak{c}^{\mathfrak{a}} = \mathfrak{c}.$$

Let us note that $\mathfrak{c}^2$ is the power of the plane, and more generally: $\mathfrak{c}^n$ is the power of $n$-dimensional Euclidean space $\mathscr{E}^n$. Formula (44) asserts that *the set of all infinite sequences whose terms are real numbers* (i.e. the infinite cartesian product $\mathscr{E}\times\mathscr{E}\times\ldots$) *is also of power* $\mathfrak{c}$.

The last formula dealing with the numbers $\mathfrak{a}$ and $\mathfrak{c}$ is

$$(45) \qquad 2^{\mathfrak{c}} = \mathfrak{a}^{\mathfrak{c}} = \mathfrak{c}^{\mathfrak{c}}.$$

In fact, $\mathfrak{c}^{\mathfrak{c}} = (2^{\mathfrak{a}})^{\mathfrak{c}} = 2^{\mathfrak{a}\mathfrak{c}} = 2^{\mathfrak{c}}$ for $\mathfrak{a}\mathfrak{c} = \mathfrak{c}$ by (44).

Let us set $2^{\mathfrak{c}} = \mathfrak{f}$. By virtue of (22), $2^{\mathfrak{c}} > \mathfrak{c}$; $\mathfrak{f}$ is therefore a cardinal number greater than $\mathfrak{a}$ and $\mathfrak{c}$. Formula (45) asserts that $\mathfrak{f}$ is the power of the family of all subsets of the real line (or more generally—of the family of all subsets of the space $\mathscr{E}^n$); it is at the same time the power of the set of all real valued functions of a real variable (as well as the power of the set of all functions of a real variable whose values are natural numbers).

R e m a r k. We now give a more direct proof of the formula $\mathfrak{c}^2 = \mathfrak{c}$ because of its fundamental importance.

Let $A$ be the square determined by the conditions $0 < x < 1$ and $0 < y < 1$. Since $\overline{\overline{A}} = \mathfrak{c}^2$, our problem depends on the definition of a one-to-one real valued function on the square $A$ (it will follow from this that $\mathfrak{c}^2 \leqslant \mathfrak{c}$; the inequality $\mathfrak{c} \leqslant \mathfrak{c}^2$ is obvious).

Let us develop the numbers $x$ and $y$ in essentially infinite decimal expansions (i.e. containing an infinite number of nonzero digits):

$$x = 0.a_1 a_2 \ldots, \qquad y = 0.b_1 b_2 \ldots,$$

and let

$$(46) \qquad f(x, y) = 0.a_1 b_1 a_2 b_2 \ldots a_n b_n \ldots$$

We must prove that if $f(x, y) = f(\bar{x}, \bar{y})$, then $x = \bar{x}$ and $y = \bar{y}$.

Now the development (46) contains an infinite number of digits which are different from zero; at the same time no number has two different essentially infinite developments, and therefore the formula

$$f(x, y) = 0.a_1 b_1 a_2 b_2 \ldots = 0.\bar{a}_1 \bar{b}_1 \bar{a}_2 \bar{b}_2 \ldots = f(\bar{x}, \bar{y})$$

implies that

$$a_1 = \bar{a}_1, \quad b_1 = \bar{b}_1, \quad a_2 = \bar{a}_2, \quad b_2 = \bar{b}_2, \quad \ldots,$$

i.e. $x = \bar{x}$ and $y = \bar{y}$.

**Exercises**

**1.** Let $\boldsymbol{R}$ be a family of sets each of which has power $\mathfrak{c}$ and let $\overline{\overline{\boldsymbol{R}}} = \mathfrak{c}$. Prove that $\bigcup \boldsymbol{R} = \mathfrak{c}$.

**2.** Let $\overline{\overline{A}}_n = \mathfrak{c}$ for $n = 1, 2, \ldots$ Prove that

$$\overline{\overline{A_1 \times A_2 \times \ldots}} = \mathfrak{c}.$$

**3.** Let $\overline{\overline{T}} = \mathfrak{n}$ and $\overline{\overline{F}}_t = \mathfrak{m}$ for every $t \in T$. Calculate $\overline{\overline{\Pi_t F_t}}$.

**4.** Prove that a necessary and sufficient condition that the set $A$ be of power equal to that of one of its proper subsets (i.e. to some subset distinct from $A$) is that $\mathfrak{a} \leqslant \overline{\overline{A}}$.

*Hint*: In the proof of necessity take into consideration an element $a \in A - f(A)$, then $f(a)$, $ff(a)$, and so on. In the proof of sufficiency consider the sequence $a_1, a_2, \ldots$ contained in $A$ and the function $f$ defined as follows:

$$f(x) = x \quad \text{for} \quad x \neq a_n \quad (n = 1, 2, \ldots) \quad \text{and} \quad f(a_n) = a_{n+1}.$$

CHAPTER VII

# ORDER RELATIONS

## § 1. Definitions

Let the relation $\varrho$, written $\leqslant$, be defined for elements of a given set $X$. Consider the following four conditions:

1. $x \leqslant x$, for each $x$,
2. if $x \leqslant y$ and $y \leqslant x$, then $x = y$,
3. if $x \leqslant y$ and $y \leqslant z$, then $x \leqslant z$,
4. for each pair $x, y$, either $x \leqslant y$ or $y \leqslant x$.

If the conditions 1–3 are satisfied, we say that the relation $\leqslant$ is an *ordering* of $X$ (or that the set $X$ *is ordered*); the relation $\leqslant$ is a *quasi-ordering* if it satisfies conditions 1 and 3 only; it is a *linear ordering* if it satisfies conditions 1–4.

For example, the family $2^X$ is ordered by the relation of inclusion $X \subset Y$. If the family $\boldsymbol{R} \subset 2^X$ is linearly ordered by the above relation, we say that $\boldsymbol{R}$ is *monotonic.*

A quasi-ordered set is called *directed* if to each pair $x, y$ there exists $z$ such that $x \leqslant z$ and $y \leqslant z$. Again such is the family $2^X$ (since $X \subset X \cup Y$ and $Y \subset X \cup Y$).

An ordered set $A$ is said to be *cofinal* with the set $B \subset A$, if for each $x \in A$ there is $y \in B$ such that $x \leqslant y$.

For example the set of all real numbers ordered by the relation $\leqslant$ is cofinal with the set of positive integers.

Obviously, if $A$ contains the last element $a$, it is cofinal with $\{a\}$.

## § 2. Similarity. Order types

We say that the relation $\leqslant$ which orders the set $A$ and the relation $\leqslant^*$ which orders the set $B$ establish *similar* orderings of $A$ and $B$ if there exists a one-to-one mapping $f$ (called a *similarity mapping*) of $A$ onto $B$ such that

$$(x \leqslant y) \equiv [f(x) \leqslant^* f(y)],$$

i.e. $x$ precedes $y$ in the set $A$ if and only if $f(x)$ precedes $f(y)$ in the set $B$.

For example, the relation $\leqslant$ establishes similar orderings of the set of natural numbers and the set of numbers of the form $1-1/n$.

Just as cardinal numbers were assigned to sets, we assign *order types* to order relations or, as we say, to ordered sets. Namely we assign the same order type to two ordered sets if and only if they are similar. We depend here on the property of the similarity relation of being an equivalence relation, i.e.

(a) every ordered set is similar to itself,

(b) if $A$ is similar to $B$, then $B$ is similar to $A$,

(c) if $A$ is similar to $B$ and $B$ is similar to $C$, then $A$ is similar to $C$.

We omit the simple proofs of these properties.

Obviously two similar sets have the same cardinality.

The following order types are particularly important: $\omega$—the type of the set of natural numbers, $\omega^*$—the type of the set of negative integers, $\eta$—the type of the set of rational numbers, and $\lambda$—the type of the set of all real numbers (all these sets are considered to be ordered by the relation $\leqslant$).

The type of a finite set, consisting of $n$ numbers, is denoted by $n$.

THEOREM. *Every countable linearly ordered set $A$ is similar to some subset of the set $\mathscr{R}$ of all rational numbers (ordered with respect to the relation $\leqslant$).*

Proof. Let us arrange the elements of the set $A$ (ordered by the relation $\varrho$) in a sequence $a_1, a_2, \ldots, a_n, \ldots$ consisting of distinct terms (we assume that $A$ is infinite; for finite sets the theorem is obvious).

We define a similarity mapping $f$ of $A$ onto a subset of $\mathscr{R}$, in the following way.

Let $f(a_1) = 0$; $f(a_2)$ is defined as an (arbitrary) rational number which is less than $f(a_1)$ if $a_2 \varrho a_1$, but larger than $f(a_1)$ if $a_1 \varrho a_2$. The inductive definition of the number $f(a_{n+1})$ is the following: if, in the set $A$, $a_{n+1}$ precedes all the elements $a_1, a_2, \ldots, a_n$, then $f(a_{n+1})$ is a rational number less than all the numbers $f(a_1)$, $f(a_2), \ldots, f(a_n)$; analogously if $a_{n+1}$ follows all the elements $a_1$, $a_2, \ldots, a_n$, then the number $f(a_{n+1})$ is larger than all the numbers $f(a_1), f(a_2), \ldots, f(a_n)$; finally, if none of these cases holds, then let $a_k$ be the last among those elements $a_1, a_2, \ldots, a_n$, which precede $a_{n+1}$ and let $a_m$ be the first among those which follow $a_{n+1}$; then let us set

$$f(a_{n+1}) = \{f(a_k)+f(a_m)\}/2.$$

The function $f$ defined in this way is obviously one-to-one. Moreover, for every $n$ it is a similarity mapping of the set $\{a_1, a_2, \ldots, a_{n+1}\}$ onto the set $\{f(a_1), f(a_2), \ldots, f(a_{n+1})\}$. But from this it follows that the function $f$ is a similarity mapping of the entire set $A$ onto $f(A)$. For if $a_i \varrho a_j$, then, denoting by $n+1$ the larger of the two numbers $i$ and $j$, we deduce from the similarity of the sets $\{a_1, a_2, \ldots, a_{n+1}\}$ and $\{f(a_1), f(a_2), \ldots, f(a_{n+1})\}$ that $f(a_i) < f(a_j)$.

## § 3. Dense ordering

We say that a linear ordering of the set $A$ is *dense* if whenever $a < b$, then there exists $c$ such that $a < c$ and $c < b$.

An example of a dense ordering is the ordering of the rational numbers (with respect to the "less than" relation).

Every countable set with dense ordering, without a first and last element, is of type $\eta$.†

## § 4. Continuous ordering

Definitions. A subset $B$ of an ordered set $A$ is said to be an *initial interval* of $A$ if together with each of its elements $x \in B$ it contains all the elements of the set $A$ which precede $x$, i.e. if

$$(y \leqslant x \in B) \Rightarrow (y \in B).$$

Given a set $Z \subset A$, the earliest element $a$ of the set $A$ which satisfies the condition

$$(x \in Z) \Rightarrow (x \leqslant a)$$

(if it exists) is called the *least upper bound* of $Z$.

---

† For a proof, see Hausdorff, *Set Theory*, Chapter 3, § 11, Theorem IV; or Kuratowski and Mostowski, *Set Theory*, p. 217.

An ordering of the set $A$ is *continuous* if it is dense, and if, furthermore, for each of its initial intervals $B$ which is nonvoid and distinct from $A$ there exists a least upper bound.

The set $\mathscr{E}$ of all real numbers is of continuous type. This is in fact only a different formulation of the known Dedekind axiom of continuity.

The ordering of the set of rational numbers is not continuous; for we can take as $B$ the set of rational numbers less than $\sqrt{2}$. (We also say that $\sqrt{2}$ determines a "*gap*" in the set of rational numbers.)

Remark. The following theorem which we give here without proof contains the most essential part of the theory of irrational numbers due to Dedekind.

*Let $\mathscr{R}$ denote the set of all rational numbers and let $\boldsymbol{K}$ denote the family of all its initial intervals which are non-empty, distinct from $\mathscr{R}$, and which do not possess a last element. Then the relation $\subset$ establishes an ordering of the family $\boldsymbol{K}$ of type $\lambda$.*

Hence, real numbers can be defined as the initial intervals of the set $\mathscr{R}$ of all rational numbers which are non-empty, distinct from $\mathscr{R}$, and which do not possess a last element.

## § 5. Inverse systems, inverse limits

Let $T$ be a directed set. Let $X$ be a set-valued mapping, $X\colon T \to 2^A$; thus $X_t \subset A$ for each $t \in T$. Let $f$ be a mapping defined on $T \times T$ for pairs $\langle t_0, t_1 \rangle$ where $t_0 \leqslant t_1$, and such that

(1) $$f_{t_0 t_1}\colon X_{t_1} \to X_{t_0}.$$

We assume further that

(2) $$f_{t_0 t_1} \circ f_{t_1 t_2} = f_{t_0 t_2} \quad \text{for} \quad t_0 < t_1 < t_2 \quad \text{(transitivity)}$$

and

(3) $$f_{tt} = \text{identity}.$$

Then we call the triple $(T, X, f)$ an *inverse system*.†

The inverse limit of the system $(T, X, f)$, denoted

$$X_\infty \quad \text{or} \quad \underset{\longleftarrow}{\mathrm{Lim}}(T, X, f) \quad \text{or} \quad \underset{t, t_0 < t_1}{\mathrm{Lim}} \{X_t, f_{t_0 t_1}\},$$

† See S. Eilenberg and N. Steenrod, *Foundations of Algebraic Topology*, Princeton, 1952, Chapter VIII. Comp. P.S. Alexandroff, *Ann. of Math.* **30** (1928).

is the subset of the cartesian product $\prod_{t\in T} X_t$ composed of elements $z = \{z^t\}$ such that

$$(4)\qquad f_{t_0t_1}(z^{t_1}) = z^{t_0}.$$

In other words, we have for $z \in X_\infty$

$$(5)\qquad f_{t_0t_1}\circ \pi_{t_1} = \pi_{t_0}.$$

We shall agree to write

$$(6)\qquad f_t = \pi_t | X_\infty, \quad \text{i.e.} \quad f_t(z) = z^t.$$

Consequently

$$(7)\qquad f_{t_0t_1}\circ f_{t_1} = f_{t_0}, \quad \text{hence} \quad f_{t_0}^{-1} = f_{t_1}^{-1}\circ f_{t_0t_1}^{-1}.$$

Consider two inverse systems $(T, X, f)$ and $(T, Y, g)$. Suppose that $h$ assigns to each $t$ a mapping

$$h_t\colon X_t \to Y_t$$

such that (for $t_0 < t_1$) commutativity holds in the diagram

$$(8)\qquad \begin{array}{ccc} X_{t_0} & \xleftarrow{f_{t_0t_1}} & X_{t_1} \\ {\scriptstyle h_{t_0}}\downarrow & & \downarrow{\scriptstyle h_{t_1}} \\ Y_{t_0} & \xleftarrow[g_{t_0t_1}]{} & Y_{t_1} \end{array}$$

i.e.

$$(9)\qquad h_{t_0}\circ f_{t_0t_1} = g_{t_0t_1}\circ h_{t_1}.$$

Then we may define a mapping

$$h_\infty\colon X_\infty \to Y_\infty$$

so that the following diagram is commutative for each $t \in T$:

$$(10)\qquad \begin{array}{ccc} X_t & \xleftarrow{f_t} & X_\infty \\ {\scriptstyle h_t}\downarrow & & \downarrow{\scriptstyle h_\infty} \\ Y_t & \xleftarrow[g_t]{} & Y_\infty \end{array}$$

We put $y = \{y^t\} = h_\infty(z)$ for $z \in X_\infty$, where

$$(11)\qquad h_t(z^t) = y^t.$$

It is easily seen that

(12) *if each $h_t$ is a one-to-one mapping onto, so is $h_\infty$.*

**Exercises**

**1.** Let $X$ and $Y$ be two subsets of the ordered set $A$ such that $X \cup Y = A$, $X \cap Y = \emptyset$ and $(x \in X)(y \in Y) \Rightarrow (x < y)$. We say that the pair $X$, $Y$ is a *cut* of the set $A$.

Prove that if $X_1, Y_1$ and $X_2, Y_2$ are cuts of the set $A$, then either $X_1 \subset X_2$ or $X_2 \subset X_1$.

**2.** Prove that every linearly ordered set is similar to some monotonic family of subsets of this set (ordered by the relation $\subset$).

**3.** Let $\boldsymbol{R}$ be a monotonic family of subsets of the set $Z$. Prove that the family of all sets $\bigcup X$ and $\bigcap X$, where $X \subset \boldsymbol{R}$, is also monotonic.

**4.** Give an example of a linearly ordered set which is not of type $\omega$, but which, despite this fact, possesses a first element and which is such that to every element there exists an element immediately following and (except to the first) an element immediately preceding it.

**5.** A subset $G$ of a linearly ordered set $A$ is said to be *dense with respect to $A$* if between every two elements $x$ and $y$ of the set $A$ there is an element $z$ of the set $G$.

Prove that a set $A$ of type $\lambda$ contains a countable subset which is dense with respect to $A$.

**5a.** Show the inverse theorem (see Remark of § 4): if $A$ is continuously ordered, contains a countable subset dense in $A$, but contains no first nor last element, then $A$ is of order type $\lambda$.

R e m a r k. The following conjecture, called the *Souslin problem*, is independent of the axioms of set theory (that is, it can neither be proved nor disproved with the help of these axioms):† let every family of disjoint intervals of a continuously ordered set $A$ be countable; is $A$ necessarily of type $\lambda$ (assuming that $A$ contains no first and last elements)?

**6.** Let us establish an ordering for the set $\mathscr{E}^2$ (of all complex numbers) by assuming that of two complex numbers with distinct imaginary parts that one is earlier whose imaginary part is smaller, and of two numbers with equal imaginary parts that one is earlier which has the smaller real part.

Prove that in the set $\mathscr{E}^2$ there does not exist a countable part which is dense with respect to $\mathscr{E}^2$.

**7.** (a) The family of all infinite sequences with real terms can be ordered in the following way: the sequence $a_1, a_2, \ldots$ precedes the sequence $b_1, b_2, \ldots$ if there exists a $k$ such that $a_n < b_n$ for $n > k$.

(b) A family of real valued functions is ordered by the relation

$$(f \leqslant g) \equiv \bigwedge_x [f(x) \leqslant g(x)].$$

† As recently shown by Tennenbaum and Solovay.

CHAPTER VIII

# WELL ORDERING

## § 1. Well ordering

D e f i n i t i o n. We say that a linear ordering of a set $A$ is a *well ordering* if every non-empty subset of the set $A$ has a first element.

We call the order types of well ordered sets *ordinal numbers* (concisely: *ordinals*).

EXAMPLES. The set of all natural numbers is a well ordered set (this follows directly from the principle of finite induction). Therefore $\omega$ is an ordinal number. On the other hand, none of the order types $\omega^*$, $\eta$, $\lambda$ is an ordinal number.

It follows from the definition of well ordering that every subset of a well ordered set is well ordered. It also follows that for every element $a$ of a well ordered set (with the exception of the last element, provided the set contains a last element) there exists an element $b$ which is its immediate successor. Namely, $b$ is the first element of the set $\{x\colon a < x\}$.

On the other hand, a well ordered set can contain an element (which is not its first element), for which there does not exist an element which is an immediate predecessor. For example, the set consisting of the numbers $1-1/n$ $(n = 1, 2, ...)$ together with the number 1 is well ordered, but there does not exist an element in this set which immediately precedes the number 1.

If the set $A$ is well ordered, then for every initial interval $B$ which is distinct from $A$ there exists one and only one element $b$ in $A$ such that

$$B = \{x\colon x < b\}.$$

Namely, $b$ is the first element of the set $A-B$. It is therefore the least upper bound of the interval $B$ if $B$ does not contain a last element; but if $B$ contains a last element, then $b$ is the element which immediately succeeds this element.

Let us set

(1) $$P(a) = \{x\colon x < a\}.$$

$P$ is a one-to-one correspondence between the elements of $A$ and the family $\boldsymbol{R}$ of all initial intervals of $A$ which are distinct from $A$.

Thus $P\colon A \to \boldsymbol{R}$ is a similarity mapping (where $\boldsymbol{R}$ is ordered by the inclusion relation: $X \subset Y$).

For if $a < b$, then $x < a \Rightarrow x < b$, i.e. $P(a) \subset P(b)$ and conversely.

## § 2. Theorem on transfinite induction

*Let $A$ be a well ordered set and let $\varphi(x)$ be a propositional function, where $x \in A$, satisfying the following condition for every $x$:*

(2) $$\textit{if } \bigwedge_y[(y < x) \Rightarrow \varphi(y)], \textit{ then } \varphi(x).$$

*Then every element of the set $A$ satisfies the propositional function $\varphi(x)$, i.e. $\bigwedge_x \varphi(x)$.*

Let us assume that this is not the case, i.e. that the set $Z$ of elements of the set $A$ which do not satisfy the propositional function $\varphi(x)$ is nonvoid. Let $x_0$ be the first element of the set $Z$. Therefore

$$\bigwedge_y[(y < x_0) \Rightarrow \varphi(y)].$$

But it follows from this by virtue of (2) that the proposition $\varphi(x_0)$ is true. But then $x_0 \notin Z$.

R e m a r k. The principle of finite induction known from arithmetic is a particular case of the preceding theorem; namely, the case where $A$ is the set of natural numbers.

## § 3. Theorems on the comparison of ordinal numbers

D e f i n i t i o n. Let $\alpha$ and $\beta$ be ordinal numbers; let $\alpha$ be the order type of the set $A$ and let $\beta$ be that of the set $B$. We write $\alpha < \beta$ if the set $A$ is similar to some initial interval of the set $B$ which is distinct from $B$.

We assume the above definition of the "less than" relation in connection with the following theorems.

THEOREM 1. *No well ordered set is similar to an initial interval which is distinct from the set itself, i.e.*

(3) $$\alpha \not< \alpha.$$

Let us assume the contrary. That is, let us assume that $f\colon A \to P(a)$ is a similarity mapping for some $a \in A$. Since $f(a) \in P(a)$ we have $f(a) < a$. Therefore the set

$$Z = \{x\colon f(x) < x\}$$

is not empty. Let $x_0$ be its first element. Hence

$$(4) \qquad f(x_0) < x_0,$$

and since $f$ is a similarity mapping of $A$ onto $P(a)$, we deduce that

$$(5) \qquad f[f(x_0)] < f(x_0),$$

but then—comparing formula (5) with (4)—it follows that $x_0$ is not the first element of the set $Z$.

THEOREM 2. *No two initial intervals of a well ordered set are similar.*

This follows directly from the preceding theorem, for of two distinct initial intervals $P(a)$ and $P(b)$ one is an initial interval of the other (depending on whether $a < b$ or $b < a$).

Theorem 2 can also be expressed in the following manner:

$$(6) \qquad \textit{if } \alpha < \beta \textit{ then } \beta \nless \alpha.$$

Since an initial interval of an initial interval of the set $A$ is an initial interval of this set, we have:

$$(7) \qquad \textit{if } \alpha < \beta \textit{ and } \beta < \gamma \textit{ then } \alpha < \gamma.$$

We shall now prove the following fundamental theorem:

THEOREM 3. *If* $\alpha \neq \beta$ *then* $\alpha < \beta$ *or* $\beta < \alpha$. In other words, *if the sets $A$ and $B$ are well ordered, then either the set $A$ is similar to an initial interval of the set $B$ or the set $B$ is similar to an initial interval of the set $A$.*

Proof. We shall denote by $P_A(x)$ the initial intervals of $A$ and by $P_B(y)$ the initial intervals of $B$. We shall write $M \simeq N$ if $M$ and $N$ are similar.

We set

$$(8) \qquad X = \{x\colon \bigvee_y [P_A(x) \simeq P_B(y)]\}.$$

By Theorem 2, for every $x \in X$, there exists only one element $y$ such that $P_A(x) \simeq P_B(y)$. Hence we can denote this $y$ by $f(x)$. Therefore the equivalence

$$(9) \qquad [y = f(x)] \equiv [P_A(x) \simeq P_B(y)]$$

holds for every $x \in X$.

We shall prove that $X$ is an initial interval of $A$. Let $x' < x \in X$. We must prove that $x' \in X$. Since $x \in X$, there exists (by virtue of (8)) a similarity mapping of the interval $P_A(x)$ onto the interval $P_B[f(x)]$; but since $P_A(x')$ is an initial interval of $P_A(x)$, under this mapping the interval $P_A(x')$ goes over into an initial interval of $P_B[f(x)]$, and hence onto an initial interval of $B$. This means that $x' \in X$, i.e. that $X$ is an initial interval of $A$.

Analogously, the set $f(X)$ is an initial interval of the set $B$. For by virtue of (9) and formula (12) of Chapter IV, § 4, we have

$$(10) \quad f(X) = \{y: \bigvee_x [y = f(x)]\} = \{y: \bigvee_x [P_B(y) \simeq P_A(x)]\}.$$

Moreover, as we have already proved, the condition $x' < x$ implies that the interval $P_B[f(x')]$ is an initial interval of the interval $P_B[f(x)]$, and hence that $f(x') < f(x)$. This means that

$$(11) \qquad X \simeq f(X).$$

It remains to prove that either $X = A$ or $f(X) = B$. Let us assume the contrary, that $X \neq A$ and that $f(X) \neq B$. Since the sets $X$ and $f(X)$ are initial intervals of $A$ and $B$, there exist $a \in A$ and $b \in B$ such that

$$X = P_A(a) \quad \text{and} \quad f(X) = P_B(b).$$

By virtue of (11) we therefore have $P_A(a) \simeq P_B(b)$, whence it follows by (8) that $a \in X$, i.e. that $a \in P_A(a)$, hence $a < a$, which is a contradiction.

Theorem 3 implies the following:

THEOREM 4. *If the sets $A$ and $B$ are well ordered, then their powers satisfy the trichotomy condition, i.e.*

$$\textit{either } \overline{\overline{A}} = \overline{\overline{B}}, \textit{ or } \overline{\overline{A}} < \overline{\overline{B}}, \textit{ or } \overline{\overline{B}} < \overline{\overline{A}}.$$

A question of fundamental significance which arises here naturally is: can every set be well ordered?

We shall consider this question in § 7.

## § 4. Sets of ordinal numbers

We shall use the following notation:

(12) $$\Gamma(\alpha) = \{\xi\colon\ \xi < \alpha\}.$$

THEOREM 1. *The set $\Gamma(\alpha)$ is well ordered (by the relation $\leqslant$) and the order type of this ordering is $\alpha$.*

Proof. Let $A$ be a well ordered set of type $\alpha$ and let $\tau(x)$ for $x \in A$ be the order type of the interval $P(x)$.

$\tau$ is a similarity mapping of $A$ onto $\Gamma(\alpha)$. For if $x' < x$, then the set $P(x')$ is distinct from $P(x)$ and is an initial interval of $P(x)$, and hence (cf. Theorem 1, § 3) $\tau(x') < \tau(x)$. At the same time, every $\xi \in \Gamma(\alpha)$ is a value of $\tau$. For let $\xi \in \Gamma(\alpha)$, i.e. $\xi < \alpha$; by the definition of the relation $<$ for ordinal numbers, a set of type $\xi$ is similar to some initial interval $P(x')$ of $A$; and hence $\xi = \tau(x')$.

THEOREM 2. *Every set of ordinals is well ordered (by the relation $\leqslant$).*

We have to prove that every non-empty set $\Phi$ of ordinals contains a least number. Let $\alpha \in \Phi$. If $\alpha$ is not the least number of $\Phi$ then the set $\Phi \cap \Gamma(\alpha)$ is nonvoid and therefore, being a subset of the well ordered set $\Gamma(\alpha)$, it contains a least number $\beta$. The number $\beta$ is the least number of the set $\Phi$. For if $\xi \in [\Phi - \Gamma(\alpha)]$ then $\xi \geqslant \alpha$ and hence $\xi > \beta$.

THEOREM 3. *For every set $\Phi$ of ordinals there exists an ordinal number which is greater than every number of this set.*

Such a number is $\alpha+1$, where $\alpha$ is the order type of the set

$$\Psi = \bigcup_\xi \Gamma(\xi) \qquad \text{where} \qquad \xi \in \Phi,$$

and $\alpha+1$ denotes the type of the set $\Psi \cup \{\alpha\}$ (cf. § 6).

In fact, for every $\xi$ the set $\Gamma(\xi)$ is an initial interval of the set $\Psi$. If $\Gamma(\xi) = \Psi$, then $\xi = \alpha$ (by virtue of Theorem 1); and in the contrary case $\xi < \alpha$. Therefore for every $\xi$ we have $\xi < \alpha+1$.

THEOREM 4. *There does not exist the set of all ordinal numbers.*

## § 5. The number $\Omega$

Definition. Let us denote by $\Xi$ the set of all order types of countable well ordered sets and by $\Omega$ the order type of the set $\Xi$.

By Theorem 2 of § 4, $\Omega$ is an ordinal number.

We shall prove that

$$\Xi = \Gamma(\Omega). \tag{13}$$

By virtue of Theorem 3 of § 4 there exists an ordinal $\alpha$ greater than every number of the set $\Xi$. Therefore $\Xi \subset \Gamma(\alpha)$. Further $\Xi$ is an initial interval of the set $\Gamma(\alpha)$. For let $\xi' < \xi \in \Xi$; $\xi'$ is therefore an order type of some subset of a countable well ordered set (of type $\xi$); this subset is obviously countable and hence $\xi' \in \Xi$.

Since $\Xi$ is an initial interval of $\Gamma(\alpha)$, there exists (cf. (1)) a number $\gamma \leqslant \alpha$ such that $\Xi = \Gamma(\gamma)$. In order to prove formula (13) it remains to show that $\gamma = \Omega$. But this follows immediately from the definition of $\Omega$ and from Theorem 1, § 4, by virtue of which $\Gamma(\gamma)$ has the type $\gamma$.

THEOREM 1. *The set $\Gamma(\Omega)$ is noncountable, i.e.*

$$\overline{\overline{\Gamma(\Omega)}} > \mathfrak{a}. \tag{14}$$

In fact, if the set $\Gamma(\Omega)$ were countable, then its order type would belong to $\Xi$, i.e. $\Omega \in \Xi$, whence by (13) we should have $\Omega \in \Gamma(\Omega)$, i.e. $\Omega < \Omega$ which is impossible.

Remark 1. The cardinal number $\overline{\overline{\Gamma(\Omega)}}$ is denoted by the symbol $\aleph_1$ ("aleph" 1). Hence we have $\aleph_1 > \mathfrak{a}$, as well as $\mathfrak{c} > \mathfrak{a}$ (Chapter VI, § 3, (21)). However, we were led to the number $\aleph_1$ by entirely different reasoning than that used to define the number $\mathfrak{c}$. Are these numbers equal? The hypothesis, called the *continuum hypothesis*, asserting that

$$\aleph_1 = \mathfrak{c} \tag{15}$$

is independent of the axioms of set theory.†

THEOREM 2. *$\aleph_1$ is the number immediately following the number $\mathfrak{a}$, i.e. if $\mathfrak{m} < \aleph_1$ then $\mathfrak{m} \leqslant \mathfrak{a}$.*

Proof. Let $\overline{\overline{A}} = \mathfrak{m}$. Since $\mathfrak{m} < \aleph_1$, there is $B \subset \Xi$ such that $\overline{\overline{B}} = \mathfrak{m}$. Let $\beta$ denote the order type of $B$. Therefore the sets $B$ and $\Gamma(\beta)$ are similar and hence of equal power, i.e. $\overline{\overline{\Gamma(\beta)}} = \mathfrak{m}$.

---

† See P. J. Cohen, *Set Theory and the Continuum Hypothesis.*

It follows that $\beta < \Omega$, for otherwise $\Omega \leqslant \beta$, whence $\Gamma(\Omega) \subset \Gamma(\beta)$, and therefore $\aleph_1 = \overline{\overline{\Gamma(\Omega)}} \leqslant \overline{\overline{\Gamma(\beta)}} = \mathfrak{m}$ contrary to the assumption. It follows from the inequality $\beta < \Omega$ (by the definition of $\Omega$) that the set $B$ is countable, i.e. $\mathfrak{m} \leqslant \mathfrak{a}$.

Remark 2. The alephs $\aleph_2, \aleph_3, \ldots, \aleph_n, \ldots$ can be defined similarly to $\aleph_0 = \mathfrak{a}$ and $\aleph_1$. Namely $\aleph_2$ is the cardinality of the set of all order types of well ordered sets of power $\aleph_1$; $\aleph_n$ is defined by induction. It follows that

$$\aleph_n < \aleph_{n+1}.$$

Here we assume obviously the

Axiom of infinity. *There exist infinite sets.*

Another axiom is needed in order to prove the existence of cardinal numbers greater than all $\aleph_n$, $n = 1, 2, \ldots$ This is the

Axiom of replacement. *If to every element $x$ of a set $A$ there corresponds an element $y$ (which belongs or does not belong to $A$), then the totality of all these $y$'s is a set.*

Consequently, if we denote $y$ by $f(x)$, then $f$ is a mapping in the usual sense (Chapter IV, § 1).

Here we use this axiom as follows. As shown, there exists, for each $n$, a set $Z_n$ of power $\aleph_n$. Denote by $A$ the set of positive integers and by $\boldsymbol{R}$ (according to the axiom of replacement) the family of all sets $Z_n$ for $n = 1, 2, \ldots$, and consider the union $\bigcup_{n=1}^{\infty} Z_n$ (which exists according to the generalized union axiom, cf. Chapter IV, § 5, Remark). Its power exceeds each $\aleph_n$, and it is natural to denote it $\aleph_\omega$.

## § 6. The arithmetic of ordinal numbers

Let $\alpha$ and $\beta$ be two ordinal numbers (or more generally, two order types). Let $A$ and $B$ be two sets with order types $\alpha$ and $\beta$, respectively; let us assume also that $A \cap B = \emptyset$ (see Chapter VI, § 1, concerning the possibility of making such an assumption). Let us establish an ordering of the set $A \cup B$ by assuming that every element of the set $A$ precedes every element of the set $B$ and that in the domain of each of the sets $A$ and $B$ individually the ordering does not change.

We denote the *order type of the set $A \cup B$ by $\alpha+\beta$.*

We shall prove that, *under the assumption that* $\alpha$ *and* $\beta$ *are ordinal numbers,* $\alpha+\beta$ *is also an ordinal number.*

We have to prove that the set $A \cup B$ with the above-established ordering of its elements is well ordered. Hence, let $\emptyset \neq X \subset A \cup B$. If $X \cap A \neq \emptyset$, then—since the set $A$ is well ordered—the set $X \cap A$ contains an earliest element; this element is the earliest element of the entire set $X = (X \cap A) \cup (X \cap B)$, inasmuch as it precedes, by the definition of the ordering of the set $A \cup B$, each of the elements of the set $X \cap B$.

Now, if $X \cap A = \emptyset$, then $X \subset B$ and therefore there exists an earliest element in the set $X$.

EXAMPLES. $\alpha+1 > \alpha$ whereby $\alpha+1$ follows immediately after $\alpha$. The number $\omega+\omega$ is the type of the set of numbers of the form $1-1/n$ together with the numbers of the form $2-1/n$ where $n = 1, 2, \ldots$ Let us note that $1+\omega = \omega$; and hence addition is not commutative.

We denote *by* $\alpha \cdot \beta$ *the order type of the cartesian product* $A \times B$ *ordered as follows*:

$$[\langle x, y\rangle \leqslant \langle u, v\rangle] \equiv [(y \leqslant v) \vee ((y = v)(x \leqslant u))].$$

*Under the assumption that* $\alpha$ *and* $\beta$ *are ordinal numbers,* $\alpha \cdot \beta$ *is also an ordinal number.*

For, let $\emptyset \neq Z \subset A \times B$. Let $Y$ denote the projection of the set $Z$ onto the $B$-axis. Hence we have $\emptyset \neq Y \subset B$. Let $b$ be the earliest element of the set $B$ and let $X = \{x: \langle x, b\rangle \in Z\}$. Finally let $a$ be the first element of the set $X$. It is easy to verify that $\langle a, b\rangle$ is the first element of the set $Z$.

EXAMPLES. $2 \cdot \omega$ is the order type of the cartesian product $\{1, 2\} \times J$ (where $J$ is the set of natural numbers) ordered as follows:

$$\langle 1, 1\rangle, \langle 2, 1\rangle, \langle 1, 2\rangle, \langle 2, 2\rangle, \ldots,$$

and hence $2 \cdot \omega = \omega$.

On the other hand, $\omega \cdot 2 = \omega+\omega$ is the order type of the product $J \times \{1, 2\}$ (see the example given above).

As we see, multiplication is not commutative.

$\omega \cdot \omega$ is the type of the set of all numbers of the form $k-1/n$ where $k = 1, 2, \ldots$ and $n = 1, 2, \ldots$

Instead of $\omega \cdot \omega$ we write $\omega^2$. In general, $\alpha^{n+1} = \alpha^n \cdot \alpha$.

We denote by $a^\omega$ (for $a > 1$) the least number larger than any $a^n$, where $n = 1, 2, \ldots$

More generally, the definition of *exponentiation* (and of many other operations) can be introduced with the aid of the following definitions of the concept of transfinite sequence and of its limit.

Definitions. By a *transfinite sequence* of type $\alpha$ we mean a mapping $\varphi$ whose domain is $\Gamma(\alpha)$; instead of $\varphi$, we write usually

$$a_0, a_1, \ldots, a_\xi, \ldots, \qquad \xi < \alpha.$$

If $\alpha$ is a limit ordinal and the terms of the above transfinite sequence are ordinals, then its *limit*, denoted

$$\lim_{\xi<\alpha} a_\xi,$$

is the least ordinal larger than all $a_\xi$ for $\xi < \alpha$.

With the help of the above definitions, we define the *power* $\alpha^\beta$ (for $\alpha > 1$) as follows

1. $$\alpha^0 = 1,$$
2. $$\alpha^{\xi+1} = \alpha^\xi \cdot \alpha,$$
3. $$\alpha^\lambda = \lim_{\xi<\lambda} \alpha^\xi,$$

where $\lambda$ is a limit ordinal (cf. the theorem on the definition by transfinite induction, § 8, Theorem 2).

Remark. The arithmetic of ordinal numbers forms at present a well established theory which we shall not develop any further here.† The main objective of the theory given above was a study of ordinals of countable sets; all these ordinals can be obtained with the aid of sets of real (or even rational) numbers (see Chapter VII, § 2).

Definition 1. The ordinal $\alpha$ is *cofinal with* the (limit) ordinal $\beta$ if the set $\Gamma(\alpha)$ is cofinal with a subset of type $\beta$ (cf. Chapter VII, § 1).

Thus, for example $\omega_\omega$ (see below) is cofinal with $\omega$, while $\Omega$ is not.

† See e.g. W. Sierpiński, *Cardinal and Ordinal Numbers*, or F. Hausdorff, *Set Theory*, Chapter III.

D e f i n i t i o n 2. The ordinal $\alpha$ is called *initial* relative to the cardinal number $\mathfrak{m}$ if $\alpha$ is the smallest possible ordinal number of a well ordered set $Z$ such that $\overline{\overline{Z}} = \mathfrak{m}$.

For example, $\omega$ and $\Omega$ are initial numbers. Let us denote them $\omega_0$ and $\omega_1$, and more generally let us denote by $\omega_n$ the initial number relative to $\aleph_n$.

As stated in § 5, the existence of $\aleph_\omega$ follows from the axiom of replacement. One can show similarly the existence of initial ordinals $\omega_\omega, \omega_\Omega$ and so on. The two quoted numbers are cofinal with smaller numbers (namely with their indices). This leads to the problem of the existence of an initial ordinal $\omega_\lambda$, where $\lambda$ is a limit number ($\neq 0$), which is not cofinal with a smaller number. The existence of numbers of that kind, called *inaccessible* numbers, does not follow from our axioms (including the axiom of infinity and the axiom of replacement). One could, of course, assume axiomatically the existence of inaccessible numbers, but this would not solve analogous problems concerning greater and greater ordinals.

Here we have before us a problem which has not been solved satisfactorily thus far.

R e m a r k. It can be shown that to each ordinal $\alpha$ there corresponds an initial ordinal $\omega_\alpha$. This leads to the definition: $\aleph_\alpha = \overline{\overline{\Gamma(\omega_\alpha)}}$.

Let us add that we write, instead of $X^{\Gamma(\omega_\alpha)}$, more concisely $X^{\aleph_\alpha}$ (as we did in the case of $\alpha = 0$).

## § 7. The well-ordering theorem

We shall deduce this theorem, which is of fundamental importance for the theory of sets (cf. e.g. Theorem 4, § 3), from the axiom of choice. To this end, we shall prove first of all, the following theorem which is a generalization of the axiom of choice.

THEOREM 1. (General principle of choice.) *For every set $A$ there exists a mapping $e$ which assigns to every non-empty subset of $A$ one of its elements, i.e.*

(16) $$e(X) \in X \quad \textit{for every} \quad \emptyset \neq X \subset A.$$

P r o o f. Let $F(X) = \{X\} \times X$, i.e. the set $F(X)$ consists of ordered pairs of the form $\langle X, x\rangle$ where $x \in X$. Let $\boldsymbol{R}$ denote the family consisting of all sets $F(X)$, where $\emptyset \neq X \subset A$. This is a family consisting of nonvoid disjoint sets. By the axiom of choice (Chapter III, § 7) there exists therefore a set consisting of elements, one chosen from each of the sets belonging to $\boldsymbol{R}$; this set is the desired mapping $e$.

THEOREM 2. (Zermelo theorem.) *For every set $A$ there exists a relation which establishes its well ordering.*

Proof. Let us consider the ordinal numbers $\beta$ with the following properties: there exists a transfinite sequence $f_\beta$ of type $\beta+1$ such that

$$(17)\quad f_\beta(0) = e(A), \quad f_\beta(\xi) = e[A - f_\beta(\Gamma(\xi))] \quad \text{for} \quad \xi \leqslant \beta;$$

in particular

$$f_2(1) = e[A - \{e(A)\}],$$

$$f_2(2) = e\big(A - \{e(A), e[A - \{e(A)\}]\}\big).$$

$f_\beta$ is one-to-one. For if $\xi' < \xi \leqslant \beta$, then $\xi' \in \Gamma(\xi)$ and hence $f_\beta(\xi') \in f_\beta(\Gamma(\xi))$ but $f_\beta(\xi) \in [A - f_\beta(\Gamma(\xi))]$ by (16) and (17).

It follows that the set of values of $f_\beta$, i.e. the set $f_\beta(\Gamma(\beta+1))$, is of order type $\beta+1$.

Hence the numbers $\beta$ form a subset $\Phi$ of the set of all order types of subsets of $A$ which can be well ordered. By virtue of Theorem 3, § 4, there exist ordinal numbers which do not belong to $\Phi$. Let $\alpha$ be the least of them. Therefore, there does not exist $f_\alpha$ satisfying conditions (17) (replacing $\beta$ by $\alpha$), and on the other hand, for every $\beta < \alpha$, there exists $f_\beta$ which satisfies these conditions.

We shall prove that $A$ can be well ordered, its order type being $\alpha$.

To this end, let us first note that if $\beta' \leqslant \beta$ and the transfinite sequence $g_{\beta'}$, of type $\beta'+1$, satisfies conditions analogous to those of (17), i.e.

$$(18)\quad g_{\beta'}(0) = e(A), \quad g_{\beta'}(\xi) = e[A - g_{\beta'}(\Gamma(\xi))] \quad \text{for} \quad \xi \leqslant \beta',$$

then for each $\xi \leqslant \beta'$ the identity

$$(19)\qquad g_{\beta'}(\xi) = f_\beta(\xi)$$

is satisfied (this means that, if $\beta' = \beta$, $f_\beta$ is uniquely determined and that if $\beta' < \beta$, $f_\beta$ is an extension of $f_{\beta'}$).

In fact, let us denote by $\varphi(\xi)$ the propositional function (19), having the set $\Gamma(\beta'+1)$ for its domain.

Let us apply to this function the theorem on transfinite induction (see § 2, where we substitute $\Gamma(\beta'+1)$ for $A$). Hence let us assume that for given $\xi \leqslant \beta'$ the condition $\gamma < \xi$ implies that $g_{\beta'}(\gamma) = f_\beta(\gamma)$ and therefore that $g_{\beta'}(\Gamma(\xi)) = f_\beta(\Gamma(\xi))$, which in turn, by virtue of (18) and (17), implies (19). By the theorem on transfinite induction, (19) holds for every $\xi \leqslant \beta'$.

Let

$$f(\beta) = f_\beta(\beta) \tag{20}$$

for every $\beta < \alpha$.

In order to show that $A$ admits a well ordering of type $\alpha$, it obviously suffices to prove that $f$ is one-to-one and that $A$ is its set of values.

Hence, let $\beta' < \beta$. As we proved (cf. (19)) $f_{\beta'}(\xi) = f_\beta(\xi)$ for every $\xi \leqslant \beta'$, and hence $f_{\beta'}(\beta') = f_\beta(\beta')$ in particular. But since $f_\beta$ is one-to-one, we therefore have $f_\beta(\beta') \neq f_\beta(\beta)$, i.e. $f(\beta') \neq f(\beta)$.

It remains to prove that $f(\Gamma(\alpha)) = A$. Let us suppose that $A - f(\Gamma(\alpha)) \neq \emptyset$, and define $f_\alpha$ as follows:

$$f_\alpha(\beta) = f(\beta) \text{ for } \beta < \alpha \quad \text{and} \quad f_\alpha(\alpha) = e[A - f_\alpha(\Gamma(\alpha))].$$

As can easily be seen, $f_\alpha$ satisfies condition (17) (replacing $\beta$ by $\alpha$). But this contradicts the definition of $\alpha$.

## * § 8. Definitions by transfinite induction

THEOREM 1. *For every set $A$, for every ordinal $\alpha$ and for every mapping $h\colon 2^A \to A$, i.e.*

$$h(X) \in A \quad \textit{for} \quad X \subset A, \tag{21}$$

*there is a transfinite sequence $f$ of type $\alpha+1$ such that*

$$f(\xi) = h[f(\Gamma(\xi))], \tag{22}$$

*where $f(\Gamma(\xi))$ denotes (as always) the set of all $f(\eta)$ with $\eta < \xi$.*

S k e t c h  o f  t h e  p r o o f. Let us assume that the theorem is false and that $\alpha$ is the least number for which there does not exist a transfinite sequence $f$ of type $\alpha+1$ satisfying condition (22). Therefore for every $\beta < \alpha$ there exists $f_\beta$ such that

$$f_\beta(\xi) = h[f_\beta(\Gamma(\xi))] \quad \text{for} \quad \xi \leqslant \beta. \tag{23}$$

We can prove as above that $f_\beta$ is uniquely determined. The transfinite sequence $f$ defined by the formulas

$$f(\beta) = f_\beta(\beta) \text{ for } \beta < \alpha \quad \text{and} \quad f(\alpha) = h[f(\Gamma(\alpha))]$$

then satisfies the conditions of the theorem — contrary to our assumption.

This completes the proof of Theorem 1.

Remark. The Zermelo theorem can be deduced from Theorem 1 by substituting

$$h(X) = e(A-X) \quad \text{for} \quad X \neq A,$$

and denoting by $h(A)$ an arbitrary element of $A$.

We denote by $\Phi$ the set of ordinals $\beta$ for which there exists a transfinite sequence $f_\beta$ satisfying condition (23) and the inequality $f_\beta(\Gamma(\beta)) \neq A$. Let $\alpha$ be the least ordinal which does not belong to $\Phi$. Then $f(\Gamma(\alpha)) = A$, whence it easily follows that $A$ can be well ordered, its order type being $\alpha$.

Another way of defining by transfinite induction is based on the following theorem.†

THEOREM 2. *Let $A$ be a given set, $a$ an element of $A$, $\alpha$ an ordinal, and let $g$: $A \to A$ and $h$: $2^A \to A$ be two given mappings. Then there exists a transfinite sequence $f$ of type $\alpha+1$ such that*:

(i) $$f(0) = a,$$

(ii) $$f(\xi+1) = g(f(\xi)), \quad \textit{for} \quad \xi \leqslant \alpha,$$

(iii) $$f(\lambda) = h[f(\Gamma(\lambda))]$$

*when $\lambda$ is a limit ordinal $\leqslant \alpha$.*

The proof is quite similar to the proof of Theorem 1.

Definition. A set $M \in \boldsymbol{R}$ is called *maximal* (or *saturated*) in $\boldsymbol{R}$ if it is contained in no other member of $\boldsymbol{R}$, i.e. if

$$(M \subset Z \in R) \Rightarrow (Z = M).$$

Theorem 2 implies the following theorem.

† For a statement including both theorems, 1 and 2, see Kuratowski–Mostowski, *Set Theory*, p. 239.

THEOREM 3.† *Let $\boldsymbol{R}$ be a family of subsets of a given set $E$ such that for every monotonic and well ordered (by the inclusion relation) subfamily $\boldsymbol{X} \subset \boldsymbol{R}$ the union $\bigcup \boldsymbol{X}$ belongs to $\boldsymbol{R}$. Then every set $A_0 \in \boldsymbol{R}$ is a subset of a maximal set $M \in \boldsymbol{R}$.*

P r o o f. For each $Z \in \boldsymbol{R}$, which is not a maximal member of $\boldsymbol{R}$, denote by $G(Z)$ a member of $\boldsymbol{R}$ such that $Z \subset G(Z) \neq Z$. If $Z$ is maximal, then we put $G(Z) = Z$ (we assume also that if $Z \notin \boldsymbol{R}$ but $Z \subset E$, then $G(Z) = Z$). Finally put $H(\boldsymbol{X}) = \bigcup \boldsymbol{X}$ whenever $\boldsymbol{X} \subset 2^E$.

Then, we can define by transfinite induction a sequence $A_0, A_1, \ldots, A_\xi, \ldots$ such that

$$A_{\xi+1} = G(A_\xi) \quad \text{and} \quad A_\lambda = \bigcup_{\xi<\lambda} A_\xi,$$

where $\lambda$ is a limit ordinal.

Obviously, this sequence is monotonic and for each $\xi$, $A_\xi \in \boldsymbol{R}$. Hence, if $\alpha$ is a number such that all $A_\xi$ with $\xi < \alpha$ are different, then $\overline{\overline{\Gamma(\alpha)}} \leqslant \overline{\overline{\boldsymbol{R}}}$. Now let $\alpha$ be such that $\overline{\overline{\Gamma(\alpha)}} > \overline{\overline{\boldsymbol{R}}}$. Then there is $\beta < \alpha$ such that $A_{\beta+1} = A_\beta$, i.e. $G(A_\beta) = A_\beta$. This means that $A_\beta$ is a maximal set of $\boldsymbol{R}$.

The following statement will be proved in a similar way.

THEOREM 4.‡ *Let $A$ be an ordered set such that, for each linearly ordered $X \subset A$, there is $c \in A$ such that $x \leqslant c$ whatever $x \in X$ is. Then there is a maximal element in $A$.*

P r o o f. Let $a_0$ be an arbitrary element of $A$. If $a$ is not a maximal element of $A$ denote by $g(a)$ any $b \in A$ such that $a < b$; if $a$ is maximal, let $g(a) = a$. Similarly, for $Z \subset A$, let $h(Z)$ be an element $z$ of $A$ following all elements of $Z$; if such a $z$ does not exist, let us agree that $h(Z) = a_0$.

By Theorem 2 there exists a transfinite sequence

$$a_0 < a_1 < \ldots < a_\xi < \ldots$$

such that

$$a_{\xi+1} = g(a_\xi) \quad \text{and} \quad a_\lambda = h(Z_\lambda)$$

† See my paper in *Fundamenta Mathematicae* 3 (1922), p. 89. As shown in this paper, the theorem can also be proved without using ordinal numbers.

‡ This statement is frequently called *Zorn Lemma.*

where $Z_\lambda$ denotes the set of all $a_\xi$ such that $\xi < \lambda$ ($\lambda$ being a limit ordinal).

As in the preceding proof, we show the existence of $\beta$ such that $a_{\beta+1} = a_\beta$. It follows that $a_\beta$ is a maximal element in $A$.

**Exercises**

**1.** Prove that the conditions $\alpha < \Omega$ and $\beta < \Omega$ imply that $\alpha+\beta < \Omega$ and $\alpha \cdot \beta < \Omega$.

**2.** Every ordinal number is of the form $\lambda+n$, where $\lambda$ is a limit ordinal and $n$ is a natural number or zero.

*Hint*: Make use of the fact that in a well ordered set there does not exist an infinite sequence of the form $a_1 > a_2 > a_3 > \ldots$

**3.** Prove the following implications:

(a) $$(\alpha < \beta) \Rightarrow (\gamma+\alpha < \gamma+\beta),$$

(b) $$(\alpha \leqslant \beta) \Rightarrow (\alpha+\gamma \leqslant \beta+\gamma).$$

Does the condition $\beta > 0$ imply the inequality $\gamma < \beta+\gamma$?

**4.** Prove the distributive law:

$$\alpha \cdot (\beta+\gamma) = \alpha \cdot \beta + \alpha \cdot \gamma.$$

Show by means of an example that the formula $(\beta+\gamma) \cdot \alpha = \beta \cdot \alpha + \gamma \cdot \alpha$ is not true.

**5.** Prove that if $\alpha \geqslant \beta$ then there exists one and only one ordinal number $\gamma$ such that $\alpha = \beta+\gamma$ (we call the number $\gamma$ the *difference* $\alpha-\beta$ of the numbers $\alpha$ and $\beta$).

**6.** Prove that for every two ordinal numbers $\alpha \neq 0$ and $\beta$ there exists a pair of numbers $\delta$ and $\varrho < \alpha$ such that

$$\beta = \alpha \cdot \delta + \varrho$$

and that the numbers $\delta$ (*quotient*) and $\varrho$ (*remainder*) are uniquely determined.

**7.** A transfinite sequence $\varphi$ whose values are ordinals $< \alpha_0$ is said to be *continuous* if for every limit ordinal $\lambda < \alpha_0$ the following identity holds:

$$\varphi(\lambda) = \lim_{\xi<\lambda} \varphi(\xi).$$

Prove that the transfinite sequences $\varphi(\xi) = \alpha+\xi$ and $\varphi(\xi) = \alpha \cdot \xi$ (for $\alpha > 0$) are increasing and continuous.

**8.** Prove that every increasing transfinite sequence $\varphi$ satisfies the inequality $\xi \leqslant \varphi(\xi)$ for every $\xi$.

*Hint*: Assuming that the theorem is false, denote by $\alpha$ the least number such that $\varphi(\alpha) < \alpha$.

**9.** Let $\varphi$ be an increasing continuous transfinite sequence. Let us form the sequence

$$\alpha_0 = \alpha, \quad \alpha_1 = \varphi(\alpha_0), \quad \ldots, \quad \alpha_n = \varphi(\alpha_{n-1}), \quad \ldots$$

Let $\varkappa = \lim_{n<\omega} \alpha_n$. Prove that $\varphi(\varkappa) = \varkappa$ (under the assumption that the numbers under consideration belong to the domain of the function $\varphi$).

**10.** The number $\varkappa$ in Exercise 9 is said to be a *critical number* of the sequence $\varphi$. Find the critical numbers of the sequences

$$\varphi(\xi) = \alpha+\xi, \quad \varphi(\xi) = \alpha\cdot\xi, \quad \varphi(\xi) = \alpha^{\xi}.$$

**11.** Using the generalized principle of choice (see § 7) prove that every infinite cardinal number $\mathfrak{m}$ satisfies the inequality $\mathfrak{m} \geqslant \mathfrak{a}$.

**12.** Show that every proper ideal is contained in a maximal ideal.

*Hint*: Use Theorem 3 of § 8

# INTRODUCTION TO PART II

Topology is the study of those properties of spaces, sets, geometric figures, etc., which remain invariant relative to homeomorphisms (see Chapter XII, § 2). We call such properties topological invariants. For example, the property of a circle of separating the plane into two regions is a topological invariant; if we transform the circle into an ellipse or into the boundary of a triangle, this property is retained. On the other hand, the property of a curve of having a tangent line at every point is not a topological property; the circle has this property but the boundary of a triangle has not, although it may be obtained from the circle by means of a homeomorphism.

As can already be seen from the above example topology operates with more general concepts than analysis; differential properties of a given transformation are nonessential for topology, but continuity is essential. As a consequence, topology is often suitable for the solution of problems to which analysis cannot give the answer.

The generality of topological methods rests not only on the generality of the assumptions concerning the transformations considered but also on the generality of the sets considered to which these transformations are applied. These can be arbitrary point sets on the real line or in the plane, or in $n$-dimensional space, or still more general sets, provided only that they be sets for which—roughly speaking—it is possible to define the concept of open set, i.e. provided that they are topological spaces (see Chapter X). This generality has not only a methodological significance; in modern mathematics there is a trend to confer upon the set of objects considered in a given investigation (be these functions, sequences or curves) a topology, and hence—to a geometrization or rather to a topologization—of the investigation. This gives rise to numerous applications. Thus, e.g. theorems on the existence of a solution of certain types of differential equations can be expressed as theorems on the existence of invariant points of a function space (the space of continuous functions) under continuous transformations; these theorems can be proved by topological methods in a more general form and in a simpler way than was formerly done without the aid of topology.

How much more general ought the spaces considered in topology to be in order that they suffice for applications and yet, because of undue generality, not become too artificial? The answer to this question depends on the aims which a given topological work is to serve.

In this book we are mainly concerned with *topological spaces*; their definition is given in Chapter X. The Chapter IX on metric spaces and on Euclidean spaces has a rather auxiliary character: it serves for the interpretation of theorems and notions of general topology on simple and familiar examples. At the same time, the metric spaces—though they are simple and geometrically elementary—have many important applications (e.g. in differential equations and in functional analysis).

In Chapters X–XIII we give the fundamental concepts used in all parts of topology, and which are basic for *general topology*. The reader knows many of these concepts from analysis for the space of real or complex numbers (such as accumulation point, neighbourhood, closed set, and so on); this refers especially to Chapter XII on continuous functions. Theorems known from analysis, e.g. on uniform continuity, uniform convergence, the Darboux property, are proved here (and in Chapters XVI and XVII) under significantly more general hypotheses. This permits us to recognize the proper validity of these theorems (which also is of didactic significance).

In the further chapters (XIV–XVIII) we are gradually leaving the scope of general topology, confining ourselves to more specific spaces: we consider spaces with a countable base (and in particular metric separable spaces), complete spaces (with the Baire theorem and its consequences), compact spaces (which generalize the concept of closed bounded subsets of Euclidean space), connected spaces (connectedness is the precise formulation of the concept of the continuity of a set) and locally connected spaces (as it turns out, curves, surfaces, multi-dimensional varieties or manifolds, with which we have to deal in differential geometry, are as a rule locally connected continua). Many problems of that part of the book belong to *analytic topology* (devoted chiefly to the study of mappings).

Chapter XIX deals with the concept of dimension. This concept—even though it dates from antiquity (it appears in Euclid's *Elements*)—was properly defined only in recent times and this thanks to the use of topological methods.

We shall concern ourselves in more detail with the properties of the $n$-dimensional simplex, which is the fundamental concept of classical multi-dimensional geometry, in Chapter XX. In particular, we give a proof of the famous fixed point theorem, due to L. E. J. Brouwer (which has extensive applications in the theory of differential equations).

Chapter XXI, conceptually closely related to geometry, concerns theorems on the separation of the plane. A detailed proof is given here of the Jordan theorem, which is a classical theorem of analysis, and some important far-reaching generalizations due to Janiszewski and Eilenberg.

Most of the material contained in Part II (with the exception, for instance, of Chapter XX) belongs to *point-set topology.*

An introduction to *algebraic topology* is given in the *Supplement* written by Prof. Engelking.

In its initial stages, point-set topology and algebraic topology developed entirely independently and possessed completely different topics. Point-set topology, formerly called the theory of point sets, and concerning applications of set theory to arbitrary subsets of Euclidean space, was begun by G. Cantor, the creator of the theory of sets (*circa* 1880). Algebraic topology was created by H. Poincaré in the last years of the past century; its objects were $n$-dimensional polygons and polyhedra. Some unification of these two theories came rather late, about 40 years ago; this was, to a large degree, the work of P. S. Aleksandrov. This period was preceded by the transition from the investigation of subsets of Euclidean space in set-theoretic topology to the investigation of arbitrary topological spaces. This extension of the thematics of topology appeared to a significant degree in connection with the new mathematical investigations concerning the concept of function space and infinite-dimensional spaces introduced by Hilbert.

In the last forty years or so there has appeared an unusually rich flourishing of topology; many fundamental problems of topology have been solved, new methods created and new branches developed; such is the fascinating *differential topology*, constantly increasing in importance and interest. Topology, which was a conglomeration of loosely related theorems, became a systematic science, and topological methods penetrated into a great many other domains of mathematics.

The following list contains books which may be of interest to the reader who wishes to increase his knowledge of topology:

P. S. Aleksandrov and H. Hopf, *Topologie*, I, Edwards, Ann Arbor, 1945.

C. Berge, *Espaces topologiques*, Paris, 1959 (also an English translation).

K. Borsuk, *Theory of Retracts*, Monografie Matematyczne, Warszawa, 1967.

N. Bourbaki, *Topologie générale*, Actualités Scientifiques, Nos. 1045, 1084, 1142, 1143, 1235, Paris, 1949–1961.

D. C. J. Burgess, *Analytical Topology*, Van Nostrand, New York, 1966.

E. Čech, *Topological Spaces*, Praha, 1966.

G. Choquet, *Cours d'analyse*, t. II, Paris, 1964 (also an English translation).

J. Dugundji, *Topology*, Allyn-Bacon, 1966.

R. Engelking, *Outline of General Topology*, N.-Holland Publ. Comp. and PWN, Amsterdam–Warszawa, 1968.

S. A. Gaal, *Point Set Topology*, Academic Press, 1964.

F. Hausdorff, *Set Theory*, Chelsea, New York, 1957.

J. G. Hocking and Gail S. Young, *Topology*, Reading–London, 1961.

S. T. Hu, *Elements of General Topology*, Holden-Day, 1965.

—— *Homotopy Theory*, New York–London, 1959.

—— *Theory of Retracts*, Detroit, 1965.

W. Hurewicz and H. Wallman, *Dimension Theory*, Princeton, 1948.

J. L. Kelley, *General Topology*, Van Nostrand, New York, 1955.

H. J. Kowalsky, *Topologische Räume*, Basel–Stuttgart, 1961.

K. Kuratowski, *Topology*, vol. I, 1966, vol. II, 1968, Academic Press and PWN, New York–London–Warszawa,

S. Lefschetz, *Introduction to Topology*, Princeton Univ. Press, Princeton, 1949.

T. O. Moore, *Elementary General Topology*, Prentice-Hall, 1964.

J. Nagata, *Modern General Topology*, N.-Holland Publ. Comp., 1968.

M. H. A. Newman, *Elements of the Topology of Plane Sets of Points*, Cambridge Univ. Press, Cambridge, 1952.

G. Nöbeling, *Grundlagen der analytischen Topologie*, Springer, Berlin, 1954.

W. J. Pervin, *Foundations of General Topology*, Academic Press, 1964.

H. Schubert, *Topology*, Macdonald, London, 1968.

H. Seifert and W. Threlfall, *Lehrbuch der Topologie*, Chelsea, New York, 1947.

W. Sierpiński, *General Topology*, Univ. of Toronto Press, Toronto, 1952.

W. J. Thron, *Topological Structures*, Holt, Rinehart and Winston, 1966.

G. T. Whyburn, *Analytic Topology*, Coll. Public., New York, 1942.

—— *Topological Analysis*, Princeton, 1964.

R. L. Wilder, *Topology of Manifolds*, Coll. Public., New York, 1949.

We wish to quote also the following more elementary books:

P. S. Aleksandrov, *Einfachste Grundbegriffe der Topologie*, Berlin, 1932.

B. H. Arnold, *Intuitive Concepts in Elementary Topology*, Prentice-Hall, 1962.

J. D. Baum, *Elements of Point Set Topology*, Prentice-Hall, 1964.

D. Bushaw, *Elements of General Topology*, J. Wiley, New York, 1963.

W. D. Chinn and N. E. Steenrod, *First Concepts of Topology*, New York–Toronto, 1966.

E. T. Copson, *Metric Spaces*, Cambridge Tracts, 1968.

W. Franz, *Allgemeine Topologie*, Göschen, Berlin, 1960.

M. Mansfield, *Introduction to Topology*, Princeton, 1963.

E. M. Patterson, *Topology*, Interscience Publ., New York, 1956.

F. Simmons, *Introduction to Topology and Modern Analysis*, McGraw-Hill, New York, 1963.

G. L. Spencer and D. W. Hall, *Elementary Topology*, J. Wiley, New York, 1955.

R. Vaidyanathaswamy, *Treatise on Set Topology*, Part I, Indian Mathematical Society, Madras, 1947.

CHAPTER IX

# METRIC SPACES. EUCLIDEAN SPACES

## § 1. Metric spaces

Definition. A set $X$ is said to be a *metric space* if to every pair of its elements, i.e. to every pair of *points* $x$, $y$ belonging to the set $X$, there is assigned a real number $|x-y| \geqslant 0$, called the *distance from the point x to the point y*, which satisfies the following three conditions:

(1) $$|x-y| = 0 \textit{ if and only if } x = y,$$

(2) $$|x-y| = |y-x|,$$

(3) $$|x-y|+|y-z| \geqslant |x-z|;$$

the last condition is the so-called *triangle inequality*.

It follows immediately from this definition that every subset of a metric space is itself a metric space (the definition of distance remaining the same).

EXAMPLES. 1. Every set of real or complex numbers forms a metric space if the distance between two numbers $x$ and $y$ is understood to be the absolute value of the difference of these numbers. This justifies the symbol we are using for the distance.

2. Euclidean $n$-space, $\mathscr{E}^n$, whose points are sequences of $n$ real numbers $(x_1, x_2, \ldots, x_n)$, is a metric space under the usual definition of the distance from the point $x = (x_1, x_2, \ldots, x_n)$ to the point $y = (y_1, y_2, \ldots, y_n)$ given by the Pythagorean formula

(4) $$|x-y| = \Big\{ \sum_{i=1}^{n} |x_i-y_i|^2 \Big\}^{1/2}.$$

This same formula "metrizes" the *cartesian product* $X_1 \times X_2 \times \ldots$ $\ldots \times X_n$ of any $n$ metric spaces, $X_1, X_2, \ldots, X_n$.

3. *Hilbert space*. This space is the set of all sequences of real numbers $x = (x_1, x_2, \ldots, x_i, \ldots)$ such that the series $\sum_{i=1}^{\infty} x_i^2$ is

convergent. Here the distance between two such sequences is understood to be

$$(5) \qquad |x-y| = \left\{ \sum_{i=1}^{\infty} |x_i-y_i|^2 \right\}^{1/2}.$$

4. The set of continuous real valued functions defined on the closed interval $0 \leqslant x \leqslant 1$ forms a metric space if the distance between two functions $f$ and $g$ is defined by the formula

$$(6) \qquad |f-g| = \sup |f(x)-g(x)|.$$

R e m a r k. An arbitrary set can be considered to be a metric space if we assume that the distance between each pair of distinct points is 1.

## § 2. Diameter of a set. Bounded spaces. Bounded mappings

D e f i n i t i o n 1. The least upper bound of the distances $|x-y|$ between all pairs of points $x$ and $y$ in the metric space $X$ is called the *diameter* of the space $X$ and is denoted by the symbol $\delta(X)$. If $X$ is a circle or sphere, then its diameter $\delta(X)$ is the diameter in the usual sense.

Metric spaces with finite diameter are said to be *bounded.*

For example, the closed interval $0 \leqslant x \leqslant 1$ is bounded. The same is true of a square and the $n$-dimensional cube. On the other hand, the half-line $x \geqslant 0$, the real line, and the space $\mathscr{E}^n$ are examples of unbounded spaces.

D e f i n i t i o n 2. A mapping $f$: $X \to Y$ where $Y$ is a metric space, is called *bounded* if the set $f(X)$ is bounded.

THEOREM 1. *If $f$ and $g$ are bounded mappings of the (arbitrary) set $X$ into the metric space $Y$, their distance $|f-g|$ given by formula* (6) *is finite.*

P r o o f. Let $a$ be a given element of $X$. Then

$$|f(x)-g(x)| \leqslant |f(x)-f(a)|+|f(a)-g(a)|+|g(a)-g(x)|,$$

hence

$$|f-g| \leqslant \delta[f(X)]+|f(a)-g(a)|+\delta[g(X)].$$

THEOREM 2. *The set $\Phi(X, Y)$ of all bounded mappings $f$: $X \to Y$, where $X$ is an arbitrary set and $Y$ a metric space, is a metric space with distance defined by formula* (6).

This follows easily from Theorem 1.

## § 3. The Hilbert cube

Under the assumption that the spaces $X_1, X_2, \ldots, X_m, \ldots$ are uniformly bounded (i.e. the upper bound of their diameters is finite; see also Chapter XII, § 3, Remark 2), we define the distance between two points $x = (x_1, x_2, \ldots, x_m, \ldots)$ and $y = (y_1, y_2, \ldots, y_m, \ldots)$ of the infinite cartesian product $X_1 \times X_2 \times \ldots \times X_m \times \ldots$, by means of the formula

$$|x-y| = \sum_{m=1}^{\infty} (1/2^m)|x_m - y_m|. \tag{7}$$

We shall leave it to the reader to prove that the distance defined in this way satisfies conditions (1)–(3), i.e. that the space $X_1 \times X_2 \times \ldots$ is metric.

We denote the closed interval $0 \leqslant x \leqslant 1$ by $\mathscr{I}$. The space $\mathscr{H} = \mathscr{I} \times \mathscr{I} \times \ldots$ is called the *Hilbert cube*; it is a space all "coordinates" $x_m$ of whose points $x = (x_1, x_2, \ldots, x_m, \ldots)$ are contained in the closed interval $[0, 1]$. The space $\mathscr{H}$, or the infinite countable power of the closed interval $[0, 1]$, is clearly the natural generalization of the $n$-dimensional cube.

## § 4. Convergence of a sequence of points

We define the concept of the limit of a sequence of points, which is a fundamental concept in topology, by making use of the concept of the limit of a sequence of real numbers which is known from elementary analysis.

Definition. A sequence of points $p_1, p_2, \ldots, p_n, \ldots$ of a metric space is *convergent* to the point $p$ of this space if the sequence of real numbers $|p_n - p|$ is convergent to zero. We then call the point $p$ the *limit* of the sequence $p_1, p_2, \ldots, p_n, \ldots$ and we write $p = \lim\limits_{n \to \infty} p_n$.

Using the symbolism of logic, we write this definition in the following form:

$$(\lim_{n\to\infty} p_n = p) \equiv (\lim_{n\to\infty} |p_n - p| = 0) \tag{8}$$

$$\equiv \bigwedge_\varepsilon \bigvee_k \bigwedge_n [(n > k) \Rightarrow (|p_n - p| < \varepsilon)]. \tag{9}$$

A sequence of points does not need to converge. However, if it does converge, then there is just *one limit* of this sequence.

The definition of the convergence of a sequence of points in a metric space can be given in another form, very suitable for considerations in the sequel, by introducing the concept of ball.

An (*open*) *ball* with centre $p$, or more briefly $K(p, \varepsilon)$, is the set of points $x$ whose distance from the point $p$ is less than $\varepsilon$:

$$K(p, \varepsilon) = \{x\colon |x-p| < \varepsilon\}. \tag{10}$$

In the space of real numbers an open ball is an open interval and in the plane it is a circular disk without the boundary. Hence our terminology corresponds to Euclidean space.

Let us add that by replacing in (10) $<$ by $\leqslant$ we obtain the definition of a *closed* ball.

## § 5. Properties of the limit

THEOREM 1. *A necessary and sufficient condition that* $\lim\limits_{n\to\infty} p_n = p$, *is that every ball $K$ with centre $p$ contains all the points of the sequence $p_1, p_2, \ldots$, with perhaps the exception of a finite number (i.e. there exists a $k$ such that $p_n \in K$ for all $n > k$).*

In order to prove this we substitute $p_n \in K(p, \varepsilon)$ into formula (9) instead of $|p_n - p| < \varepsilon$ (which we can do by virtue of (10)).

THEOREM 2. *Every convergent sequence is bounded*; in other words: *the set of terms in a convergent sequence is bounded.*

For let $p = \lim\limits_{n\to\infty} p_n$ and let $Z$ be the set of terms of the sequence $p_1, p_2, \ldots, p_n, \ldots$ By virtue of our assumption there exists a $k$ such that for $n > k$ we have $|p_n - p| < 1$. Let $\varrho$ denote the maximal of the $k+1$ numbers

$$|p_1 - p|, |p_2 - p|, \ldots, |p_n - p|, 1.$$

Hence we have $|p_n - p| \leqslant \varrho$ for every $n$. Therefore

$$|p_n - p_m| \leqslant |p_n - p| + |p - p_m| \leqslant 2\varrho, \quad \text{i.e.} \quad \delta(Z) \leqslant 2\varrho.$$

The proofs of the following theorems do not deviate from the proofs given in elementary analysis for sequences of real numbers.

THEOREM 3. *If $p_n = p$ for $n = 1, 2, \ldots$, then* $\lim\limits_{n\to\infty} p_n = p$.

THEOREM 4 (ON SUBSEQUENCES). *If* $\lim_{n\to\infty} p_n = p$ *and* $k_1 < k_2 < \ldots$, *then*

$$\lim_{n\to\infty} p_{k_n} = p.$$

THEOREM 5. *Every sequence* $p_1, p_2, \ldots$, *which is not convergent to* $p$, *contains a subsequence none of whose subsequences is convergent to* $p$.

THEOREM 6. *Neither the convergence of a sequence nor its limit depend on the initial finite number of terms of this sequence.*

This means that the addition or the omission of a finite number of terms of a convergent sequence does not affect either its convergence or the value of its limit.

THEOREM 7. *If* $\lim_{n\to\infty} p_n = p = \lim_{n\to\infty} q_n$, *then the sequence* $p_1, q_1, p_2, q_2, \ldots$ *is convergent to* $p$.

## § 6. Limit in the cartesian product

Let $Z = X \times Y$ be the cartesian product of the metric spaces $X$ and $Y$.

THEOREM 1. *A necessary and sufficient condition that a sequence of points* $z_n = \langle x_n, y_n\rangle$ *of the space* $X\times Y$ *be convergent to the point* $z = \langle x, y\rangle$ *is that* $\lim_{n\to\infty} x_n = x$ *and* $\lim_{n\to\infty} y_n = y$.

P r o o f. Let $\lim_{n\to\infty} z_n = z$ and let $\varepsilon > 0$. Hence there exists $k$ such that $|z_n - z| < \varepsilon$ for $n > k$. But since

$$|z_n - z| = \{|x_n - x|^2 + |y_n - y|^2\}^{1/2} \geqslant |x_n - x|$$

(cf. § 1, (4)), we also have $|x_n - x| < \varepsilon$ for $n > k$, i.e. $\lim_{n\to\infty} x_n = x$.

In an analogous manner we can prove that $\lim_{n\to\infty} y_n = y$.

Let us assume conversely that $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} y_n = y$. Let $\varepsilon > 0$. Then there exists a $k$ such that for $n > k$ we have

$$|x_n - x| < \varepsilon \quad \text{and} \quad |y_n - y| < \varepsilon,$$

whence

$$|z_n - z| = \{|x_n - x|^2 + |y_n - y|^2\}^{1/2} < \varepsilon\sqrt{2}.$$

Therefore $\lim_{n\to\infty} z_n = z$.

THEOREM 2. *Let $X_1, X_2, \ldots, X_m, \ldots$ be uniformly bounded spaces* (see also Chapter XII, § 3, Remark 2). *Let $\varrho > \delta(X_m)$ for $m = 1, 2, \ldots$ Let $x^n = (x_1^n, x_2^n, \ldots, x_m^n, \ldots)$ for $n = 1, 2, \ldots$ be a point of the space $X_1 \times X_2 \times \ldots \times X_m \times \ldots$ (i.e. $x_m^n \in X_m$ for $m = 1, 2, \ldots$), metrized with the aid of formula* (7) *of* § 3. *A necessary and sufficient condition that this sequence be convergent to the point $x = (x_1, x_2, \ldots, x_m, \ldots)$ is that $\lim\limits_{n\to\infty} x_m^n = x_m$ for $m = 1, 2, \ldots$, i.e.*

$$(\lim_{n\to\infty} x^n = x) \equiv \bigwedge_m (\lim_{n\to\infty} x_m^n = x_m).$$

P r o o f. Let $\lim\limits_{n\to\infty} x_n = x$ and let $\varepsilon > 0$. Therefore, for a fixed $m$ there exists a $k$ such that

$$|x^n - x| < \varepsilon/2^m$$

for $n > k$.

Since, however,

$$(1/2^m)|x_m^n - x_m| \leqslant |x^n - x|$$

by (7), we have

$$|x_m^n - x_m| \leqslant 2^m |x^n - x| < 2^m \cdot \varepsilon/2^m = \varepsilon$$

for $n > k$.

This means that

$$\lim_{n\to\infty} x_m^n = x_m. \tag{11}$$

Let us assume next that formula (11) holds for $m = 1, 2, \ldots$ Let $\varepsilon > 0$. Let $i$ be an integer such that

$$1/2^i < \varepsilon. \tag{12}$$

Formula (11) for $m = 1, 2, \ldots, i$ implies that there exists $k$ such that for $n > k$ the inequalities

$$|x_1^n - x_1| < \varepsilon, \quad |x_2^n - x_2| < \varepsilon, \quad \ldots, \quad |x_i^n - x_i| < \varepsilon \tag{13}$$

hold. Therefore, because of (12) and (13),

$$|x^n - x| = \sum_{m=1}^{\infty} (1/2^m)|x_m^n - x_m|$$

$$= \sum_{m=1}^{i} (1/2^m)|x_m^n - x_m| + \sum_{m=i+1}^{\infty} (1/2^m)|x_m^n - x_m|$$

$$< \sum_{m=1}^{i} (\varepsilon/2^m) + \sum_{m=i+1}^{\infty} \delta(X_m)/2^m < \varepsilon + \varepsilon \cdot \varrho$$

for all $n > k$, i.e.

$$|x^n - x| < \varepsilon(1+\varrho).$$

Hence

$$\lim_{n\to\infty} x^n = x.$$

## § 7. Uniform convergence

This concept can be introduced in the same way as in elementary analysis.

Definition. Let $f_n\colon X \to Y$, $n = 1, 2, \dots$, where $X$ is an arbitrary set and $Y$ is metric. We say that the sequence $f_1, f_2, \dots$ converges *uniformly* to $f$ if

$$\bigwedge_\varepsilon \bigvee_k \bigwedge_x \bigwedge_{n\geq k} |f_n(x) - f(x)| < \varepsilon. \tag{14}$$

Let us consider (as in § 2) the space $\Phi(X, Y)$ of all bounded mappings $f\colon\ X \to Y$ with the distance defined by formula (6) of § 1.

The definition of the limit in a metric space (see § 4) implies that

$$(\lim_{n\to\infty} f_n = f) \equiv \lim_{n\to\infty} |f_n - f| = 0$$

$$\equiv \bigwedge_\varepsilon \bigvee_k \bigwedge_{n\geq k} (\sup |f_n(x) - f(x)| \leqslant \varepsilon)$$

$$\equiv \bigwedge_\varepsilon \bigvee_k \bigwedge_{n\geq k} \bigwedge_x \big(|f_n(x) - f(x)| < \varepsilon\big).$$

Thus we have the following theorem.

THEOREM 1. *In the space $\Phi(X, Y)$, the condition $\lim_{n\to\infty} f_n = f$ means that the sequence of mappings $f_1, f_2, \dots$ converges uniformly to the mapping $f$.*

THEOREM 2. *The limit of a uniformly convergent sequence of bounded mappings is bounded.*

Proof. Let $\varepsilon > 0$. Let $n$ be such that $|f_n - f| < \varepsilon$. Since

$$|f(x_1) - f(x_2)| \leqslant |f(x_1) - f_n(x_1)| + |f_n(x_1) - f_n(x_2)| + |f_n(x_2) - f(x_2)|,$$

it follows that

$$\delta[f(X)] \leqslant \delta[f_n(X)]+2\varepsilon.$$

**Exercises**

**1.** Let $\mathscr{E}^2$ be the complex number plane; for points $z, z' \in \mathscr{E}^2$ (where $z \neq z'$) let $||z-z'||$ be defined as follows: in case the line $zz'$ goes through the origin of the coordinate system, take $||z-z'|| = |z-z'|$, and in the contrary case take $||z-z'|| = |z|+|z'|$, where $|z|$ denotes as always the absolute value of $z$. Furthermore, let $||z-z|| = 0$. Prove that $||z-z'||$ can be treated as the distance of $z$ from $z'$, i.e. that it satisfies the conditions (1)–(3).

**2.** Show that if the sets $A$ and $B$ are not void and if $A \subset B$, then $\delta(A) \leqslant \delta(B)$.

**3.** Prove the inequality

$$\delta(A \cup B) \leqslant \delta(A)+\delta(B)$$

under the assumption that $A \cap B \neq \emptyset$.

**4.** A set of arbitrary elements is called an $\mathscr{L}^*$-*space* if to certain infinite sequences $p_1, p_2, \ldots$ of its elements (called convergent sequences) there corresponds an element $p = \lim_{n\to\infty} p_n$, called the limit of the considered sequence so that the Theorems 3–5 of § 5 hold true. Thus, metric spaces are $\mathscr{L}^*$-spaces. Show that Theorems 6 and 7 of § 5 hold in $\mathscr{L}^*$-spaces.

**5.** Let $\lim_{n\to\infty} p_n = p$ in an $\mathscr{L}^*$-space. Show that, if the sequence $q_1, q_2, \ldots$ is derived from the sequence $p_1, p_2, \ldots$ by finite repetition of its elements, then $\lim_{n\to\infty} q_n = p$.

CHAPTER X

# TOPOLOGICAL SPACES

## § 1. Definition. Closure axioms

A *topological space* is a set $X$ and a mapping assigning to each set $A \subset X$ a set $\bar{A} \subset X$ satisfying the following four axioms:

$$\text{(I)} \qquad \overline{A \cup B} = \bar{A} \cup \bar{B},$$

$$\text{(II)} \qquad A \subset \bar{A},$$

$$\text{(III)} \qquad \bar{\emptyset} = \emptyset,$$

$$\text{(IV)} \qquad \overline{(\bar{A})} = \bar{A}.$$

The elements of the space $X$ are called *points* and the set $\bar{A}$ is called the *closure* of $A$.

## § 2. Relations to metric spaces

We are going to show that every metric space (and in particular Euclidean space) can be regarded as a topological space if the following definition of the closure is assumed.

D e f i n i t i o n. $p$ belongs to $\bar{A}$ if and only if $p$ is the limit of a sequence of points belonging to $A$.

First, let us show the following theorem.

THEOREM 1. *$p \in \bar{A}$ if and only if*

$$(1) \qquad K \cap A \neq \emptyset$$

*for every open ball K of centre p.*

For, if $p = \lim_{n\to\infty} p_n$, where $p_n \in A$, then $K \cap A \neq \emptyset$ by virtue of Theorem 1 of Chapter IX, § 5.

Next, let us assume that condition (1) is satisfied for every $K$. Let $K_n = K(p, 1/n)$. By assumption, $K_n \cap A \neq \emptyset$, i.e. for every $n$ there exists a point $p_n \in K_n \cap A$. By the definition of $K_n$ we have $|p_n - p| < 1/n$, and therefore $p = \lim_{n\to\infty} p_n$. Inasmuch as $p_n \in A$ we have $p \in \bar{A}$.

Remark. The above theorem can be formulated as follows: *A necessary and sufficient condition that the point p does not belong to the set $\overline{A}$ is, that there exist a ball with centre p which is disjoint from the set A.*

THEOREM 2. *Let X be a metric space. If the above definition of closure is assumed, X becomes a topological space.*

We have to show that the conditions (I)–(IV) are fulfilled.

Proof of property (I). Let $p \in \overline{A \cup B}$. This means that $p = \lim_{n\to\infty} p_n$, where $p_n \in A \cup B$. It follows that there exists a sequence of indices $k_1 < k_2 < \ldots$ such that for every $n$ we have $p_{k_n} \in A$ or for every $n$ we have $p_{k_n} \in B$. Since $p = \lim_{n\to\infty} p_{k_n}$ (by virtue of Theorem 4 of Chapter IX, § 5) in the first case we obtain $p \in \overline{A}$, and in the second case $p \in \overline{B}$. Hence, in every case we have $p \in \overline{A} \cup \overline{B}$.

We have thus proved that

$$\overline{A \cup B} \subset \overline{A} \cup \overline{B}. \tag{2}$$

In order to prove the converse inclusion we shall show that

$$(A \subset B) \Rightarrow (\overline{A} \subset \overline{B}). \tag{I'}$$

If $p \in \overline{A}$, then $p = \lim_{n\to\infty} p_n$, where $p_n \in A$. Because of the inclusion $A \subset B$ it follows that $p_n \in B$, and hence $p \in \overline{B}$.

Since $A \subset A \cup B$ and $B \subset A \cup B$, we deduce from (I′) that

$$\overline{A} \subset \overline{A \cup B} \quad \text{and} \quad \overline{B} \subset \overline{A \cup B},$$

and hence, adding these two inclusions memberwise, we obtain

$$\overline{A} \cup \overline{B} \subset \overline{A \cup B}. \tag{3}$$

Inclusions (2) and (3) yield (I).

Proof of inclusion (II). It suffices to note that if $p \in A$ then $p = \lim_{n\to\infty} p_n$, where $p_n = p$ for $n = 1, 2, \ldots$ (see Theorem 3 of Chapter IX, § 5).

Formula (IV) remains to be proved. By virtue of inclusion (II) we have $\overline{A} \subset \overline{(\overline{A})}$. Therefore, it suffices to prove that $\overline{(\overline{A})} \subset \overline{A}$.

Hence, let $p \in \overline{(\overline{A})}$. By virtue of Theorem 1 for every ball $K$ of centre $p$ we have $K \cap \overline{A} \neq \emptyset$. Hence, let $q \in K \cap \overline{A}$. Let us choose a ball $L$ of centre $q$ such that $L \subset K$. Since $q \in \overline{A}$, $L$ is a ball with centre $q$, and hence (by virtue of Theorem 1) we have $L \cap A \neq \emptyset$. But since $L \subset K$, we therefore have $(L \cap A) \subset (K \cap A)$, whence $K \cap A \neq \emptyset$. We deduce from this that $p \in \overline{A}$ (by virtue of the very same Theorem 1).

## § 3. Further algebraic properties of the closure operation

Let $X$ denote a topological space. The following formulas hold.

1. $$(A \subset B) \Rightarrow (\overline{A} \subset \overline{B}).$$

Proof (this formula denoted by (I′) was proved for metric spaces in § 2). Obviously

$$(A \subset B) \equiv (B = A \cup B).$$

Hence $\overline{B} = \overline{A \cup B}$ and by (I) $\overline{B} = \overline{A} \cup \overline{B}$. But this means that $\overline{A} \subset \overline{B}$.

2. $$\overline{A} - \overline{B} \subset \overline{A - B}.$$

Proof. $A \cup B = (A - B) \cup B$, and therefore

$$\overline{A \cup B} = \overline{(A - B) \cup B}.$$

From this, by virtue of formula (I), we have $\overline{A} \cup \overline{B} = \overline{A - B} \cup \overline{B}$ and hence $\overline{A} \subset \overline{A - B} \cup \overline{B}$, whence $\overline{A} - \overline{B} \subset \overline{A - B}$.

3. $$\overline{A \cap B} \subset \overline{A} \cap \overline{B}.$$

Proof. Since $A \cap B \subset A$ and $A \cap B \subset B$, we have, by virtue of (I′), $\overline{A \cap B} \subset \overline{A}$ and $\overline{A \cap B} \subset \overline{B}$, and therefore $\overline{A \cap B} \subset \overline{A} \cap \overline{B}$.

More generally, the following formula is valid:

4. $$\overline{\bigcap_t A_t} \subset \bigcap_t \overline{A_t},$$

where the variable $t$ ranges over an arbitrary set $T$.

P r o o f. Since for every $s \in T$ we have $\bigcap_t A_t \subset A_s$, hence by virtue of formula 1 we have $\overline{\bigcap_t A_t} \subset \overline{A}_s$, and from this we get $\overline{\bigcap_t A_t} \subset \bigcap_s \overline{A}_s$. Replacing the index $s$ by $t$ we obtain formula 4.

5. $$\bigcup_t \overline{A}_t \subset \overline{\bigcup_t A_t}.$$

P r o o f. For every $s$ we have $A_s \subset \bigcup_t A_t$, and hence by virtue of formula 1 we have $\overline{A}_s \subset \overline{\bigcup_t A_t}$, and $\bigcup_s \overline{A}_s \subset \overline{\bigcup_t A_t}$. From this we obtain formula 5.

6. $$\overline{X} = X.$$

This follows directly from axiom (II).

## § 4. Closed sets. Open sets

D e f i n i t i o n s. A set is said to be *closed* if $\overline{A} = A$, that is (because of axiom (II)) if $\overline{A} \subset A$.

A set $A$ is said to be *open* if its complement is closed, or in other words, if $A = X - \overline{X - A}$, where $X$ is the entire space.

Thus, in a metric space, the condition for a set $A$ to be closed can be expressed by the implication

$$(\lim_{n\to\infty} x_n = x)(x_n \in A) \Rightarrow (x \in \overline{A}).$$

The condition for a set $A$ to be open (in a metric space) is that each point of $A$ belongs to an open ball contained in $A$ (compare § 2, Remark).

EXAMPLES. 1. The null set is a closed set, i.e. $\overline{\emptyset} = \emptyset$ (§ 1, axiom (III)); the entire space is a closed set (§ 3, property 6). It also follows from this that the null set and the entire space are open sets.

2. In the space of real numbers the closed interval $a \leqslant x \leqslant b$ is a closed set. Our terminology is therefore in agreement with the terminology used in analysis. On the other hand, the open interval $a < x < b$ is an open set (which is not closed).

3. If $f$ is a continuous real valued function defined on the closed interval $a \leqslant x \leqslant b$, then this function, i.e. the set of points

$$A = \{\langle x, y\rangle: [y = f(x)]\,(a \leqslant x \leqslant b)\},$$

is a closed set.

For, let $p \in \bar{A}$, i.e. $p = \lim_{n\to\infty} p_n$, where $p_n \in A$. The points $p_n$ are therefore of the form

$$p_n = \langle x_n, f(x_n)\rangle, \tag{4}$$

$$a \leqslant x_n \leqslant b. \tag{5}$$

Let $p = \langle x, y\rangle$. Since $p = \lim_{n\to\infty} p_n$, we have (compare Theorem 1 of Chapter IX, § 6)

$$\lim_{n\to\infty} x_n = x, \tag{6}$$

$$\lim_{n\to\infty} f(x_n) = y. \tag{7}$$

It follows from (5) and (6) that $a \leqslant x \leqslant b$.

But because of the continuity of the function $f$, it follows from (6) that

$$\lim_{n\to\infty} f(x_n) = f(x),$$

and hence $y = f(x)$ by virtue of (7), i.e. $p = \langle x, f(x)\rangle$ and by the definition of the set $A$ we have $p \in A$.

We have thus proved that $\bar{A} \subset A$, i.e. that $A$ is closed.

Remark. As seen in Example 1, every nonvoid space $X$ contains two closed-open subsets, namely $\emptyset$ and $X$. A space which contains no other closed-open subset is called *connected* (connected spaces will be studied in Chapter XVII).

## § 5. Operations on closed sets and open sets

THEOREM 1. *The union of two closed sets is a closed set.*

For, if the sets $A$ and $B$ are closed, i.e. $\bar{A} = A$ and $\bar{B} = B$, then

$$\overline{A \cup B} = \bar{A} \cup \bar{B} = A \cup B.$$

This theorem can be generalized (by induction) to an arbitrary *finite* number of sets. The union of an infinite number of closed sets may be a non closed set; if e.g. $A_n = \{1/n\}$ then the union $A_1 \cup A_2 \cup \ldots$ is not a closed set (in the space of real numbers), since the point 0 does not belong to it but it belongs to its closure.

THEOREM 2. *The intersection of an arbitrary number of closed sets is a closed set.*

In fact, if the sets $A_t$ are closed, i.e. $\overline{A_t} = A_t$, then by formula 4 of § 3, we have

$$\overline{\bigcap_t A_t} \subset \bigcap_t \overline{A_t} = \bigcap_t A_t,$$

and hence the set $\bigcap_t A_t$ is closed.

THEOREM 1′. *The intersection of a finite number of open sets is an open set.*

THEOREM 2′. *The union of an arbitrary number of open sets is an open set.*

These properties follow from Theorems 1 and 2 using De Morgan formulas (see Chapter II, § 4, (30) and Chapter IV, § 2, (4)):

$$X-(A \cap B) = (X-A) \cup (X-B), \quad X-\bigcup_t A_t = \bigcap_t (X-A_t).$$

For, if the sets $A$ and $B$ are open, then the sets $X-A$ and $X-B$ are closed, and hence the set $X-(A \cap B) = (X-A) \cup (X-B)$ is also closed, i.e. the set $A \cap B$ is open. The generalization of the theorems to the case of an arbitrary finite number of sets is immediate.

If the sets $A_t$ are open, i.e. the set $X-A_t$ are closed, then the set $X-\bigcup_t A_t = \bigcap_t (X-A_t)$ is closed, and hence the set $\bigcup_t A_t$ is open.

Remark. Theorems 1, 1′ and 2, 2′ are examples of a *duality* in topology: to every theorem on closed sets there corresponds, by virtue of the De Morgan formulas, a theorem on open sets, and conversely.

THEOREM 3. *The set $\overline{A}$ is the smallest closed set containing $A$; in other words, $\overline{A}$ is the intersection of all closed $F$ such that $A \subset F$.*

Equivalently

(8) $p \in \overline{A} \equiv \bigwedge_F (A \subset F) \Rightarrow (p \in F)$, *whenever $F$ is closed*;

otherwise stated

(9) $p \in \overline{A} \equiv \bigwedge_G (p \in G) \Rightarrow (A \cap G \neq \emptyset)$, *whenever $G$ is open.*

Proof. Let $p \in \overline{A}$ and $A \subset F$. Hence $\overline{A} \subset \overline{F} = F$ and consequently $p \in F$.

Conversely, if $(A \subset F) \Rightarrow (p \in F)$, we obtain $p \in \overline{A}$ substituting $\overline{A}$ for $F$ (according to axiom (IV)).

(9) is equivalent to (8), since the implication $(A \subset F) \Rightarrow (p \in F)$ is equivalent to $(p \notin F) \Rightarrow (A \not\subset F)$, hence to

$$(p \in X-F) \Rightarrow [A \cap (X-F) \neq \emptyset];$$

thus it is sufficient to substitute the variable $G$ (open) for $F$ (closed).

## § 6. Interior points. Neighbourhoods

Definition 1. The set $\mathrm{Int}(A) = X-\overline{X-A}$ is called the *interior* of the set $A$.

Obviously, the condition $p \in X-\overline{X-A}$ means that $p \notin \overline{X-A}$. Therefore (see § 5, (9)), $p$ is an *interior point* of $A$ (i.e. $p \in \mathrm{Int}\,(A)$) if and only if there is an open $G$ containing $p$ and such that $G \cap (X-A) = \emptyset$.

It follows that $A$ is open iff $A = \mathrm{Int}(A)$.

The interior operation is dual to the closure operation. Thus the following statements (which can be easily proved) are dual to the statements (I)–(IV) of § 1:

(I′) $$\mathrm{Int}\,(A \cap B) = \mathrm{Int}\,(A) \cap \mathrm{Int}\,(B),$$

(II′) $$\mathrm{Int}\,(A) \subset A,$$

(III′) $$\mathrm{Int}\,(X) = X,$$

(IV′) $$\mathrm{Int}\,[\mathrm{Int}\,(A)] = \mathrm{Int}\,(A).$$

It should be noted that the interior operation can be taken as the *primitive term* in the definition of topological space instead of the closure operation.

Also one easily sees that $\mathrm{Int}\,(A)$ is the *largest open* set contained in $A$ (this is dual to Theorem 3 of § 5).

Definition 2. A set $A$ is said to be a *neighbourhood* of a point $p$ if $p \in \mathrm{Int}(A)$, i.e. if $p$ is an interior point of the set $A$.

Hence an open set is a neighbourhood of each of its points. Every neighbourhood of the point $p$ contains an open neighbourhood of the point $p$, namely its interior.

We say, more generally, that *$A$ is a neighbourhood of the set $B$ if $B \subset \mathrm{Int}(A)$.*

Every set containing a neighbourhood of $p$ is itself a neighbourhood of $p$. The intersection of two neighbourhoods of $p$ is a neighbourhood of $p$ (by (I')). Thus the *family of all neighbourhoods of $p$ is a filter.*

Definition 3. The set $\mathrm{Fr}(A) = \overline{A} \cap \overline{X-A}$ is called the *boundary* of $A$.

Clearly, every neighbourhood of a point $p \in \mathrm{Fr}(A)$ intersects $A$ and $X-A$.

EXAMPLES. 1. The interior of the closed interval $a \leqslant x \leqslant b$ in the space of real numbers is the open interval $a < x < b$ and its boundary is the set consisting of its endpoints $a$ and $b$.

2. The interior of the closed disk $\{x: |x-p| \leqslant \varrho\}$ on the plane is the open disk $\{x: |x-p| < \varrho\}$ and its boundary is the circumference $\{x: |x-p| = \varrho\}$.

## § 7. The concept of open set as the primitive term of the notion of topological space

Let $X$ be an arbitrary set and let $\boldsymbol{R}$ be a family of subsets of $X$ such that

(i) the union of an arbitrary number of sets belonging to $\boldsymbol{R}$ belongs to $\boldsymbol{R}$,

(ii) the intersection of a finite number of sets belonging to $\boldsymbol{R}$ belongs to $\boldsymbol{R}$,

(iii) $\emptyset \in \boldsymbol{R}$,

(iv) $X \in \boldsymbol{R}$.

THEOREM. *$\boldsymbol{R}$ gives a topological structure to $X$ by assuming that $\boldsymbol{R}$ is the family of open sets of $X$.*

More precisely: define the closure operation by the condition (9) of § 5 where $G$ ranges over $\boldsymbol{R}$; then the axioms (I)–(IV) are fulfilled. Moreover, the family of sets which are open relative to this definition (according to the definition given in § 4) is identical with $\boldsymbol{R}$.

Proof. In order to show that the above defined closure satisfies (I), put $p \in \overline{A \cup B}$. We have to show that either $p \in \overline{A}$ or $p \in \overline{B}$. Suppose the contrary is true. Then there exist by (9) two sets $G \in \boldsymbol{R}$ and $H \in \boldsymbol{R}$ such that $p \in G$, $p \in H$, $G \cap A = \emptyset$ and $H \cap B = \emptyset$. It follows that

$$p \in (G \cap H) \quad \text{and} \quad (G \cap H) \cap (A \cap B) = \emptyset.$$

Hence $p \notin \overline{A \cup B}$ because by (ii) $(G \cap H) \in \boldsymbol{R}$.

Thus $\overline{A \cup B} \subset \bar{A} \cup \bar{B}$. The converse inclusion follows directly from (9). The case of conditions (II) and (III) is similar.

Finally, in order to prove (IV), it is sufficient to show that $(p \in \bar{\bar{A}}) \Rightarrow (p \in \bar{A})$. Now let $p \in G \in \boldsymbol{R}$; we have to show that $A \cap G \neq \emptyset$. Substituting in (9) $\bar{A}$ for $A$, we get $\bar{A} \cap G \neq \emptyset$. Let $q \in \bar{A} \cap G$. It follows by (9) (where $p$ has to be replaced by $q$) that $A \cap G \neq \emptyset$.

We are now going to show that the family of open sets in the sense of the closure defined according to the condition (9) (with $G \in \boldsymbol{R}$) is identical with $\boldsymbol{R}$. Let $A$ be open. Hence by (9)

$$(p \in A) \equiv (p \notin \overline{X-A}) \equiv \bigvee_G (p \in G \in \boldsymbol{R})[(X-A) \cap G = \emptyset].$$

Thus, $A$ is the union of some elements of $\boldsymbol{R}$ (namely, of elements contained in $A$), and therefore $A \in \boldsymbol{R}$ (by (i)).

Conversely, if $A \in \boldsymbol{R}$, $A$ is obviously open.

R e m a r k 1. Another equivalent form of defining topological space (dual to the preceding) consists in assuming the term *closed* set as the primitive term. One will only need to replace in the conditions (i) and (ii) "union" by "intersection" and vice versa.

R e m a r k 2. Conditions (iii) and (iv) follow from (i) and (ii) if we agree that the number of elements of $\boldsymbol{R}$ may be zero.

## § 8. Base and subbase

D e f i n i t i o n 1. A family $\boldsymbol{B}$ of open subsets of the space $X$ is called its (open) *base* if every open $G \subset X$ is the union of a certain number of members of $\boldsymbol{B}$.

A family $\boldsymbol{S}$ of open subsets of $X$ is called its (open) *subbase* if the family of all finite intersections of members of $\boldsymbol{S}$ is a base of $X$.

Thus the family $\boldsymbol{R}$ of all open subsets of $X$ is generated by $\boldsymbol{B}$ by means of the union operation. $\boldsymbol{R}$ is generated by $\boldsymbol{S}$ by means of two operations: the finite intersection operation and the union operation.

R e m a r k s. The notions of a base and of a subbase lead to a very general method of introducing topology in an arbitrary set $X$. Namely if $\boldsymbol{R}$ fulfills condition (ii) of § 7, then $X$ becomes a topological space if we declare that $\boldsymbol{R}$ has to be its base.

Similarly, if $\boldsymbol{F}$ is an arbitrary family of subsets of $X$, then $X$ becomes a topological space if we assume that $\boldsymbol{F}$ is its subbase.

This follows easily from the formula

$$\left(\bigcup_s A_s\right) \cap \left(\bigcup_t B_t\right) = \bigcup_{s,t} (A_s \cap B_t)$$

where $A_s \in \boldsymbol{B}$ and $B_t \in \boldsymbol{B}$ (hence $(A_s \cap B_t) \in \boldsymbol{B}$).

Definition 2. A space having a countable base (but having no finite base) is said to be of *weight* $\mathfrak{a}$.

Generally, the *weight of a space* is the least cardinal $\mathfrak{m}$ such that the space has a base of cardinality $\mathfrak{m}$.

Definition 3. A (metric separable) space is called *0-dimensional* if it contains a base composed of sets which are simultaneously closed and open.

EXAMPLES. In the space $\mathscr{E}$ the family of all open intervals with rational endpoints is a base. The open rays $x > r$ and $x < r$ form a subbase of $\mathscr{E}$.

In the space $\mathscr{E}^n$ the family of all open balls $K(p, r)$ where $r$ is rational and $p$ has rational coordinates is a base.

The spaces: $\mathscr{Z}$ (of integers), $\mathscr{R}$ of rational numbers, $\mathscr{N}$ of irrational numbers are 0-dimensional.

## § 9. Relativization. Subspaces

Each subset $E$ of a topological space $X$ can be regarded as a topological space when we assume that the set $Q \subset E$ is *open relative to* $E$ if $Q$ is the intersection of $E$ and of a set $G$ open in $X$. It is easily seen that the family of open sets relative to $E$ satisfies conditions (i)–(iv) of § 7, and thus $E$ becomes a topological space (with the topology *induced* by the topology of $X$).

It follows that a subset $P$ of $E$ is *closed relative to* $E$ if and only if it can be written in the form $P = E \cap F$ where $F$ is closed in $X$. The closure of $A \subset E$ relative to $E$ equals $E \cap \overline{A}$.

## § 10. Comparison of topologies

Given a set $X$, there are—generally speaking—various ways of introducing topology in $X$. One can, for instance, assume that every subset of $X$ is open; this is the topology called *discrete*. One can, on the contrary, assume that the only two open sets are the void set and the whole space; this is the *coarsest* topology, the first one is the *richest* (the *finest*) topology.

The totality of all topologies of a given set $X$ can be ordered. Namely denote by $(X, \boldsymbol{R})$ the topology of $X$ having $\boldsymbol{R}$ as the family of all open sets and let us agree that

$$[(X, \boldsymbol{R}_1) \leqslant (X, \boldsymbol{R}_2)] \equiv (\boldsymbol{R}_1 \subset \boldsymbol{R}_2).$$

Thus $(X, \boldsymbol{R}_1)$ is coarser than (or equal) $(X, \boldsymbol{R}_2)$. Of course, the family of closed sets of the first is contained in the family of closed sets of the second; the closure of a set $A$ in the first topology contains its closure in the second.

## § 11. Cover of a space

Definition. A family $\boldsymbol{C}$ of open subsets of $X$ is called an (open) *cover* of $X$ if $X = \bigcup \boldsymbol{C}$, i.e. if each point of $X$ belongs to some member of $\boldsymbol{C}$ (compare Chapter IV, § 5).

$\boldsymbol{C}_1$ is called a *subcover* of $\boldsymbol{C}$ if $\boldsymbol{C}_1 \subset \boldsymbol{C}$.

The concept of a cover gives rise to a number of important topological notions which will be studied later.

For instance, $X$ is said to be *compact*, if every cover of $X$ contains a finite subcover. $X$ is called *countably compact* if every countable cover contains a finite subcover. $X$ is called a *Lindelöf space* if every cover contains a countable subcover.

THEOREM. *Let $\boldsymbol{B}$ be an open base of $X$ and $\boldsymbol{C}$ an open cover of $X$. Then there exists a cover $\boldsymbol{R}$ which is a refinement* (see Chapter IV, § 5) *of $\boldsymbol{C}$ such that $\boldsymbol{R} \subset \boldsymbol{B}$.*

Proof. Let $\boldsymbol{C} = \{G_t\}$ and let $\{H_s\}$ be the family of members of $\boldsymbol{B}$ such that for each $s$ there is $t(s)$ such that

$$(1) \qquad H_s \subset G_{t(s)}.$$

$\{H_s\}$ is a cover of $X$. For let $p$ be a given point of $X$. Since $\{G_t\}$ is a cover of $X$, there is $t_0$ such that $p \in G_{t_0}$, and since $\boldsymbol{B}$ is a base of $X$ there is $s_0$ such that $p \in H_{s_0} \subset G_{t_0}$. Hence $H_{s_0}$ is a member of $\{H_s\}$.

Remark. Let $\boldsymbol{B}$ be an open base of the space $X$. If every cover of $X$ composed of elements belonging to $\boldsymbol{B}$ contains a finite (resp. countable) cover, then $X$ is compact (resp. is a Lindelöf space); in other words, in order to show the compactness or the Lindelöf property of a space $X$ we can limit ourselves to the consideration of covers contained in $\boldsymbol{B}$.

**Exercises**

**1.** Prove that if the set $G$ is open, then the following rules are valid for every set $A$:

(a) $$G \cap \bar{A} \subset \overline{G \cap A},$$

(b) $$\overline{G \cap \bar{A}} = \overline{G \cap A}.$$

**2.** Prove the formulas:

(a) $$A \subset B \quad \text{implies} \quad \operatorname{Int}(A) \subset \operatorname{Int}(B),$$

(b) $$\bigcup_t \operatorname{Int}(A_t) \subset \operatorname{Int}\left(\bigcup_t A_t\right),$$

(c) $$\operatorname{Fr}(A) = A \cap \overline{X - A} \cup (\bar{A} - A),$$

(d) $$\bar{A} = A \cup \operatorname{Fr}(A),$$

(e) $$\operatorname{Fr}(A \cup B) \cup \operatorname{Fr}(A \cap B) \cup (\operatorname{Fr}(A) \cap \operatorname{Fr}(B)) = \operatorname{Fr}(A) \cup \operatorname{Fr}(B),$$

(f) $$\operatorname{Fr}[\operatorname{Int}(A)] \in \operatorname{Fr}(A),$$

(g) $$\operatorname{Int}(A) \cap \operatorname{Fr}(A) = \emptyset,$$

(h) $$\operatorname{Fr}\{\operatorname{Fr}[\operatorname{Fr}(A)]\} = \operatorname{Fr}[\operatorname{Fr}(A)].$$

**3.** Prove the equivalences:

$$(A \text{ is closed}) \equiv (\operatorname{Fr}(A) = A \cap \overline{X - A}),$$

$$(A \text{ is open}) \equiv (\operatorname{Fr}(A) = \bar{A} - A).$$

$$(A \text{ is a difference of two closed sets}) \equiv (\bar{A} - A \text{ is closed}).$$

**4.** Let $\{A_t\}$ be an arbitrary family of sets open relative to their union. Prove that

(i) $$\operatorname{Int}\left(\bigcup_t A_t\right) = \bigcup_t \operatorname{Int}(A_t),$$

(ii) $$\overline{\operatorname{Int}\left(\overline{\bigcup_t A_t}\right)} = \overline{\bigcup_t \operatorname{Int}(\bar{A}_t)}.$$

**5.** Prove that the following two properties of a topological space are equivalent:

(i) the closure of each open set is open (a space with this property is called *extremally discontinuous*),

(ii) whenever two open sets are disjoint, then their closures are disjoint.

**6.** Suppose that $X$ contains a countable base. Then every base contains a countable subfamily which itself is a base.

**7.** A family of sets $\{A_t\}$ is called *locally finite*, if each point $p$ is contained in an open set $G$ which intersects only a finite number of sets $A_t$. Show that

$$\overline{\bigcup_t A_t} = \bigcup_t \overline{A_t}.$$

**8.** Let $C_1$ and $C_2$ be covers of $X$. Recall that $C_2$ is a *refinement* of $C_1$ if each member of $C_2$ is a subset of a member of $C_1$. Write in this case $C_1 \leqslant C_2$. Show that this relation is a quasi-order (see Chapter VII, § 1) in the totality of all covers of $X$ and that this totality is directed by this relation.

R e m a r k. $X$ is called *paracompact* if each cover has a locally finite refinement. Each metric space is paracompact (A. H. Stone theorem).†

**9.** Two sets $A$ and $B$ are said to be *separated* if

$$\overline{A} \cap B = \emptyset = A \cap \overline{B}.$$

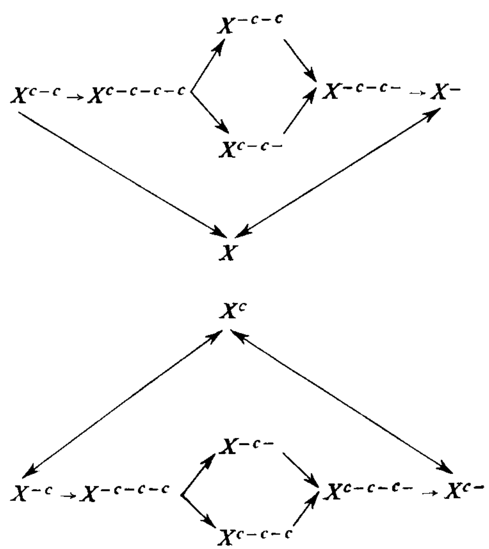
Show that $A$ and $B$ are separated iff they are disjoint and closed in $A \cup B$.

**10.** If two subsets $A$ and $B$ of $X$ are separated, the set $X-(A \cup B)$ is said to *separate A from B.*

Show that $\mathrm{Fr}(A)$ separates $\mathrm{Int}(A)$ from $\mathrm{Int}(X-A)$.

**11.** Prove that if we apply to a set $A \subset X$ two operations, $\overline{A}$ and $X-A$ then the maximal number of sets that we can obtain is 14. Namely (write $A^-$ instead of $\overline{A}$ and $A^c$ instead of $X-A$): $A, A^-, A^c, A^{-c}, A^{c-}, \ldots, A^{-c-c-c}$, $A^{c-c-c-c}$.

Verify that all the general valid inclusions are exhibited in the following table:‡




**12.** Prove that in a metric space $\delta(\overline{A}) = \delta(A)$.

† A simple proof of this theorem is given by M. E. Rudin in *Proc. Amer. Math. Soc.* **20** (1969), p. 603.

‡ For typographic reasons we write $\rightarrow$ instead of $\subset$ .

**13.** Let $X$ be a metric space. We say that a point $p \in X$ belongs to the *lower topological limit* of the sequence of sets $A_1, A_2, \ldots$ contained in $X$, i.e.

$$p \in \operatorname*{Li}_{n \to \infty} A_n$$

if every neighbourhood of $p$ intersects all $A_n$ with sufficiently great indices.

Prove the following formulas:

1. $\overline{\mathrm{Li} A_n} = \mathrm{Li} A_n = \mathrm{Li} \overline{A_n}$,
2. $A_n \subset B_n \Rightarrow \mathrm{Li} A_n \subset \mathrm{Li} B_n$,
3. $\mathrm{Li} A_n \cup \mathrm{Li} B_n \subset \mathrm{Li}(A_n \cup B_n)$,
4. $\mathrm{Li}(A_n \cap B_n) \subset (\mathrm{Li} A_n) \cap (\mathrm{Li} B_n)$,
5. $\liminf A_n \subset \mathrm{Li} A_n$ (comp. Exercise 6 of Chapter IV),
6. $\mathrm{Li}(A_n \times B_n) = \mathrm{Li} A_n \times \mathrm{Li} B_n$.

**14.** Let $X$ be as before a metric space. We say that a point $p \in X$ belongs to the *upper topological limit* of the sequence of sets $A_1, A_2, \ldots$ contained in $X$, i.e.

$$p \in \operatorname*{Ls}_{n \to \infty} A_n$$

if every neighbourhood of $p$ intersects $A_n$ for infinitely many $n$.

Prove the following formulas:

1. $\overline{\mathrm{Ls} A_n} = \mathrm{Ls} A_n = \mathrm{Ls} \overline{A_n}$,
2. $A_n \subset B_n \Rightarrow \mathrm{Ls} A_n \subset \mathrm{Ls} B_n$,
3. $\mathrm{Ls}(A_n \cup B_n) = \mathrm{Ls} A_n \cup \mathrm{Ls} B_n$,
4. $\mathrm{Ls}(A_n \cap B_n) \subset \mathrm{Ls} A_n \cap \mathrm{Ls} B_n$,
5. if $A_n = A$ then $\mathrm{Ls} A_n = \overline{A}$,
6. $\mathrm{Ls} A_n = \overline{\bigcap_{n=0}^{\infty} \bigcup_{k=0}^{\infty} A_{n+k}}$,
7. $\mathrm{Ls}(A_n \times B_n) \subset \mathrm{Ls} A_n \times \mathrm{Ls} B_n$,
8. $\mathrm{Ls} A_n - \mathrm{Ls} B_n \subset \mathrm{Li}(A_n - B_n)$,
9. $\mathrm{Li}(A_n \cup B_n) \subset \mathrm{Li} A_n \cup \mathrm{Li} B_n \cup (\mathrm{Ls} A_n \cap \mathrm{Ls} B_n)$.

**15.** If $\operatorname*{Li}_{n \to \infty} A_n = \operatorname*{Ls}_{n \to \infty} A_n$ we use the notation

$$\operatorname*{Lt}_{n \to \infty} A_n = \operatorname*{Li}_{n \to \infty} A_n = \operatorname*{Ls}_{n \to \infty} A_n$$

and call the set $\operatorname*{Lt}_{n \to \infty} A_t$ the *topological limit* of the sequence $A_1, A_2, \ldots$.

Prove the following formulas:

1. $\overline{\mathrm{Lt} A_n} = \mathrm{Lt} A_n = \mathrm{Lt} \overline{A_n}$,
2. $\mathrm{Lt}(A_n \cup B_n) = \mathrm{Lt} A_n \cup \mathrm{Lt} B_n$,
3. if $A_n = A$, then $\mathrm{Lt} A_n = \overline{A}$,
4. if $A_1 \subset A_2 \subset \ldots$, then $\mathrm{Lt} A_n = \overline{\bigcup_{n=1}^{\infty} A_n}$,
5. if $A_1 \supset A_2 \supset \ldots$, then $\mathrm{Lt} A_n = \bigcap_{n=1}^{\infty} \overline{A_n}$,
6. $\mathrm{Lt}(A_n \times B_n) = \mathrm{Lt} A_n \times \mathrm{Lt} B_n$.

CHAPTER XI

# BASIC TOPOLOGICAL CONCEPTS

## § 1. Borel sets

Borel sets are sets which belong to the smallest family $\boldsymbol{R}$ of subsets of a given space satisfying the following conditions:

(a) *every closed set belongs to* $\boldsymbol{R}$,

(b) *if* $X_n \in \boldsymbol{R}$ *for* $n = 1, 2, \ldots$, *then* $\bigcup_{n=1}^{\infty} X_n \in \boldsymbol{R}$,

(c) *if* $X_n \in \boldsymbol{R}$ *for* $n = 1, 2, \ldots$, *then* $\bigcap_{n=1}^{\infty} X_n \in \boldsymbol{R}$.

A family of Borel sets is therefore, in the sense of the terminology of Chapter IV, § 7, a Borel family generated by the family of closed sets.

Making use of ordinal numbers we can classify Borel sets in classes $\boldsymbol{R}_\alpha$, where $\alpha < \Omega$, in the following manner.

1. The class $\boldsymbol{R}_0$ is the family of all closed sets.

2. For $\alpha = \lambda + n > 0$, where $\lambda$ is a limit ordinal and $n$ is a non-negative integer, the class $\boldsymbol{R}_\alpha$ is the family of all sets of the form

$$\bigcap_{k=1}^{\infty} X_k \quad \text{or} \quad \bigcup_{k=1}^{\infty} X_k,$$

according to whether $n$ is even or odd, and the sets $X_1, X_2, \ldots$ belong to classes of indices smaller than $\alpha$.

In particular, the class $\boldsymbol{R}_1$ is the family of all countable unions of closed sets; they are called $F_\sigma$-sets. The class $\boldsymbol{R}_2$ is the family of intersections of a countable number of $F_\sigma$-sets (they are called $F_{\sigma\delta}$-sets), and so forth.

It can be proved that for every $\alpha < \Omega$ there exists in the space of real numbers a set of the class $\boldsymbol{R}_\alpha$ which does not belong to any class with index smaller than $\alpha$.

Remark. If we start with open sets, instead of closed sets (cf. condition (a)), we obtain the Borel family generated by the family of open sets (which, as can be proved, is in metric spaces identical with the Borel family, considered above, generated by the family of closed sets; see Chapter XII, § 4). Here the open sets form the zero class, the $G_\delta$-sets, i.e. countable intersections of open sets form the first class, the $G_{\delta_\sigma}$-sets form the second class, and so forth. This classification is dual to the classification previously considered.

## § 2. Dense sets and boundary sets

A set $A$ is said to be *dense* if $\overline{A} = X$. A set $A$ is said to be a *boundary set* if its complement is a dense set, i.e. if $\overline{X-A} = X$. (A set whose closure is a boundary set is also said to be a *nowhere dense* set.)

Obviously, every set which contains a dense set is dense and a subset of a boundary set is a boundary set.

In the space $\mathscr{E}$ of all real numbers, the set of rational numbers is both a dense and a boundary set. In the plane $\mathscr{E}^2$ a straight line is a boundary set.

It can be easily proved (applying formula (9) of Chapter X, § 5) that the following theorems are valid:

THEOREM 1. *A set $A$ is dense iff in every open set ($\neq \emptyset$) there exist points which belong to $A$.*

THEOREM 2. *A set $A$ is a boundary set iff in every open set ($\neq \emptyset$) there exist points which do not belong to $A$.*

THEOREM 3. *A closed set $A$ is a boundary set iff for every open set ($\neq \emptyset$) $G$ there exists an open set ($\neq \emptyset$) $H \subset G$ such that $H \cap A = \emptyset$.*

The union of two boundary sets might not be a boundary set. For example, the set of all rational numbers and the set of all irrational numbers are boundary sets (in the space of real numbers), but their union is not a boundary set.

On the other hand, the following theorem can be proved:

THEOREM 4. *If a set $A$ is a boundary set and the set $B$ is a closed boundary set, then $A \cup B$ is a boundary set.*

Hint for the proof. Applying formula 2 of Chapter X, § 3, we have

$$X-B = \overline{X-A}-B \subset \overline{(X-A)-B} = \overline{X-(A \cup B)}.$$

## § 3. $\mathscr{T}_1$-spaces. $\mathscr{T}_2$-spaces

Definition 1. A topological space $X$ is called a *$\mathscr{T}_1$-space* if each single element set is closed:

$$\overline{\{p\}} = \{p\} \quad \text{for each} \quad p \in X. \tag{1}$$

Clearly, every metric space is a $\mathscr{T}_1$-space.

On the other hand, there are topological spaces which are not $\mathscr{T}_1$-spaces. Such is e.g. each space $X$ containing more than one point and containing only two open sets: $\emptyset$ and $X$.

The topological spaces which are under consideration in this book are—as a rule—assumed to be $\mathscr{T}_1$-spaces; this assumption will not always be explicitly formulated.

Definition 2. A topological space is called a *Hausdorff space* (or a *$\mathscr{T}_2$-space*, or a *separated space*) if for each pair of points $p \neq q$ there are two open sets $G$ and $H$ such that

$$p \in G, \quad q \in H \quad \text{and} \quad G \cap H = \emptyset. \tag{2}$$

Clearly, each metric space is a $\mathscr{T}_2$-space.

THEOREM. *Each $\mathscr{T}_2$-space is a $\mathscr{T}_1$-space.*

Let $p$ be a given point. By assumption, each $x \neq p$ belongs to an open $G_x$ such that $p$ is not in $G_x$. Consequently, $X - \{p\} = \bigcup_{x \neq p} G_x$. Thus $X - \{p\}$ is open, and $\{p\}$ is closed.

Remark 1. *There exist $\mathscr{T}_1$-spaces which are not $\mathscr{T}_2$.* Such is the space composed of the points $1/n$ for $n = 1, 2, \ldots$ and the point 0 with the following topology: Open sets which do not contain the point 1 are identical with open sets in the usual topology of real numbers; sets containing 1 are open if and only if they are complements of finite sets. Obviously, each open set containing 0 is infinite, and consequently the points 0 and 1 cannot be separated by means of disjoint open sets.

Remark 2. It is easy to show that the properties of being a $\mathscr{T}_1$-space and of being a $\mathscr{T}_2$-space are *hereditary*; that means that every subset of a $\mathscr{T}_1$-space is $\mathscr{T}_1$ and every subset of a $\mathscr{T}_2$-space is $\mathscr{T}_2$.

## § 4. Regular spaces, normal spaces

A notion less general than that of a $\mathscr{T}_2$-space is the notion of a regular space.

Definition 1. A topological space is called *regular* if every point $p$ and every closed set $F$ which does not contain $p$ can be separated by disjoint open sets; i.e. if there are two open sets $G_0$ and $G_1$ such that

(3) $$p \in G_0, \quad F \subset G_1 \quad \text{and} \quad G_0 \cap G_1 = 0.$$

This can be stated equivalently: if there exists an open set $G$ such that

(4) $$p \in G \quad \text{and} \quad \overline{G} \cap F = 0.$$

A still less general notion than that of a regular space is the notion of a normal space.

Definition 2. A topological space is called *normal* if for each pair of disjoint closed sets $A$ and $B$, there are two disjoint open sets $G$ and $H$ such that

(i) $$A \subset G \quad \text{and} \quad B \subset H.$$

Equivalently stated, if for each pair $A \subset C$ where $A$ is closed and $C$ open, there is $G$ open such that

(ii) $$A \subset G \quad \text{and} \quad \overline{G} \subset C.$$

The implication (i) $\Rightarrow$ (ii) follows by substituting $B = X - C$. In order to obtain the converse implication we put $C = X - B$ and $H = X - \overline{G}$.

Remark 1. Obviously every normal $\mathscr{T}_1$-space is regular. The converse is not true; moreover regularity is hereditary, while normality is not (see Exercise 10).

On the other hand, *every metric space is normal* (see Chapter XII, § 4, Theorem 6).

Remark 2. *If $X$ is regular and $\boldsymbol{B}$ is its base, then every open set $G$ is the union of some members of $\boldsymbol{B}$ whose closures are contained in $G$.*

For, let $p \in G$ and let $p \in H$ where $H$ is open and $\overline{H} \subset G$ (the existence of $H$ follows from the regularity of $X$). Since $\boldsymbol{B}$ is a base of $X$, we have $H = \bigcup_t R_t$ where $R_t \in \boldsymbol{B}$. Hence there is $t$ such that $p \in R_t$. Obviously $\overline{R}_t \subset \overline{H} \subset G$.

Remark 3. If $X$ is regular, the Theorem of Chapter X, § 11 can be strengthened in the following way: we assume that not only the cover $\boldsymbol{R} \subset \boldsymbol{B}$ is a refinement of $\boldsymbol{C}$ but that *the family of closures of members of* $\boldsymbol{R}$ *is a refinement of* $\boldsymbol{C}$.

To show this one has only to replace in the proof $H_s$ by $\overline{H}_s$.

## § 5. Accumulation points. Isolated points

Definitions. $p$ is an *accumulation point* of the set $A$ if $p \in \overline{A - \{p\}}$.

$p$ is an *isolated* point of $A$ if $p$ belongs to $A$ but is not its accumulation point.

For example, the point 0 is the only accumulation point of the infinite set $A$ of points 1, 1/2, 1/3, ...; all these points are isolated points of $A$.

It is easy to prove the following theorems.

THEOREM 1. *$p$ is an accumulation point of $A$ iff every open set containing $p$ contains a point of $A$ different from $p$.*

THEOREM 2. *$p$ is an isolated point of $A$ iff there is an open set $G$ such that $G \cap A = \{p\}$.*

Consequently, *$p$ is an isolated point of the space iff the set $\{p\}$ is open.*

THEOREM 3. *If $X$ is metric, then $p$ is an accumulation point of $X$ iff*

$$p = \lim_{n\to\infty} p_n \quad \textit{where} \quad p_n \neq p.$$

## § 6. The derived set

The set of all accumulation points of $A$, denoted $A^d$, is called the derived set of $A$.

In $\mathscr{T}_1$-spaces the derived set has the following properties.

1. $\overline{A} = A \cup A^d$,
2. $(A \cup B)^d = A^d \cup B^d$,
3. $\bigcup_t A_t^d \subset (\bigcup_t A_t)^d$,
4. $A^{dd} \subset A^d$,
5. $\overline{A^d} = A^d$.

The formulas 1–3 can be easily proved (they hold in every topological space). Let us establish formula 4.

Suppose that $p \notin A^d$, i.e. $p \notin \overline{A-\{p\}}$. There is therefore an open $G$ such that $p \in G$ and $G \cap A-\{p\} = \emptyset$. Suppose (contrary to 4) that $p \in A^{dd}$, i.e. $p \in \overline{A^d-\{p\}}$. Since $p \in G$, we have $G \cap A^d - p \neq \emptyset$, and there is $q \neq p$ such that $q \in G \cap A^d$. It follows that

$$q \in (G-\{p\}) \cap \overline{A-\{q\}}, \qquad \text{hence} \qquad (G-\{p\}) \cap (A-\{q\}) \neq 0,$$

since $G-\{p\}$ is open (the space being supposed $\mathscr{T}_1$). But this contradicts formula $G \cap A-\{p\} = \emptyset$.

Formula 5 is an easy consequence of 1 and 4.

Remark 1. If the space is not $\mathscr{T}_1$, formula 5 (hence 4) may not hold. Such is the case if $X = \{p, q\}$ where $p \neq q$ and where the topology is trivial. For $\{p\}^d = \{q\}$ while $\overline{\{q\}} = \{p, q\}$.

Remark 2. In contrast to the closure, the second derived set need not necessarily be equal to the first. If, for example, $A$ consists of the points $1, 1/2, 1/3, \ldots$, then $A^d$ consists of the point 0, and $A^{dd}$ is the null set. If $A$ is the set of numbers of the form $1/n+1/m$ $(n, m = 1, 2, \ldots)$, then $A \neq A^{dd} \neq A^{ddd} = \emptyset$.

## § 7. Sets dense in themselves

Definition. A set each of whose point is an accumulation point of this set is said to be a set *dense in itself.*

Hence these sets are characterized by the inclusion

$$A \subset A^d$$

or—what amounts to the same—by the condition that they do not contain isolated points.

In $\mathscr{T}_1$-spaces one has the following theorems.

THEOREM 1. *The closure of a set dense in itself is dense in itself.*

Proof. Let $A$ be a set which is dense in itself and therefore satisfies formula (4). From this, by virtue of formula 1 of § 6, we infer that

$$A^d = A \cup A^d = \bar{A}, \tag{5}$$

and therefore, applying formulas 2 and 4 of § 6, we obtain

$$(\bar{A})^d = (A \cup A^d)^d = A^d \cup A^d = A^d,$$

whence by (5), we have $(\bar{A})^d = \bar{A}$. Hence, the set $\bar{A}$ is dense in itself.

THEOREM 2. *The union of an arbitrary family of sets which are dense in themselves is a set dense in itself.*

For, if $A_t \subset A_t^d$, then, by virtue of formula 3 of § 6, we have

$$\bigcup_t A_t \subset A_t^d \subset (\bigcup_t A_t)^d.$$

THEOREM 3. *Each $\mathcal{T}_1$-space is the union of two sets of which one is closed and dense in itself and the other does not contain any non-empty subset which is dense in itself.*

Proof. Let $C$ denote the union of all subsets of the given space which are dense in themselves. It follows from Theorem 2 that the set $C$ is dense in itself and therefore, by virtue of Theorem 1, the set $\bar{C}$ is also dense in itself and hence is a subset of the set $C$. Thus $\bar{C} \subset C$, i.e. the set $C$ is closed. Finally, the set $X-C$, being disjoint from $C$ does not contain non-empty sets which are dense in themselves.

Remark. Sets which are simultaneously closed and dense in themselves are also called *perfect sets*. They are therefore characterized by the identity $A = A^d$. Sets which do not contain any non-empty subset which is dense in itself are called *scattered sets*.

**Exercises**

**1.** Prove that: (a) the complement of a $G_\delta$-set is an $F_\sigma$-set, (b) the union of an infinite sequence of $F_\sigma$-sets is an $F_\sigma$-set; the intersection of two $F_\sigma$-sets is an $F_\sigma$-set. State theorems on $G_\delta$-sets which are the duals of (b) (use the De Morgan rules).

**2.** Prove that the intersection of any collection of topologies for $X$ is a topology for $X$.

**3.** Every open subset of a dense in itself space is dense in itself.

**4.** If the sets $A$ and $X-A$ are boundary sets, then the space $X$ is dense in itself.

**5.** The set Int [Fr($A$)] is dense in itself.

**6.** We say that a $\mathcal{T}_2$-space $X$ has the property (A) if

$$(A, B \subset X \text{ and } A \cap B = \emptyset) \Rightarrow A^d \cap B^d = \emptyset.$$

Prove that $X$ has property (A) iff for every accumulation point $x \in X$ the family of sets $\{V : V \cup \{x\}$ is a neighbourhood of $X\}$ is a maximal filter.

**7.** Show that every subset of a regular space is regular.

**8.** Prove that $X$ is normal iff the condition $X = G \cup H$, where $G$ and $H$ are open sets, implies the existence of closed sets $A$ and $B$ such that

$$X = A \cup B, \quad A \subset G \quad \text{and} \quad B \subset H.$$

**9.** Prove that, if every open set of a normal space is normal, then the set is *hereditarily normal* (i.e. every subset of the space is normal).

**10.** Let $X = \{\alpha : \alpha \leqslant \Omega\}$ and $Y = \{\beta: \beta \leqslant \omega\}$ with the natural (order) topology (i.e. with topology generated by sets $\{\alpha: \alpha < \beta\}$ and $\{\alpha: \alpha > \beta\}$). Show that the cartesian product (see Chapter XIII, § 1) $X \times Y$ is normal, while the set $(X \times Y) - \langle \Omega, \omega \rangle$ (called the *Tychonov plank*) is not.

**11.** Show that a space is hereditarily normal iff for each pair of separated sets $A$ and $B$ there are two disjoint open sets $G$ and $H$ such that $A \subset G$ and $B \subset H$.

**12.** Let $X$ be a normal space. Two systems of sets $A_1, \ldots, A_n \subset X$ and $B_1, \ldots, B_n \subset X$ are called *combinatorially equivalent* if the equivalence

$$(A_{i_1} \cap \cdots \cap A_{i_k} = 0) \equiv (B_{i_1} \cap \cdots \cap B_{i_k} = 0)$$

holds for every sequence of indices ($\leqslant n$).

Prove that for every system $F_1, \ldots, F_n$ of closed sets there exists a combinatorially equivalent system $G_1, \ldots, G_n$ of open sets such that $F_i \subset G_i$ for $i = 1, \ldots, n$.

*Hint*: The proof is by induction. First define an open set $G_1$ in such a way that the systems $\overline{G_1}, F_2, \ldots, F_n$ and $F_1, \ldots, F_n$ are combinatorially equivalent

**13.** Deduce the following corollary from the preceding theorem:

Let $X$ be a normal space. If the sets $G_1, G_2, \ldots, G_n$ are open and $X = G_1 \cup G_2 \cup \ldots \cup G_n$, then there exist closed sets $F_1, F_2, \ldots, F_n$ such that

$$F_1 \cup F_2 \cup \ldots \cup F_n = X \quad \text{and} \quad F_i \subset G_i \quad \text{for} \quad i = 1, 2, \ldots, n.$$

**14.** Consider in the space of rational numbers the subbase $S$ consisting of open bounded intervals and of the set of dyadic numbers (of the form $k/2^n$).

Show that the space generated by $S$ is a non-regular $\mathcal{T}_2$-space.

CHAPTER XII

# CONTINUOUS MAPPINGS

## § 1. Continuity

D e f i n i t i o n. Let $X$ and $Y$ be topological spaces and let $f\colon X \to Y$. $f$ is said to be *continuous at the point* $x_0$ if

$$(1) \qquad x_0 \in \overline{A} \Rightarrow f(x_0) \in \overline{f(A)} \textit{ for each } A \subset X.$$

Mappings continuous at *each point* are called, briefly *continuous*. The set of these mappings is denoted $(Y^X)_{\text{top}}$, or briefly $Y^X$—when no confusion can occur with the notation of Chapter IV, § 1.

THEOREM 1. *Condition* (1) *is equivalent to*

$$(2) \qquad x_0 \in \overline{f^{-1}(B)} \Rightarrow f(x_0) \in \overline{B} \textit{ for each } B \subset Y.$$

P r o o f. 1. Suppose that (1) is true and that $x_0 \in \overline{f^{-1}(B)}$. Put $A = f^{-1}(B)$ in (1). It follows that $f(x_0) \in \overline{f[f^{-1}(B)]}$. But $f[f^{-1}(B)] \subset B$ (by Chapter IV, (18)). Hence $f(x_0) \in \overline{B}$.

2. Suppose that (2) is true and that $x_0 \in \overline{A}$. Put $B = f(A)$ in (2). As $A \subset f^{-1}[f(A)] = f^{-1}(B)$ (by Chapter IV, (19), so $x_0 \in \overline{f^{-1}(B)}$, and (2) yields $f(x_0) \in \overline{B}$, hence $f(x_0) \in \overline{f(A)}$.

COROLLARY 1. *The continuity of $f$ at $x_0$ is equivalent to the following condition: if $C$ is a neighbourhood of $f(x_0)$, then $f^{-1}(C)$ is a neighbourhood of $x_0$.* In other words: *if $f(x_0) \in H$ and $H$ is open (in $Y$), then there is an open $G$ (in $X$) such that $x_0 \in G$ and $f(G) \subset H$.*

P r o o f. Substitute $Y-C$ for $B$ in (2). One obtains the equivalent formula (compare Chapter IV, (17a)):

$$(2') \qquad f(x_0) \in X-\overline{Y-C} \Rightarrow x_0 \in X-\overline{f^{-1}(Y-C)} = X-\overline{X-f^{-1}(C)},$$

i.e.

$$(2'') \qquad f(x_0) \in \operatorname{Int}(C) \Rightarrow x_0 \in \operatorname{Int}[f^{-1}(C)].$$

This completes the proof of the first part of the Corollary. Its second part follows from the fact that in each neighbourhood of a point there is an open set containing this point.

THEOREM 2. *$f$ is continuous iff*

$$f(\bar{A}) \subset \overline{f(A)} \text{ for each } A \subset X. \tag{3}$$

Equivalently: *iff*

$$\overline{f^{-1}(B)} \subset f^{-1}(\bar{B}) \text{ for each } B \subset Y. \tag{4}$$

P r o o f. 1. Suppose that $f$ is continuous and let $y_0 \in f(\bar{A})$. Hence there exists $x_0 \in \bar{A}$ such that $y_0 = f(x_0)$. By (1) $f(x_0) \in \overline{f(A)}$, hence $y_0 \in \overline{f(A)}$. Thus inclusion (3) is true.

2. Suppose (3) is true and let $x_0 \in \bar{A}$. Hence

$$f(x_0) \in f(\bar{A}) \subset \overline{f(A)}, \quad \text{therefore} \quad f(x_0) \in \overline{f(A)}.$$

Thus $f$ is continuous at $x_0$.

The proof of the second part of Theorem 2 is similar.

COROLLARY 2. *$f: X \to Y$ is continuous iff $f^{-1}(B)$ is closed whenever $B \subset Y$ is closed*; equivalently—*iff $f^{-1}(B)$ is open whenever $B \subset Y$ is open.*

P r o o f. To prove the necessity of the condition, we put $B = \bar{B}$ in (4).

Conversely, suppose that the inverse image of each closed set is closed. Then we have, for arbitrary $B \subset Y$, $\overline{f^{-1}(\bar{B})} = f^{-1}(\bar{B})$, and (4) follows since $\overline{f^{-1}(B)} \subset \overline{f^{-1}(\bar{B})}$.

The second part of the Corollary follows from the first one by virtue of the identity $f^{-1}(Y-B) = X-f^{-1}(B)$.

COROLLARY 3. *Let $S$ be a subbase of $Y$ and let $f: X \to Y$. If $f^{-1}(G)$ is open for each $G \in S$, then $f$ is continuous.*

P r o o f. Since the family of all open subsets of $Y$ is generated from $S$ by means of two operations, the finite intersection and the union, our conclusion follows from the formulas

$$f^{-1}\left(\bigcap_t Q_t\right) = \bigcap_t f^{-1}(Q_t), \quad f^{-1}\left(\bigcup_t Q_t\right) = \bigcup_t f^{-1}(Q_t).$$

THEOREM 3. *The composition of two continuous mappings is a continuous mapping.*

More precisely, *if $f: X \to Y$, $g: Y \to Z$, $h = g \circ f$ and $f$ is continuous at $x_0$ and $g$ is continuous at $y_0 = f(x_0)$, then $h$ is continuous at $x_0$.*

P r o o f. Let $x_0 \in \bar{A}$. Then $f(x_0) \in \overline{f(A)}$, hence $g[f(x_0)] \in \overline{g[f(A)]}$, i.e. $h(x_0) \in \overline{h(A)}$.

## § 2. Homeomorphisms

D e f i n i t i o n. If the mapping $f$ of the space $X$ onto (the whole) space $Y$ is continuous and one-to-one, and its inverse $f^{-1}$ is also continuous, then we say that $f$ is a *homeomorphism*, and the spaces $X$ and $Y$ are said to be *homeomorphic* (or of the same topological type). We then write

$$X \underset{\text{top}}{=} Y \text{ (topological equivalence).}$$

If $X = Y$, a homeomorphism is called a *topological automorphism.*

The homeomorphism relation is clearly *reflexive, symmetric,* and *transitive.*

THEOREM 1. *Each of the following conditions is necessary and sufficient for a one-to-one mapping $f$ to be a homeomorphism*:

(i) $$f(\bar{A}) = \overline{f(A)} \text{ for every } A \subset X,$$

(i′) $$f^{-1}(\bar{B}) = \overline{f^{-1}(B)} \text{ for every } B \subset Y.$$

P r o o f. By § 1, (3), the inclusion $f(\bar{A}) \subset \overline{f(A)}$ is equivalent to the continuity of $f$, while the inclusion $\overline{f(A)} \subset f(\bar{A})$, by § 1, (4), is equivalent to the continuity of $f^{-1}$. The second part of the theorem can be proved in a similar way.

THEOREM 2. *A necessary and sufficient condition for $f$ to be a homeomorphism (where $X$ is supposed to be a $\mathcal{T}_1$-space) is the following*:

(ii) $$\bar{A} = f^{-1}(\overline{f(A)}) \text{ for every } A \subset X,$$

*or equivalently*

(ii′) $$(x \in \bar{A}) \equiv \left(f(x) \in \overline{f(A)}\right).$$

P r o o f. If $f$ is one-to-one, our statement is true, for then (ii) is equivalent to (i). It remains to show that a function satisfying (ii) is one-to-one. Let $f(p) = f(q)$. Then $\overline{\{p\}} = f^{-1}(\overline{f(p)}) = f^{-1}(\overline{f(q)}) = \overline{\{q\}}$, whence $\{p\} = \{q\}$ and finally $p = q$.

Remark 1. Each property of the space which is invariant under homeomorphisms is called a *topological property*. It follows from (ii′) that every property expressed in terms of the operation $\bar{A}$ (and of operations of set theory and of logics) is topological.

More generally, if a point $a$ (or a set $A$, or a family $\boldsymbol{A}$ of sets, and so on) has a given property with respect to the space $X$ and if $f$ is a homeomorphism which maps $X$ onto a space $Y$, then the point $f(a)$ has the same property with respect to $Y$ (provided that the property is expressed as above).

Thus *it is impossible to distinguish between two homeomorphic spaces by any topological means*. Similarly, if $A$ and $B$ are two sets situated in the spaces $X$ and $Y$, respectively, and if there exists a homeomorphism of $X$ onto $Y$ which maps $A$ onto $B$, then the sets $A$ and $B$ are indistinguishable in their spaces from the topological point of view (with respect to the spaces $X$ and $Y$).

It should be remarked that two sets may be homeomorphic and, at the same time, they may be situated in the space in a different manner so that there is no equivalence between them. For example, in the space of real numbers, the set composed of a point, a segment, and a second point (in that order) is not topologically equivalent (though it is homeomorphic) to the set composed of two points and a segment following them. However, the same sets (regarded as subsets of the plane) are equivalent with respect to the plane.

EXAMPLES. 1. Let $a \leqslant x \leqslant b$ and $c \leqslant y \leqslant d$, where $a < b$ and $c < d$, be two given closed intervals of real numbers. The function

$$y = \{(d-c)/(b-a)\}x + (bc-ad)/(b-a)$$

is a homeomorphism which maps the first interval onto the second. Hence, any two closed intervals are homeomorphic.

The same function maps the open interval $a < x < b$ homeomorphically onto the open interval $c < y < d$.

2. The function $y = \tan x$ maps the open interval $-\pi/2 < x < \pi/2$ homeomorphically onto the entire set of real numbers.

3. A necessary and sufficient condition for a continuous real valued function, defined on the closed interval $a \leqslant x \leqslant b$, to be a homeomorphism, is that it be strictly monotonic.

4. Let us consider in Euclidean 3-space $\mathscr{E}^3$ the surface of the sphere $x^2+y^2+(z-1)^2 = 1$ and let us draw a line, which is not parallel to the $XY$-plane, from the point $b = (0, 0, 2)$. Let us assign to the point $p$ of intersection of this line with the surface of the sphere, the point $f(p)$ which is the point of intersection of this line with the plane $z = 0$.

The function $f$ so defined is, as is easy to verify, a homeomorphism which maps the surface of a sphere with the point $b$ removed onto the entire plane. Hence the plane is homeomorphic to the surface of the sphere with one point removed. One makes use of this fact in the theory of analytic functions when it is said that the plane of complex numbers is completed with "the point at infinity" to the surface of the sphere.

Remark 2. In the definition of homeomorphism, the condition of continuity of the inverse map is essential, which means that the continuity of the mapping $f$ does not imply the continuity of the mapping $f^{-1}$. For example, the function $z = e^{2\pi i x}$ maps the set $0 \leqslant x < 1$ onto the set of complex numbers lying on the circle with equation $|z| = 1$ in a continuous and one-to-one manner. However, the inverse mapping is not continuous at the point $z = 1$.

Remark 3. The totality of all topological spaces which have the same cardinality can be *ordered* assuming that $Y \leqslant X$ if there exists a one-to-one continuous mapping $f$ of $X$ onto $Y$. This is equivalent to saying that $Y$ is *coarser* than $X$ (compare Chapter X, § 10).

Remark 4. The space $X$ is said to be *topologically contained* (or *imbedded*) in the space $Y$ if it is homeomorphic to a subset of $Y$. We write in this case

$$X \underset{\text{top}}{\subset} Y.$$

## § 3. Case of metric spaces

First, let us suppose that $X$ is an arbitrary topological space and that $Y$ is metric.

THEOREM 1. *Let $f\colon X \to Y$. $f$ is continuous at $x_0$ iff for each $\varepsilon > 0$ there is an open $G$ containing $x_0$ and such that*

$$x \in G \Rightarrow |f(x) - f(x_0)| < \varepsilon.$$

This follows from Corollary 1 of § 1, since

$$\big(|f(x) - f(x_0)| < \varepsilon\big) \equiv \big(f(x) \in K[f(x_0), \varepsilon]\big)$$

and, on the other hand, if $H$ is a neighbourhood of $f(x_0)$, then $K[f(x_0), \varepsilon] \subset H$ whenever $\varepsilon$ is sufficiently small.

Similarly, if both spaces, $X$ and $Y$, are metric, then the continuity of $f$ at $x_0$ is equivalent to the *Cauchy condition*

$$\bigwedge_\varepsilon \bigvee_\delta \bigwedge_x \big\{(|x - x_0| < \delta) \Rightarrow \big(|f(x) - f(x_0)| < \varepsilon\big)\big\}. \tag{5}$$

It is also equivalent to the following *Heine condition.*

THEOREM 2. *$f$ is continuous at $x_0$ iff the condition*

$$\lim_{n\to\infty} x_n = x_0 \tag{6}$$

*implies*

$$\lim_{n\to\infty} f(x_n) = f(x_0), \tag{7}$$

*whatever the sequence $x_1, x_2, \ldots$ of points of $X$.*

P r o o f. 1. Let us suppose that $f$ is continuous at $x_0$ in the sense of Heine and that — contrary to our assumption — (5) is not true, i.e. that

$$\bigvee_\varepsilon \bigwedge_\delta \bigvee_x (|x - x_0| < \delta)\big(|f(x) - f(x_0)| \geqslant \varepsilon\big).$$

Put $\delta = 1/n$. Then there exists (by the choice axiom) a sequence $x_1, x_2, \ldots$ such that

$$|x_n - x_0| < 1/n, \tag{8}$$

$$|f(x_n) - f(x_0)| \geqslant \varepsilon. \tag{9}$$

(8) implies (6) hence (7) (by the Heine continuity). But this contradicts (9).

2. Let us assume (5) and (6). Hence there is a $k$ such that $|x_n - x_0| < \delta$ for $n > k$. It follows by (5) that $|f(x_n) - f(x_0)| < \varepsilon$. This means that (7) is true, and thus $f$ is continuous at $x_0$ in the sense of Heine.

R e m a r k 1. *If $f$ is a continuous real valued function, then the sets*

$$\{x: f(x) \leqslant a\}, \quad \{x: f(x) \geqslant a\}, \quad \{x: a \leqslant f(x) \leqslant b\}$$

*are closed and the sets*

$$\{x: f(x) < a\}, \quad \{x: f(x) > a\}, \quad \{x: a < f(x) < b\}$$

*are open.*

This follows from Corollary 2 of § 1 because these sets are inverse images of the closed sets

$$\{y: (y \leqslant a)\}, \quad \{y: (y \geqslant a)\}, \quad \{y: (a \leqslant y \leqslant b)\},$$

and of the open sets

$$\{y: (y < a)\}, \quad \{y: (y > a)\}, \quad \{y: (a < y < b)\},$$

respectively.

THEOREM 3. *The limit of a uniformly convergent sequence of continuous mappings $f$: $X \to Y$, where $X$ is topological and $Y$ metric, is continuous.*

P r o o f. Put $f(x) = \lim_{n=\infty} f_n(x)$. Let $\varepsilon > 0$ and let $x_0$ be a given point of $X$. By assumption, $k$ exists such that

$$(10) \qquad |f_k(x) - f(x)| < \varepsilon/3 \quad \text{for each } x \in X.$$

Therefore, substituting $x = x_0$, we have

$$(11) \qquad |f_k(x_0) - f(x_0)| < \varepsilon/3.$$

Since $f_k$ is continuous at the point $x_0$, there is an open set $G$ containing $x_0$ such that

$$(12) \qquad |f_k(x) - f_k(x_0)| < \varepsilon/3 \quad \text{for each } x \in G.$$

Inequalities (10) to (12) yield $|f(x) - f(x_0)| < \varepsilon$, which means that $x_0$ is a point of continuity of $f$.

THEOREM 4. *Let $X_1, X_2, \ldots$ be a finite or infinite sequence of (uniformly bounded) metric spaces. Put $Z = X_1 \times X_2 \times \ldots$ and $f$: $Z \to Y$ where $Y$ is metric. Put $z^0 = (x_1^0, x_2^0, \ldots)$ where $x_m^0 \in X_m$. Then $f$ is continuous at $z^0$, iff*

$$[\bigwedge_m (\lim_{n\to\infty} x_m^n = x_m^0)] \Rightarrow [\lim_{n\to\infty} f(z^n) = f(z^0)]$$

*where $z^n = (x_1^n, x_2^n, \ldots)$.*

THEOREM 5. *Under the same assumptions on $X_m$, $m = 1, 2, \ldots$, and $Z$, let $f$: $T \to Z$ (where $T$ is metric). Put $f(t) = (f_1(t), f_2(t), \ldots)$. Then $f$ is continuous at $t_0$ iff*

$$(\lim_{n\to\infty} t_n = t_0) \Rightarrow \bigwedge_m [\lim_{n\to\infty} f_m(t_n) = f_m(t_0)],$$

*which means that each $f_m$ is continuous at $t_0$.*

Theorems 4 and 5 are easy consequences of the Heine condition and of Theorem 2 of Chapter IX, § 6 (the assumption of uniform boundedness can be omitted by virtue of the Theorem 8 below).

THEOREM 6. *The distance between two points of a metric space $X$ is a continuous real-valued mapping of $X \times X$ into $\mathscr{E}$.*

Proof. Let

$$\lim_{n\to\infty} x_n = x, \quad \lim_{n\to\infty} y_n = y$$

and let $\varepsilon > 0$ be given. Then there exists a $k$ such that for $n > k$ we have

$$|x - x_n| < \varepsilon, \quad |y_n - y| < \varepsilon. \tag{13}$$

From the triangle inequality we obtain (see Fig. 5):

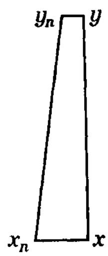



FIG. 5

$$|x - y| \leqslant |x - x_n| + |x_n - y_n| + |y_n - y|. \tag{14}$$

It follows from (13) and (14) that

$$|x - y| < |x_n - y_n| + 2\varepsilon. \tag{15}$$

Similarly, from the inequality

$$|x_n - y_n| \leqslant |x_n - x| + |x - y| + |y_n - y|$$

we obtain the inequality

$$(16) \qquad |x_n - y_n| < |x-y| + 2\varepsilon .$$

By (15) and (16) we have for $n > k$

$$\big||x_n - y_n| - |x-y|\big| < 2\varepsilon .$$

This means that $\lim_{n\to\infty} |x_n - y_n| = |x-y|$ and hence the function $|x-y|$ is continuous.

THEOREM 7. *Let $X$ and $Y$ be two metric spaces and $f$: $X \to Y$ onto. Then $f$ is a homeomorphism iff*

$$(17) \qquad (\lim_{n\to\infty} x_n = x_0) \equiv [\lim_{n\to\infty} f(x_n) = f(x_0)]$$

*for every sequence of points in $X$.*

This is an easy consequence of Theorem 2.

THEOREM 8. *Every metric space $X$ is homeomorphic to a bounded space $X^*$.*

P r o o f. We denote by $X^*$ the set $X$ with a "new distance" $||x-y||$ defined as follows. If $|x-y| \leqslant 1$, then $||x-y|| = |x-y|$; if $|x-y| > 1$, then $||x-y|| = 1$.

It is easy to see that the new distance satisfies conditions (1)–(3) of Chapter IX, § 1, which means that $X^*$ is a metric space. Moreover $\delta(X^*) \leqslant 1$.

The mapping $f$: $X \to X^*$ defined by $f(x) = x$ is a homeomorphism. This follows from the equivalence of the conditions

$$\lim_{n\to\infty} ||x_n - y|| = 0 \quad \text{and} \quad \lim_{n\to\infty} |x_n - y| = 0 .$$

R e m a r k 2. Referring to the assumption of boundedness in the definition of the distance in the infinite product $X_1 \times X_2 \times \ldots$ (see Chapter IX, § 3 and § 6), let us note that this assumption may be omitted if we denote the distance between the points $x = (x_1, x_2, \ldots)$ and $y = (y_1, y_2, \ldots)$ using the formula

$$(18) \qquad |x-y| = \sum_{m=1}^{\infty} \frac{1}{2^m} ||x_m - y_m|| .$$

Also the Theorem 2 of Chapter IX, § 6 and the above proved Theorems 4 and 5 remain valid without the assumption of uniform boundedness.

## § 4. Distance of a point from a set. Normality of metric spaces

The distance of the point $x$ from the non-empty set $A$ is defined to be the number

$$(19)\qquad \varrho(x, A) = \text{greatest lower bound of the numbers } |x-a|, \text{ where } a \in A.$$

We assume, moreover, that $\varrho(x, \emptyset) = 1$. Let us note that:

THEOREM 1. *If* $A = \{y\}$, *then* $\varrho(x, A) = |x-y|$.

THEOREM 2. *If* $\emptyset \neq A \subset B$, then $\varrho(x, B) \leqslant \varrho(x, A)$.

THEOREM 3. $[\varrho(x, A) = 0] \equiv (x \in \overline{A})$.

In fact, if $x \in \overline{A}$, then for every $\varepsilon > 0$ there exists a point $a \in A$ such that $|x-a| < \varepsilon$. This means that $\varrho(x, A) = 0$.

Conversely, if $\varrho(x, A) = 0$, then for every $\varepsilon > 0$ there exists a point $a \in A$ such that $|x-a| < \varepsilon$, and hence $x \in \overline{A}$.

From this it follows that:

THEOREM 4. *If* $A$ *is a closed set, then*

$$[\varrho(x, A) = 0] \equiv (x \in A).$$

THEOREM 5. *The function* $\varrho(x, A)$ *is continuous (for fixed* $A$*).*

P r o o f. The theorem is obvious if the set $A$ is empty. Thus, we can assume that $A \neq \emptyset$. Let $\delta > 0$ and let

$$(20)\qquad |x-x'| < \delta.$$

By virtue of (19), there exists a point $a \in A$ (see Fig. 6) such that

$$(21)\qquad |x-a| \leqslant \varrho(x, A)+\delta.$$

If follows from (20) and (21) that

$$(22)\qquad \varrho(x', A) \leqslant |x'-a| \leqslant |x-a|+|x-x'| < \varrho(x, A)+\delta+\delta.$$

Similarly, we have

$$(23)\qquad \varrho(x, A) < \varrho(x', A)+2\delta.$$

Inequalities (22) and (23) yield

$$(24)\qquad |\varrho(x, A)-\varrho(x', A)| < 2\delta.$$

This means that inequality (20) implies inequality (24). Hence the function $\varrho(x, A)$ is continuous.

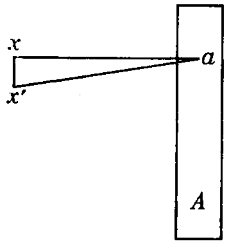



FIG. 6

THEOREM 6. *Every metric space X is normal, i.e. for every pair of disjoint closed subsets A and B, there exists a pair of disjoint open sets G and H such that*

(25) $$A \subset G \quad \textit{and} \quad B \subset H.$$

Proof. Let

$$G = \{x\colon \varrho(x, A) < \varrho(x, B)\}, \qquad H = \{x\colon \varrho(x, B) < \varrho(x, A)\}.$$

The sets $G$ and $H$ are open. In fact, by virtue of the continuity of the functions $\varrho(x, A)$ and $\varrho(x, B)$, the function $f(x) = \varrho(x, B) - \varrho(x, A)$ is also continuous. Since

$$G = \{x\colon \varrho(x, B) - \varrho(x, A) > 0\},$$

the set $G$ is open (cf. Remark 1, § 3). Similarly, the set $H$ is open.

The proof of the formula $G \cap H = \emptyset$ is immediate.

Finally, the formulas (25) hold. For, if $x \in A$, then by virtue of Theorem 4 we have $\varrho(x, A) = 0$, but $\varrho(x, B) \neq 0$, because $x$ does not belong to $B$ (since $A \cap B = \emptyset$). Therefore, $\varrho(x, A) < \varrho(x, B)$, and from this it follows that $x \in G$.

This means that $A \subset G$. Similarly, $B \subset H$.

THEOREM 7. *Every closed set in a metric space is a $G_\delta$-set.*

Proof. Let $F = \bar{F}$. Let us set

$$K(F, \varepsilon) = \{x\colon \varrho(x, F) < \varepsilon\}.$$

In view of the continuity of $\varrho(x, F)$ the set $K(F, \varepsilon)$ is open (cf. Remark 1, § 3). We shall show that

$$F = \bigcap_{n=1}^{\infty} K(F, 1/n).$$

If $x \in F$, then $\varrho(x, F) = 0$ and $x \in K(F, 1/n)$. Conversely, if $x \notin F$, then by virtue of Theorem 4, $\varrho(x, F) > 0$ and hence there exists an $n$ such that $\varrho(x, F) > 1/n$; therefore $x \notin K(F, 1/n)$.

Remarks. 1. It follows immediately from Theorem 7 that in a metric space every open set is an $F_\sigma$-set (and hence every $G_\delta$-set is an $F_{\sigma\delta}$-set). It also follows that condition (a) in the definition of Borel sets (Chapter XI, § 1) can be replaced by:

(a′) *every open set belongs to* $\boldsymbol{R}$.

2. Normal spaces in which closed sets are $G_\delta$ are called *perfectly normal*.

## § 5. Extension of continuous functions. Tietze theorem

LEMMA 1. *For every pair of disjoint closed sets $A$ and $B$ in the metric space $X$, there exists a continuous real valued function $f$ defined on the entire space $X$ and satisfying the following conditions*:

$$f(x) = \begin{cases} -1 & \text{for} \quad x \in A, \\ 1 & \text{for} \quad x \in B, \end{cases} \tag{26}$$

$$-1 < f(x) < 1 \quad \text{for} \quad x \notin A \cup B. \tag{27}$$

It is easy to prove, using Theorems 3–5, § 4, that the function $f$ defined by the formula

$$f(x) = \{\varrho(x, A) - \varrho(x, B)\}/\{\varrho(x, A) + \varrho(x, B)\}$$

satisfies the conditions set forth in the lemma.

LEMMA 2. *If $f$ is a continuous real valued function defined on a closed subset of the metric space $X$ such that $|f(x)| \leqslant \mu$ $(\neq 0)$, then there exists a continuous function $g$ defined on the entire space $X$ and satisfying the following conditions*:

$$|g(x)| \leqslant (1/3)\mu \quad \text{for all } x \in X, \tag{28}$$

$$|g(x)| < (1/3)\mu \quad \text{for all } x \in X - F, \tag{29}$$

$$|f(x) - g(x)| \leqslant (2/3)\mu \quad \text{for all } x \in F. \tag{30}$$

Proof. Let

$$A = \{x: f(x) \leqslant (-1/3)\mu\} \quad \text{and} \quad B = \{x: f(x) \geqslant (1/3)\mu\}.$$

The sets $A$ and $B$ are disjoint and closed (see Remark 1 of § 3). The function

(31) $$g(x) = (1/3)\mu\{\varrho(x, A)-\varrho(x, B)\}/\{\varrho(x, A)+\varrho(x, B)\}$$

satisfies the required conditions in virtue of Lemma 1.

THEOREM 1 (TIETZE EXTENSION THEOREM). *Every continuous real valued function $f$ defined on a closed subset $F$ of the metric space $X$ can be extended to the entire space $X$; i.e. there exists a real valued continuous function $f^*$ defined on the entire space $X$ such that*

(32) $$f^*(x) = f(x) \quad \textit{for} \quad x \in F.$$

*Moreover, if $f$ is bounded, i.e.*

(33) $$|f(x)| \leqslant \mu \ (\neq 0) \quad \textit{for every} \quad x \in F,$$

*then*

(34) $$|f^*(x)| < \mu \quad \textit{for every} \quad x \in X-F.$$

Proof. Consider first the case where the function $f$ is bounded and hence satisfies inequality (33). We define a sequence of continuous functions $g_0, g_1, \ldots$ inductively. Let $g_0(x) = 0$ for every $x \in X$. For given $n \geqslant 0$ let us assume that the functions $g_0(x), \ldots, g_n(x)$ satisfy the inequality

(35) $$\Big|f(x)-\sum_{i=0}^{n} g_i(x)\Big| \leqslant (2/3)^n\mu \quad \text{for} \quad x \in F.$$

In the case $n = 0$ this inequality reduces to inequality (33).

Replacing in the assumptions of Lemma 2: $f(x)$ by $f(x)-\sum_{i=0}^{n} g_i(x)$ and $\mu$ by $(2/3)^n\mu$, we obtain a continuous function $g_{n+1}$ defined on $X$ and such that

(36) $$|g_{n+1}(x)| \leqslant (2^n/3^{n+1})\mu \quad \text{for} \quad x \in X,$$

(37) $$|g_{n+1}(x)| < (2^n/3^{n+1})\mu \quad \text{for} \quad x \in X-F,$$

(38) $$\Big|f(x)-\sum_{i=0}^{n+1} g_i(x)\Big| \leqslant (2/3)^{n+1}\mu \quad \text{for} \quad x \in F.$$

Thus the functions $g_n$ are defined for all $n = 0, 1, 2, \ldots$

For every $x \in X$ let us set

$$(39) \qquad f^*(x) = \sum_{i=0}^{\infty} g_i(x).$$

It follows from (36) that the series (39) is uniformly convergent in the space $X$; and hence by virtue of Theorem 3, § 3, the function $f^*$ is continuous.

Moreover, condition (35) implies condition (32), and because of inequality (37), we have for $x \in X-F$:

$$|f^*(x)| = \Big|\sum_{i=0}^{\infty} g_i(x)\Big| \leqslant \sum_{i=0}^{\infty} |g_{i+1}(x)| < \mu \sum_{i=0}^{\infty} (2^i/3^{i+1}) = \mu,$$

and therefore inequality (34) is also satisfied.

Thus the theorem has been proved for the case where the function $f$ is bounded.

If $f$ is unbounded, we first apply the homeomorphism $h$ which maps the space of all real numbers onto the open interval $-1 < y < 1$, e.g. $h(x) = (2/\pi)\arctan x$. The function $h \circ f$ (the composition of the functions $f$ and $h$) is continuous and bounded; hence there exists by virtue of the part of the theorem already proved a continuous function $h^*$ defined on the space $X$ and such that

$$h^*(x) = hf(x) \text{ for } x \in F, \qquad |h^*(x)| < 1 \text{ for } x \in X.$$

Now let

$$f^*(x) = h^{-1}h^*(x)$$

for every $x \in X$. The function $f^*$ is continuous and for every $x \in F$ we have

$$f^*(x) = h^{-1}hf(x) = f(x).$$

Thus the theorem has been proved in all generality.

COROLLARY 1. *Every continuous function defined on a closed subset $F$ of a metric space $X$ with values belonging to one of the spaces $\mathcal{E}^n, \mathcal{I}^n, \mathcal{E} \times \mathcal{E} \times \ldots, \mathcal{H}$ can be extended to the entire space $X$.*

We shall prove this corollary, e.g. for the Hilbert cube $\mathcal{H} = \mathcal{I} \times \mathcal{I} \times \ldots$ (the proof in the other cases is analogous).

For every $x \in F$ we have $f(x) \in \mathcal{I} \times \mathcal{I} \times \ldots$, and hence

$$(40) \qquad f(x) = [f_1(x), f_2(x), \ldots, f_n(x), \ldots],$$

where $f_n(x)$ is the $n$th coordinate of the point $f(x)$ in the Hilbert cube, hence a continuous real valued function. Extending each of the functions $f_n$ to a continuous function $f_n^*$ defined on the entire space $X$, we obtain a function

$$f^*(x) = [f_1^*(x), f_2^*(x), \ldots, f_n^*(x), \ldots] \tag{41}$$

which is the extension of the function $f$ (see Theorem 5, § 3).

COROLLARY 2. *Every continuous function $f$ defined on a closed subset $F$ of a metric space $X$ with values belonging to the sphere $\mathcal{S}_n$ (i.e. to the set of points $x_1^2 + \ldots + x_{n+1}^2 = 1$ of the space $\mathcal{E}^{n+1}$) can be extended to some neighbourhood of the set $F$ (with respect to the space $X$).*

P r o o f. By virtue of Corollary 1 there exists an extension $f^* \in (\mathcal{E}^{n+1})^X$ of the function $f$. Let us set

$$G = \{x: f^*(x) \neq 0\}.$$

Because of the continuity of the function $f^*$, $G$ is an open set containing the set $F$ (since $|f^*(x)| = |f(x)| = 1$ for $x \in F$). Thus the function

$$g(x) = f^*(x)/|f^*(x)|$$

is the required extension of the function $f$ onto the set $G$ which assumes values belonging to $\mathcal{S}_n$.

R e m a r k s. Spaces which can be substituted in Corollary 1 for $\mathcal{E}^n$, $\mathcal{I}^n$, etc., are called *absolute retracts*. Spaces which in Corollary 2 can be substituted for $\mathcal{S}_n$ are called *neighbourhood retracts*. (These concepts were introduced by K. Borsuk.)

This terminology is connected with the concept of retraction. We say, namely, that a subset $R$ of the space $X$ is a *retract* of this space if there exists a continuous transformation $f$ of the space $X$ onto the set $R$ such that $f(x) = x$ for $x \in R$ (this transformation is called a *retraction*; a projection is an example of a retraction).

Thus, an absolute retract is, as can be proved, a space which is a retract of every metric space containing it and in which it is closed. A neighbourhood retract is not necessarily a retract of the entire space, but of some one of its neighbourhoods in this space.

These concepts are important generalizations of the concepts of classical $n$-dimensional geometry: the $n$-dimensional cube is an absolute retract, every $n$-dimensional polyhedron is (as can be proved) a neighbourhood retract.

We are going now to extend Tietze theorem (which was proved for metric spaces) to arbitrary *normal* spaces. The proof will be based on a Lemma analogous to Lemma 1 (whose proof for metric spaces was immediate).

LEMMA 1′ (of Urysohn). *Given two disjoint closed sets $A$ and $B$ in a normal space $X$, there exists a continuous function $f\colon X \to \mathcal{J}$ such that*

(42) $f(x) = 0$ *for* $x \in A$ *and* $f(x) = 1$ *for* $x \in B$.

Proof. First we shall assign to every fraction of the form $r = k/2^n$ $(k = 0, 1, \ldots, 2^n)$, an open set $G(r)$ so that

(i) $A \subset G(0)$, $X - B = G(1)$,

(ii) the condition $r < r'$ implies $\overline{G(r)} \subset G(r')$.

We proceed by induction with respect to the exponent $n$.

For $n = 0$ the conditions (i) and (ii) are fulfilled by normality of $X$. Suppose that they are fulfilled for $n-1$. We must define $G(k/2^n)$ for an odd $k$. By hypothesis

$$\overline{G[(k-1)/2^n]} \subset G[(k+1)/2^n].$$

By normality of $X$, there exists an open set, which we denote by $G(k/2^n)$, such that

$$\overline{G[(k-1)/2^n]} \subset G(k/2^n) \quad \text{and} \quad \overline{G(k/2^n)} \subset G[(k+1)/2^n].$$

Thus the function $G(r)$ is defined for every $r$.

Let $f(x) = 0$ for $x \in G(0)$ and $f(x) =$ least upper bound of the $r$'s such that $x \in X - G(r)$ for $x \notin G(0)$. By (i), $f(x) = 0$ for $x \in A$ and $f(x) = 1$ for $x \in B$.

It remains to prove that the function $f$ is continuous, i.e. that for every $x_0$ and every natural number $n$ there exists an open set $H$ containing $x_0$ such that the condition $x \in H$ implies $|f(x_0) - f(x)| < 1/2^n$.

Let $r$ be a (finite dyadic) fraction such that

(43) $$f(x_0) < r < f(x_0) + 1/2^{n+1}.$$

Let $H = G(r) - \overline{G(r-1/2^n)}$ with the convention $G(s) = 0$ for $s < 0$ and $G(s) = X$ for $s > 1$. It follows that $x_0 \in H$. For, the inequality $f(x_0) < r$ implies $x_0 \in G(r)$, while the inequality $r - 1/2^{n+1} < f(x_0)$ implies

$$x_0 \in X - G(r-1/2^{n+1}) \subset X - \overline{G(r-1/2^n)}.$$

Moreover, the hypothesis $x \in H$ implies $x \in G(r)$, hence $f(x) \leqslant r$. It also implies

$$x \in X - \overline{G(r-1/2^n)} \subset X - G(r-1/2^n); \tag{44}$$

hence $r - 1/2^n \leqslant f(x)$. Therefore

$$f(x_0) - 1/2^n < f(x) < f(x_0) + 1/2^n. \tag{45}$$

This completes the proof.

LEMMA 2′. *Given a continuous function $f$ defined in a closed subset $F$ of a normal space $X$ and satisfying $|f(x)| \leqslant c$, where $c > 0$, there exists a continuous function $g$ defined in the whole space $X$ and satisfying the conditions*

$$|g(x)| \leqslant \tfrac{1}{3}c \quad \textit{for} \quad x \in X, \tag{46}$$

$$|f(x) - g(x)| \leqslant \tfrac{2}{3}c \quad \textit{for} \quad x \in F. \tag{47}$$

P r o o f. Let

$$A = \{x: f(x) \leqslant -\tfrac{1}{3}c\}, \quad B = \{x: f(x) \geqslant \tfrac{1}{3}c\} \tag{48}$$

and

$$J = \{y: |y| \leqslant \tfrac{1}{3}c\}. \tag{49}$$

The sets $A$ and $B$ being closed and disjoint, there is, by the Urysohn Lemma, a continuous function $g: X \to J$ such that $g(x) = -\frac{1}{3}c$ for $x \in A$ and $g(x) = \frac{1}{3}c$ for $x \in B$. Obviously $g$ satisfies the conditions of the lemma.

THEOREM 1′ (GENERALIZED TIETZE EXTENSION THEOREM). *Let $X$ be normal, $F$ closed and $f$: $F \to \mathscr{E}$ continuous. Then there is a continuous $f^*$ such that*

$$f \subset f^*: X \to \mathscr{E}. \tag{50}$$

Moreover, *$\mathscr{E}$ may be replaced by $\mathscr{I}$.*

If $f$ is bounded, the proof is completely analogous to the proof of Theorem 1 (except that condition (37) is not required).

The case of $f$ unbounded can be reduced to the former case as follows.

$\mathscr{E}$ being homeomorphic to the open interval $I_0 = (-1 < y < 1)$, we may assume that $f\colon F \to I_0$. As shown, there is a continuous extension $f^*\colon X \to \bar{I}_0$ of $f$. Let $B$ denote the two-element set $\{-1, 1\}$ and let $H = f^{*-1}(B)$. Then $H$ is closed and $H \cap F = \emptyset$. By the Urysohn Lemma, there is a continuous $h\colon X \to \mathscr{I}$ such that $h(x) = 0$ for $x \in H$ and $h(x) = 1$ for $x \in F$. Put $g(x) = f^*(x) \cdot h(x)$. Then $g\colon X \to I_0$ and $g$ is the required continuous extension of $f$.

## § 6. Completely regular spaces

Definition. A topological space $X$ is called *completely regular* if for each point $p$ and each closed set $F$ which does not contain $p$ there is a continuous mapping $f\colon X \to \mathscr{I}$ such that

(51) $$f(p) = 0 \quad \text{and} \quad f(x) = 1 \quad \text{for} \quad x \in F.$$

THEOREM 1. *Each completely regular space is regular.*

Proof. Let $p$, $F$, and $f$ be as defined previously. Put $G = \{x\colon f(x) < 1/2\}$. Then $G$ is open, $p \in G$, and $\bar{G} \cap F = 0$.

THEOREM 2. *Each normal $\mathscr{T}_1$-space is completely regular.*

This follows directly from the Urysohn Lemma (see § 5).

THEOREM 3. *The range of variability of $F$ in the definition of complete regularity can be restricted to sets such that $X - F$ belongs to a subbase of $X$.*

Proof. Let us first consider a finite system $F_1, \ldots, F_n$ of closed sets, a point $p \in X - F_0$ where $F_0 = F_1 \cup \ldots \cup F_n$, and a system of continuous functions $f_1, \ldots, f_n$ satisfying conditions (51). Put

(52) $$f(x) = \max_{1 \leqslant i \leqslant n} f_i(x).$$

Obviously $f$ satisfies conditions (51) for $F_0$.

Moreover, $f$ is continuous. This follows from the identity:

$$\{x\colon u < f(x) < v\} = \bigcup_{i=1}^{n} \{x\colon u < f_i(x)\} \cap \bigcap_{j=1}^{n} \{x\colon f_j(x) < v\},$$

according to which the set $\{x\colon u < f(x) < v\}$ is open for each pair of real numbers $u < v$.

Now let us consider an arbitrary closed $F$ such that $p \in X - F$. We shall show that there is a continuous $f$ satisfying (51).

Let $\boldsymbol{B}$ be a subbase of $X$. Then there is a system of members $F_1, \ldots, F_n$ such that $(X - F_i) \in \boldsymbol{B}$,

$$p \in X - F_0 \quad \text{and} \quad F \subset F_0 \quad \text{where} \quad F_0 = F_1 \cup \ldots \cup F_n.$$

The function $f$ defined by formula (52) obviously satisfies (51).

R e m a r k s. *A regular space may fail to be completely regular.*† Moreover, there are regular $\mathscr{T}_1$-spaces on which every real continuous function is constant.

There are also *completely regular spaces which are not normal* (see Chapter XI, Exercise 10); in fact, complete regularity is hereditary, while normality is not.

**Exercises**

1. Let the sets $A$ and $B$ be both open or both closed, and let $f$ be a mapping defined on the set $A \cup B$. Prove that if $f$ is continuous on the set $A$ and on the set $B$, then it is also continuous on the set $A \cup B$.

2. Let $f$ be defined on the space $X$. If the space $X$ is a union of open sets, and if on each of these sets individually $f$ is continuous, then $f$ is continuous on the entire space $X$.

3. Let $f$ be defined on the space $X$. If $X = \bigcup_{n=1}^{\infty} A_n$ where $A_n \subset \operatorname{Int}(A_{n+1})$ and if $f$ is continuous on each of the sets $A_n$, then it is continuous on the entire space $X$.

4. The set of all sequences of natural numbers forms a metric space (the so-called *Baire space*), if for the distance between distinct sequences $x = (m_1, m_2, \ldots)$ and $y = (n_1, n_2, \ldots)$ we take the number $1/r$, where $r$ is the smallest index such that $m_r \neq n_r$. Show that this space is homeomorphic to the set of all irrational numbers of the interval $[0, 1]$.

*Hint*: Assign the continued fraction

$$f(x) = \frac{1\,|}{|\,m_1} + \frac{1\,|}{|\,m_2} + \ldots$$

to the sequence of natural numbers $x = (m_1, m_2, \ldots)$.

5. A necessary and sufficient condition for the limit $f(x) = \lim_{n\to\infty} f_n(x)$ of the sequence of continuous mappings $f_1, f_2, \ldots$ defined in the space $X$ and with values in a metric space to be continuous is, that for every $\varepsilon > 0$, $X$ be the union of open sets $A_n(\varepsilon)$, where

$$A_n(\varepsilon) = \{x\colon |f_n(x) - f(x)| < \varepsilon\}.$$

*Hint*: In order to establish the continuity of $f$ under the assumption of our condition at an arbitrary point $x_0 \in X$, we find an index $n_0$ such that

† See R. Engelking, *Outline of General Topology*, p. 76.

$x_0 \in A_{n_0}(\varepsilon/3)$. Further, we make use of the fact that the set $A_{n_0}(\varepsilon/3)$ is open and the function $f_{n_0}$ is continuous.

**6.** Introducing a "new" distance into the metric space $X$ with the aid of the formula

$$\varphi(x, y) = |x-y|/\{1+|x-y|\},$$

we define a homeomorphism of $X$ onto $X$.

Deduce from this that the set of all sequences with real terms $x = (x_1, x_2, \ldots, x_m, \ldots)$ is a metric space under the following definition of distance:

$$|x-y| = \sum_{m=1}^{\infty} (1/2^m)|x_m-y_m|/\{1+|x_m-y_m|\}$$

(this is the so-called *Fréchet space*).

**7.** Let $B(X)$ denote the family of nonvoid closed bounded subsets of a metric space $X$. By the distance between two sets $A, B \notin B(X)$ we understand the maximum of the two numbers

least upper bound $\varrho(x, B)$ for $x \in A$, and least upper bound $\varrho(y, A)$ for $y \in B$.

Prove that the distance defined in this way, which we denote by the symbol dist$(A, B)$, metrizes the set $B(X)$ (i.e. it satisfies conditions (1)–(3) of Chapter IX, § 1).

**8.** A metric (bounded) space $X$ is called *totally bounded* iff for every $\varepsilon > 0$, there is a cover composed of a finite number of sets of diameter $< \varepsilon$.

Show that $X$ is totally bounded iff for every $\varepsilon > 0$, there is a finite set $F_\varepsilon$ such that, for each $x$, $\varrho(x, F_\varepsilon) < \varepsilon$.

**9.** Show that if $X$ is totally bounded, then so is $B(X)$ (see Ex. 7)

**10.** Let $X_1, X_2, \ldots$ be totally bounded spaces such that $\delta(X_n) \leqslant 1$. Show that their product $X_1 \times X_2 \times \ldots$ is totally bounded (the distance being defined by formula (7) of Chapter IX).

**11.** Show that in Theorem 6 of § 4, the property of $A$ and $B$ of being closed and disjoint sets can be replaced by the weaker assumption of being separated. Deduce that each metric space is hereditarily normal.

**12.** Show that the Hilbert cube $\mathscr{H}$ is homeomorphic to the subset of the Hilbert space (cf. Chapter IX, § 1, Example 3) composed of points $x = (x_1, x_2, \ldots, x_i, \ldots)$ such that $0 \leqslant x_i \leqslant 1/i$.

CHAPTER XIII

# CARTESIAN PRODUCTS

## § 1. Cartesian product $X \times Y$ of topological spaces

$X$ and $Y$ being two topological spaces, the topology in $Z = X \times Y$ is introduced in the following way:

Definition. A set $\mathfrak{G} \subset X \times Y$ is called *open in* $X \times Y$ iff it is the union of cartesian products $G \times H$ where $G$ and $H$ are open subsets of $X$ and $Y$ respectively.

In other words, the family of all sets $G \times H$ is a *base* of $X \times Y$.

By Chapter III, (21), $(G_1 \times H_1) \cap (G_2 \times H_2) = (G_1 \cap G_2) \times \times (H_1 \cap H_2)$.

Therefore, assuming that $G_1, G_2, H_1$ and $H_2$ are open sets (in $X$ and $Y$ respectively), the intersection $(G_1 \times H_1) \cap (G_2 \times H_2)$ is open (in $X \times Y$). It follows that the intersection of any two open sets in $X \times Y$ is open.

As the union of an arbitrary family of open sets in $X \times Y$ is open, we have the following theorem:

THEOREM 1. *The cartesian product of two topological spaces is a topological space.*

As $(X \times Y) - (x_0, y_0) = ((X - x_0) \times Y) \cup (X \times (Y - y_0))$ we have:

THEOREM 2. *The cartesian product of two $\mathcal{T}_1$-spaces is a $\mathcal{T}_1$-space.*

EXAMPLE. In the case of the plane $\mathcal{E}^2$, the usual topology agrees with the above definition. For every open set in $\mathcal{E}^2$ can be represented as the union of open squares with sides parallel to the $X$ and $Y$ axes.

Another theorem can also be easily shown (compare Chapter IV, Exercise 3).

THEOREM 3. *If $\{B_t\}$ is a base of $X$, and $\{C_s\}$ a base of $Y$, then $\{B_t \times C_s\}$ is a base of $X \times Y$.*

The same remains true of subbases.

THEOREM 4. *$X$ and $Y$ being topological spaces, the family of sets $G\times Y$ and of sets $X\times H$, where $G$ is open in $X$ and $H$ in $Y$, is a subbase of $X\times Y$.*

Because $G\times H = (G\times Y) \cap (X\times H)$.

Remark. Recall that a relation is a subset of $X\times Y$, namely the set $\{\langle x, y\rangle: x\varrho y\}$. Consequently a *relation* will be called *closed* if this set is closed (in $X\times Y$).

## § 2. Projections and continuous mappings

Given $z = \langle x, y\rangle \in X\times Y$ consider $x$ as function of $z$. Put $x = \pi_1(z)$ and similarly $y = \pi_2(z)$. Thus

$$\pi_1: X\times Y \to X \quad \text{and} \quad \pi_2: X\times Y \to Y,$$

where $\pi_1$ and $\pi_2$ are called *projections* of $X\times Y$ on the axes $X$ and $Y$; $\pi_1(z)$ is the abscissa of $z$ and $\pi_2(z)$ its ordinate.

THEOREM 1. *The projections are continuous mappings.*

Proof. If $G$ is open in $X$, we have $\pi_1^{-1}(G) = G\times Y$, which is open by definition. Thus $\pi_1$ is continuous.

THEOREM 2. *Let $h$: $T \to X\times Y$. Put $h(t) = \langle h_1(t), h_2(t)\rangle$ where $h_1(t) \in X$ and $h_2(t) \in Y$. Then $h$ is continuous iff $h_1$ and $h_2$ are continuous.*

More precisely, *$h$ is continuous at $t_0$ iff $h_1$ and $h_2$ are continuous at $t_0$.*

Proof. Suppose that $h$ is continuous at $t_0$. As $h_1(t) = \pi_1 h(t)$, it follows by Theorem 1 that $h_1$ is continuous at $t_0$.

Suppose that $h_1$ and $h_2$ are continuous at $t_0$. Let $\mathfrak{G} \subset X\times Y$ be open and let $t_0 \in h^{-1}(\mathfrak{G})$. We have to show that $t_0 \in \mathrm{Int}\, h^{-1}(\mathfrak{G})$. According to Corollary 3 of Chapter XII, § 1, we can assume that $\mathfrak{G}$ belongs to a subbase of $X\times Y$. Put (see § 1, Theorem 4) $\mathfrak{G} = G\times \times Y$. Hence $h^{-1}(\mathfrak{G}) = h_1^{-1}(G)$ and therefore $t_0 \in h_1^{-1}(G)$. As $h_1$ is continuous at $t_0$, it follows that $t_0 \in \mathrm{Int}\, h_1^{-1}(G) = \mathrm{Int}\, h^{-1}(\mathfrak{G})$.

THEOREM 3. *Every continuous mapping of two variables is continuous relative to each variable.*

In other words, *if $f$: $X\times Y \to W$ is continuous and if $y_0 \in Y$, then the mapping $f_1$: $X \to W$ defined by the condition $f_1(x) = f(x, y_0)$ is continuous.*

P r o o f. Define the mapping $h\colon X \to X \times \{y_0\}$ by the condition $h(x) = \langle x, y_0 \rangle$. Obviously $h$ is continuous (more precisely, $h$ is a homeomorphism), and as $f_1 = f \circ h$, $f_1$ is continuous.

R e m a r k. Elementary examples show that the converse of Theorem 3 does not hold.

THEOREM 4. *Let* $f\colon U \to X$ *and* $g\colon V \to Y$. *The product-mapping* $h = (f \times g)\colon U \times V \to X \times Y$ (compare Chapter IV, Exercise 17) *is continuous iff* $f$ *and* $g$ *are continuous.*

P r o o f. By definition $h(u, v) = \langle f(u), g(v) \rangle$. Hence, if we put, for a given $v_0$, $h_1(u) = h(u, v_0) = \langle f(u), g(v_0) \rangle$, then $f = \pi_1 \circ h_1$. If we assume that $h$ is continuous, then so is $f$ (by Theorem 3); similarly, $g$ is continuous.

On the other hand, if $f$ and $g$ are continuous, then so is $h$. This follows from the identity (see Chapter IV, Exercise 17):

$$h^{-1}(G \times H) = f^{-1}(G) \times g^{-1}(H).$$

## § 3. Invariants of cartesian multiplication

THEOREM 1. *The product of two closed sets is closed.*

Let $A$ be closed in $X$ and $B$ in $Y$. By Chapter III, (24)

$$(X \times Y) - (A \times B) = [(X - A) \times Y] \cup [X \times (Y - B)].$$

Thus $(X \times Y) - (A \times B)$ is the union of two open sets, hence is open.

THEOREM 2. *The following properties are invariant under cartesian multiplication:*

(a) *of being a* $\mathscr{T}_2$*-space,*

(b) *of being completely regular.*

P r o o f. (a) Suppose that $X$ and $Y$ are $\mathscr{T}_2$-spaces. Let $z_1 = \langle x_1, y_1 \rangle$, $z_2 = \langle x_2, y_2 \rangle$, and $z_1 \neq z_2$. Then, either $x_1 \neq x_2$ or $y_1 \neq y_2$. We may suppose that $x_1 \neq x_2$. Since $X$ is a $\mathscr{T}_2$-space, there are two open sets $G_1 \subset X$ and $G_2 \subset X$ such that $x_1 \in G_1$, $x_2 \in G_2$, and $G_1 \cap G_2 = \emptyset$. It follows that $z_1 \in (G_1 \times Y)$, $z_2 \in (G_2 \times Y)$, and that $G_1 \times Y$ and $G_2 \times Y$ are open and disjoint.

(b) Let $z_0 = \langle x_0, y_0 \rangle \in \big((X \times Y) - \mathfrak{F}\big)$ where $\mathfrak{F}$ is closed. In order to show that $X \times Y$ is completely regular, we may restrict ourselves to the case where $\mathfrak{F} = A \times Y$ and $A = \overline{A} \subset X$ (see Theorem 3 of Chapter XII, § 6). It follows that $x_0 \in X - A$. Now, as $X$ is supposed to be completely regular, there exists a continuous mapping $g$: $X \to \mathscr{I}$ such that $g(x_0) = 0$ and $g(x) = 1$ for $x \in A$. Put $f(x, y) = g(x)$. Thus, $f(z_0) = 0$ and $f(z) = 1$ for $z \in \mathfrak{F}$.

Remark. One can show also that *regularity* is invariant under cartesian multiplication and that *normality* is not; in fact, there is a normal $X$ such that $X \times \mathscr{I}$ is not normal (M. E. Rudin).

## § 4. Diagonal

We call the set

$$\Delta = \{\langle x, y\rangle : (x = y)\}$$

the *diagonal* of $X^2 = X \times X$.

THEOREM 1. $\Delta \underset{\text{top}}{=} X$.

The required homeomorphism is the projection $\langle x, y\rangle \to x$.

THEOREM 2. *If $X$ is a $\mathscr{T}_2$-space, the diagonal is closed (in $X^2$).*

Proof. Put $V = X^2 - \Delta$. We have to show that $V$ is open, i.e. that given a point $(x, y)$ of $V$, there are two open sets $G$ and $H$ such that $x \in G$, $y \in H$ and $G \times H \subset V$. Now, as $x \neq y$ there are ($X$ being $\mathscr{T}_2$) two open sets $G$ and $H$ such that $x \in G$, $y \in H$, and $G \cap H = \emptyset$. Hence $(G \times H) \cap \Delta = \emptyset$, i.e. $(G \times H) \subset V$.

Remark. The assumption of $X$ being $\mathscr{T}_2$ is essential. In fact, if the diagonal of $X \times X$ is closed, $X$ is $\mathscr{T}_2$.

The following theorem is an important generalization of Theorem 2.

THEOREM 3. *If $f$: $X \to Y$ is continuous, its graph*

$$F = \{\langle x, y\rangle : y = f(x)\}$$

*is homeomorphic to $X$.*

*If $Y$ is a $\mathscr{T}_2$-space, $F$ is closed (in $X \times Y$).*

Proof. Put $h(x) = \langle x, f(x)\rangle$. As in the proof of Theorem 1, one sees that $h$ is a homeomorphism of $X$ onto $F$.

To prove the second part, put $g(x, y) = \langle f(x), y\rangle$. Obviously

$$[g(x, y) \in \Delta] \equiv [y = f(x)] = [\langle x, y\rangle \in F], \quad \text{i.e.} \quad F = g^{-1}(\Delta).$$

Since $g$ is continuous (by Theorem 4 of § 2) and $\Delta$ is closed, so $F$ is closed.

## § 5. Generalized cartesian products

Let us now consider the general case

$$Z = \prod_{t\in T} X_t,$$

where $T$ is an arbitrary set and $X_t$ is a topological space.

As in the case of two factors, we denote by $\pi_t(f)$ the $t$th coordinate of $f \in Z$, i.e. (compare Chapter IV, (32))

(1) $$\pi_t(f) = f(t), \quad \text{hence} \quad \pi_t\colon Z \to X_t.$$

$\pi_t$ is the *projection* of $Z$ on the $X_t$-axis.

We introduce topology in $Z$ (called *Tychonov topology*) by means of the following definition.

Definition. The family of sets of the form

(2) $$Q_{t,G} = \pi_t^{-1}(G) = \{f\colon f(t) \in G\},$$

where $G$ is open in $X_t$, is a subbase of $Z$.

Thus $Q_{t,G}$ is the product of $G$ and of all $X_{t'}$ with $t' \neq t$. Consequently, the products $\prod_t G_t$ where $G_t$ is open in $X_t$ and, except for a finite set of indices, is identical with $X_t$, form a base of $Z$.

It follows also that if $T$ is composed of two elements, the above definition agrees with the definition given in § 1.

The theorems of the preceding paragraphs can be easily extended to generalized products. In particular, we have the following statements.

THEOREM 1. *$\pi_t$ is continuous.*

Because $\pi_t^{-1}(G)$ is open in $Z$ if $G$ is open in $X_t$.

THEOREM 2. *Let $h\colon W \to Z$. $h$ is continuous iff $h_t = \pi_t \circ h$ is continuous for each $t \in T$.*

Because, if $h_t$ is continuous, then $h^{-1}(Q_{t,G})$ is open since

$$h^{-1}(Q_{t,G}) = h^{-1}[\pi_t^{-1}(G)] = h_t^{-1}(G).$$

THEOREM 3. *If $A_t = \overline{A_t} \subset X_t$, then $\prod_t A_t = \overline{\prod_t A_t} \subset Z$.*

Since $\pi_t$ is continuous, the set $\pi_t^{-1}(A_t)$ is closed, and so is $\prod_t A_t$ because

$$\prod_t A_t = \bigcap_t \pi_t^{-1}(A_t).$$

R e m a r k. The range of variability of $G$ in (2) can be restricted to a *subbase* $\boldsymbol{S}$ of $X_t$.

Because, if $G$ is open in $X_t$, then $\pi_t^{-1}(G)$ is generated by the sets $\pi_t^{-1}(H)$, where $H \in \boldsymbol{S}$, in the same way that $G$ is generated by the sets $H$ (with the aid of the union operation and the finite intersection operation).

THEOREM 4. *The properties of being a $\mathcal{T}_1$-space, a $\mathcal{T}_2$-space, a completely regular space, are invariant under the generalized cartesian multiplication.*

The proofs are similar to those given in § 3.

## § 6. $X^T$ considered as a topological space. The cube $\mathcal{I}^T$

Let $X$ be a topological space and $T$ an arbitrary set. $X^T$ is obviously a particular case of the product $\prod_t X_t$, namely when $X_t = X$ for each $t \in T$. Thus $X^T$ can be regarded as a topological space with the Tychonov topology. In other words, it has a subbase composed of sets $Q_{t,G}$ defined by formula (2) of § 5, where $G$ is an open subset of $X$ and $\pi_t$ is the evaluation of $X^T$ at $t$, i.e. (compare Chapter IV, (32)):

$$\pi_t(f) = f(t), \quad \text{hence} \quad \pi_t\colon X^T \to X. \tag{1}$$

Theorems 1 and 2 of § 5 imply the two following theorems.

THEOREM 1. *The mapping $\pi_t\colon X^T \to X$ is continuous.*

THEOREM 2. *Let $W$ be a topological space. Then the mapping $\varphi\colon W \to X^T$ is continuous iff each mapping $\varphi_t = \pi_t \circ \varphi$ is continuous.*

THEOREM 3. *Let $T$ be a topological space. Assume that the elements of $\Phi \subset X^T$ are continuous mappings. Then the evaluation $e$ of $\Phi$ is a continuous mapping of $T$ into $X^\Phi$.*

P r o o f. Let $g_f$ denote the evaluation of $X^\Phi$ at $f \in \Phi$; this means that (compare Chapter IV, (34)):

$$g_f(h) = h(f) \text{ for each } h \in X^\Phi, \text{ hence } g_f : X^\Phi \to X. \tag{2}$$

Therefore $g_f \circ e = f$ for each $f \in \Phi$ (compare Chapter IV, (35)), and it follows by Theorem 2 that, since the mapping $g_f \circ e$ is continuous (because $f$ is continuous), so is $e$.

R e m a r k. We will now consider the important case where $X = \mathscr{I}$. Thus $\mathscr{I}^T$ is a generalized cube. In particular $\mathscr{I}^N$ is the Hilbert cube ($N$ denoting the space of positive integers).

THEOREM 4. *Every completely regular $\mathscr{T}_1$-space $X$ is topologically contained in a generalized cube.*

More precisely, *let $e$ denote the evaluation of $\Phi = (\mathscr{I}^X)_{\text{top}}$ (the set of all continuous mappings $\varphi$: $X \to \mathscr{I}$). Then $e$: $X \to \mathscr{I}^\Phi$ is a homeomorphism.*

P r o o f. Recall that (see Chapter IV, (33))

(3) $\quad [e_x](f) = f(x) \quad \text{for} \quad f \in \Phi, \quad \text{hence} \quad e_x\colon \Phi \to \mathscr{I}.$

In order to show that $e$ is one-to-one, suppose that $x_1 \neq x_2$. As $X$ is a completely regular $\mathscr{T}_1$-space, there is $f \in \Phi$ such that $f(x_1) \neq f(x_2)$. Since $e_x(f) = f(x)$, we have $e_{x_1}(f) \neq e_{x_2}(f)$ and hence $e_{x_1} \neq e_{x_2}$.

It remains to show that the inverse mapping to $e$ is continuous. In other words, that if $G \subset X$ is open, then $e(G)$ is open in $e(X)$; that means that, if $x_0 \in G$ there is $Q \subset Z = \mathscr{I}^\Phi$ open and such that

(4) $\quad e_{x_0} \in Q \quad \text{and} \quad [Q \cap e(X)] \subset e(G).$

As $X$ is completely regular, there is $f \in \Phi$ such that

(5) $\quad f(x_0) = 0 \quad \text{and} \quad f(x) = 1 \quad \text{for} \quad x \in X - G.$

Define $Q$ by the condition

(6) $\quad (h \in Q) \equiv [g_f(h) \neq 1], \quad \text{i.e.} \quad Q = Z - g_f^{-1}(1).$

Since the mapping $g_f$ is continuous, the set $g_f^{-1}(1)$ is closed, hence $Q$ is open. Moreover $e_{x_0} \in Q$ since by (2) and (3)

$$g_f(e_{x_0}) = e_{x_0}(f) = f(x_0) \neq 1.$$

To show the second part of (4), note that by (6)

(7) $$Q \cap e(X) = e(X) - g_f^{-1}(1),$$

and since (see the Remark to Chapter V, § 1)

$$f \circ e^{-1} \subset g_f, \quad \text{hence} \quad [f \circ e^{-1}]^{-1} \subset g_f^{-1}, \text{ i.e. } e \circ f^{-1} \subset g_f^{-1},$$

it follows that $e[f^{-1}(1)] \subset g_f^{-1}(1)$. Therefore, by (7) and Chapter IV, (15a),

$$Q \cap e(X) \subset e(X) - e[f^{-1}(1)] \subset e[X - f^{-1}(1)] \subset e(G)$$

since $X - f^{-1}(1) \subset G$.

COROLLARY. *A $\mathcal{T}_1$-space is completely regular iff it is topologically contained in a generalized cube.*

Proof. Since a generalized cube is a product of completely regular spaces (namely, of intervals), it is completely regular (by Theorem 4 of § 5). It remains to refer to the fact that each subset of a completely regular space is completely regular (Chapter XII, § 6, Remarks).

Remark. The second part of Theorem 4 remains true and the proof remains valid if the set $(\mathcal{I}^X)_{\text{top}}$ is reduced to a set $\Phi$ such that

(8) if $x_0 \notin F = \overline{F} \subset X$, there is $f \in \Phi$ for which $f(x_0) \notin \overline{f(F)}$.

It follows that, if $\Phi$ is countable, $X$ can be imbedded in the Hilbert cube $\mathcal{I}^N$. Such is the case of a metric space $X$ with a countable base (see Chapter XIV, § 4).

## § 7. Cartesian products of metric spaces

The product $\prod_n X_n$ of a finite or infinite sequence of metric spaces $X_1, X_2, \ldots$ can be considered as a metric space if we use the definition of distance given by formulas (4) and (7) of Chapter IX. We are going to show that the topology induced by this distance agrees with the topology (of Tychonov) defined in this chapter.

THEOREM. *A set $H \subset Z = X_1 \times X_2 \times \ldots$ is open in the metric sense iff $H$ is open in the Tychonov topology.*

Proof. With no loss of generality (see Chapter XII, § 3) we can assume that $\delta(X_n) \leqslant 1$ for each $n$ and that the distance is defined by formula (7) of Chapter IX.

First, suppose that $H$ is open in the metric sense. Let $z \in H$. Hence there is $\varepsilon > 0$ such that $K(z, \varepsilon) \subset H$. Let $n$ be such that $2^{-n} < \varepsilon/2$. Put $z = (x_1, x_2, \ldots)$ where $x_i \in X_i$, and let $G_i = K(x_i, \varepsilon/2) \subset X_i$ for $i \leqslant n$ and let $G_j = X_j$ for $j > n$. Therefore $z \in G_1 \times G_2 \times \ldots \subset H$; because, if $z' \in G_1 \times G_2 \times \ldots$, it follows from (7) that $|z'-z| < \varepsilon$, i.e. $z' \in K(z, \varepsilon)$.

Thus $H$ is the union of some sets belonging to the base of $Z$ (compare § 5), and consequently $H$ is open in the Tychonov topology.

It remains to show that each member $Q$ of the subbase of $Z$, considered in § 5, is open in the metric sense. We can assume, of course, that $Q = G \times X_2 \times X_3 \times \ldots$ where $G$ is open in $X_1$. Let $z = (x_1, x_2, \ldots) \in Q$. Hence $x_1 \in G$. Put $\varepsilon = \varrho(x_1, X_1 - G)$, hence $\varepsilon > 0$. It follows that $K(z, \varepsilon/2) \subset Q$, and consequently $Q$ is open in the metric sense.

**Exercises**

**1.** Let $A \subset X, B \subset Y$. Prove the following formulas:

$$\mathrm{Int}(A \times B) = \mathrm{Int}(A) \times \mathrm{Int}(B),$$

$$\mathrm{Fr}(A \times B) = [\mathrm{Fr}(A) \times \bar{B}] \cup [\bar{A} \times \mathrm{Fr}(B)].$$

**2.** A necessary and sufficient condition for the cartesian product $A \times B$ to be dense in itself is that one of the sets $A$ and $B$ be dense in itself.

**3.** Suppose that, for each $a \in A$, $X_a$ is a topological space. Let $B$ and $C$ be disjoint subsets of $A$ such that $A = B \cup C$. Then the product space $\Pi_{b \in B} X_b \times \times \Pi_{c \in C} X_c$ is homeomorphic to the product space $\Pi_{a \in A} X_a$.

For each fixed topological space $X$ the product $X^A$ is homeomorphic to $X^B \times X^C$ and $(X^B)^C$ is homeomorphic to $X^{B \times C}$, all spaces being given the product topology.

**4.** Let $F$ be a closed subset of the metric space $X$, and let

$$f(x) = 1/\varrho(x, F) \quad \text{for} \quad x \in X - F.$$

Prove that the set

$$\{\langle x, y\rangle: [y = f(x)](x \notin F)\}$$

is closed in the space $X \times \mathscr{E}$.

Deduce from this that every open set in $X$ is homeomorphic to a closed subset of the space $X \times \mathscr{E}$ (making use of Theorem 2, § 4).

**5.** Let $Q$ be a $G_\delta$ subset of the metric space $X$, i.e. $Q = G_1 \cap G_2 \cap \ldots \cap G_n \cap \ldots$, where $G_n$ is an open set. Let

$$f_n(x) = 1/\varrho(x, X - G_n) \text{ for } x \in G_n, \quad \text{and} \quad f(x) = [f_1(x), f_2(x), \ldots].$$

Prove that the set

$$\{\langle x, y\rangle: [y = f(x)](x \in Q)\}$$

is closed in the space $X \times \mathscr{E} \times \mathscr{E} \times \mathscr{E} \times \ldots$

Deduce from this that every $G_\delta$-set is homeomorphic to a closed subset of the space $X \times \mathscr{E} \times \mathscr{E} \times \mathscr{E} \times \ldots$

**6.** Let $\{T, F, g\}$ be an inverse system (comp. Chapter VIII, § 5). Let $F_t$, for $t \in T$, be a metric space (or more generally a completely regular space), and $g_{t_0 t_1}$ be a continuous mapping (for $t_0 < t_1$).

Prove that:

(1°) the set $Z = \underleftarrow{\mathrm{Lim}}\{T, F, g\}$ is closed in $\prod_{t \in T} F_t$ (with Tychonov's topology) and hence is completely regular;

(2°) sets of the form $Z \cap \{f: f_t \in G\}$, where $t \in T$ and $G$ is an arbitrary open set in $F_t$, form a base of $Z$;

(3°) if $M \subset Z$, then $f \in \overline{M} \equiv \bigwedge_t f_t \in \overline{M_t}$, where $M_t$ is the projection of $M$ onto $F_t$;

(4°) if the mappings $h_t$ in Chapter VII, § 5, are homeomorphisms, then $h_\infty$ is a homeomorphism.

**7.** A continuous mapping which maps open sets onto open sets is called an *open* mapping (similarly we define *closed* mappings).

Prove that:

(a) an open (closed) one-to-one mapping is a homeomorphism;

(b) the projection of $X \times Y$ onto $X$ is an open mapping;

(c) extend the preceding theorem to generalized cartesian products.

**8.** We say that a *uniform structure* on $X$ is defined with respect to a (non-empty) family $U$ of subsets of the cartesian product $X \times X$, if the following axioms hold:

(a) if $V \in U$, then $V$ contains the diagonal of $X \times X$ (i.e. the set $\{\langle x, x\rangle: x \in X\}$),

(b) if $V \in U$, then the set $\{\langle x, y\rangle: \langle y, x\rangle \in V\}$ belongs to $U$,

(c) if $V_1, V_2 \in U$, then $(V_1 \cap V_2) \in U$,

(d) if $V \in U$ and $V \subset Z$, then $Z \in U$,

(e) if $V \in U$, then there exists a set $V_1 \in U$ such that the set

$$\{\langle x, y\rangle: \bigvee_z [\langle x, z\rangle \in V_1] [\langle z, y\rangle \in V_1]\}$$

is a subset of $V$.

Prove that

(f) if $\bigcap U$ is the diagonal of $X \times X$, then $X$ becomes a completely regular $\mathscr{T}_1$-space, if we assume that

$$x \in \overline{A} \equiv \bigwedge_V \{y: \langle y, x\rangle \in V\} \cap A \neq \emptyset, \quad \text{where} \quad V \in U.$$

Prove that a uniform structure is given to every metric space $X$ by the family of all sets containing sets of the form

$$V_\varepsilon = \{\langle x, y\rangle: |x-y| < \varepsilon\}, \quad \text{where} \quad \varepsilon > 0.$$

**9.** A *proximity relation* on the set $X$ is a relation $A \delta B$ defined for $A, B \subset X$ and such that

(a) the relation $\delta$ is symmetric,

(b) $[A \delta (B \cup C)] \equiv [(A \delta B)(A \delta C)]$,

(c) $\bigwedge_{x,y} (x \delta y) \equiv (x = y)$,

(d) $\emptyset$ non-$\delta X$,

(e) if $A$ non-$\delta B$, then there exists a pair of sets $C$ and $D$ such that

$$A \subset C, \quad B \subset D, \quad C \cap D = \emptyset, \quad A \text{ non-}\delta(X-C), \quad B \text{ non-}\delta(X-D).$$

Prove that $X$ is completely regular, if the closure is defined by the formula

$$(x \in \overline{A}) \equiv (x \delta A).$$

**10.** Let $X$ be a completely regular $\mathcal{T}_1$-space. Assume that $A$ non-$\delta B$ iff there exists $f \in \mathcal{I}^X$ such that $f(A) = 0$ and $f(B) = 1$. Prove that the relation $\delta$ is a proximity relation.

Prove that in an arbitrary metric space a proximity relation can be defined by the equivalence

$$(A \delta B) \equiv [\varrho(A, B) = 0] \quad (\text{where } A \neq \emptyset \neq B).$$

## CHAPTER XIV

# SPACES WITH A COUNTABLE BASE

### § 1. General properties

Let $G_1, G_2, \ldots$ denote the base (composed of open nonvoid sets) of the given $\mathcal{T}_1$-space $X$. This means that every open set is the union of some of the sets $G_n$.

THEOREM 1. *The property of containing an open countable base is hereditary.*

Namely, if $E \subset X$, then $E \cap G_1, E \cap G_2, \ldots$ is an open base of $E$.

THEOREM 2 (of Lindelöf). *Every space $X$ with a countable base is a Lindelöf space.*

This follows immediately from the Remark of § 11 of Chapter X.

* THEOREM 3 (BROUWER REDUCTION THEOREM). *Let $X$ be a space with a countable base and let $A$ be a family of closed subsets of $X$ such that, for each decreasing sequence $F_0 \supset F_1 \supset \ldots$ of members of $A$, the intersection $F_0 \cap F_1 \cap \ldots$ belongs to $A$.*† *Then each $M_0 \in A$ contains an irreducible member (i.e. a set $M \in A$ such that no proper closed subset of $M$ belongs to $A$).*

Proof. We assume that $M$ is the intersection of a sequence of members $M_0 \supset M_1 \supset \ldots$ of $A$ defined by induction as follows. Let $n > 0$ be a given integer. Let $M_{n-1} \in A$. If there is $B \in A$ such that $B \subset M_{n-1} - G_n$, then $M_n$ is such a set $B$ (thus in this case $M_n \cap G_n = \emptyset$); if such a $B$ does not exist, then we put $M_n = M_{n-1}$. Thus in any case $M_n \in A$, and consequently $M \in A$.

We have to show that $M$ is an irreducible member of $A$. Suppose it is not. So let $B \in A$, $B \subset M$ and $B \neq M$. Accordingly there is $G_n$ such that $M \cap G_n \neq \emptyset$ and $B \cap G_n = \emptyset$. Therefore $B \subset \subset M - G_n \subset M_{n-1} - G_n$ and (by the definition of $M_n$) $M_n \cap G_n = \emptyset$. But then $M \cap G_n = \emptyset$, which is impossible.

† A family of sets having the above property is called *inducible*.

THEOREM 4. *If the spaces $X_1, X_2, \ldots$ have a countable base, then so does $X_1 \times X_2 \times \ldots$*

P r o o f. Let $G_{n,1}, G_{n,2}, \ldots$ denote the base of $X_n$. According to the Remark of Chapter XIII, § 5, the base of $X_1 \times X_2 \times \ldots$ is composed of sets of the form $H_1 \times H_2 \times \ldots$ where for each index $n$, except a finite number of indices, $H_n = X_n$, while for the exceptional indices $H_n$ is a term of the sequence $G_{n,1}, G_{n,2}, \ldots$ This base is obviously countable (compare Chapter V, § 3, Theorem 5).

## § 2. Separable spaces

D e f i n i t i o n. A space is said to be *separable* if it contains a countable dense subset.

Hence, a metric space is separable if it contains a sequence of points $p_1, p_2, \ldots$ such that every point $p$ is of the form

$$p = \lim_{n \to \infty} p_{k_n}. \tag{1}$$

The space of all real numbers is a separable space, for the set of rational numbers is countable and dense. An example of a space which is not separable, is an arbitrary uncountable set in which $|x-y| = 1$ for every pair of points $x \neq y$.

THEOREM 1. *Every space with a countable base is separable.*

For the proof, it suffices (according to the axiom of choice) to choose a point $p_n$ from each $G_n$.

THEOREM 2. *Every metric separable space contains a countable base.*

*Thus, for metric spaces, the concepts of a separable space and of a space with a countable base are equivalent.*

P r o o f. Let $p_1, p_2, \ldots$ be a dense sequence in the given metric space. Let us consider the balls with centres $p_n$ and with rational radii:

$$K_{n,r} = \{x\colon (|x-p_n|) < r\}. \tag{2}$$

The set of these balls is countable (cf. Chapter V, § 3, Theorem 3) and forms a base.

In fact, for an arbitrary point $p$ and every number $\varepsilon > 0$, there exists a point $p_n$ such that $|p-p_n| < \varepsilon$. Let $r$ be a rational number such that $|p-p_n| < r < \varepsilon$. Then $p \in K_{n,r}$ and $\delta(K_{n,r}) < 2\varepsilon$, and hence the sets $K_{n,r}$ form a base.

## § 3. Theorems on cardinality in spaces with countable bases

We assume in § 3 that the space $X$ under consideration has a countable base.

THEOREM 1. *The family of all open sets is of power* $\leqslant \mathfrak{c}$.

*The same applies to closed sets.*

P r o o f. Let $G_1, G_2, \ldots$ be a base of the space. Hence to every open set $H$ there corresponds a sequence of natural numbers $k_1, k_2, \ldots$ such that

$$H = \bigcup_{n=1}^{\infty} G_{k_n}. \tag{3}$$

It follows that the number of open sets is $\leqslant$ the number of all sequences of natural numbers, i.e. is $\leqslant \mathfrak{c}$.

The second part of the theorem follows immediately from the first, for, if we assign to each open set its complement, then we map the family of open sets onto the family of closed sets in a one-to-one manner.

THEOREM 2. *Every $\mathscr{T}_1$-space with a countable base has power* $\leqslant \mathfrak{c}$.

This follows at once from the second part of Theorem 1.

* R e m a r k. More generally, we can prove that the *family of all Borel sets has power* $\leqslant \mathfrak{c}$. Hence, if the space is of power $\mathfrak{c}$, it contains non-Borel sets; and furthermore, since the family of all subsets of this space has power $2^{\mathfrak{c}}$, the family of non-Borel subsets has power $> \mathfrak{c}$ (and therefore, e.g., on the real line there exist more non-Borel than Borel sets).

THEOREM 3. *Every family $\boldsymbol{R}$ of disjoint open sets is countable.*

P r o o f. Let $p_1, p_2, \ldots$ be a sequence dense in the space under consideration. Hence, if $H$ is a non-empty set belonging to the family $\boldsymbol{R}$, then there exists an index $n$ such that $p_n \in H$; we denote this index by $n(H)$; if $\emptyset \in \boldsymbol{R}$ then we put $n(\emptyset) = 0$. We have therefore assigned to each non-empty set $H$ belonging to $\boldsymbol{R}$ a number $n(H)$ so that

$$p_{n(H)} \in H. \tag{4}$$

Distinct numbers correspond to distinct sets. For if $n(H_1) = n(H_2)$, then by (4) we have

$$p_{n(H_1)} \in H_1 \cap H_2,$$

which is possible only if $H_1 = H_2$ (because the sets belonging to the family $\boldsymbol{R}$ are disjoint).

Therefore, there are at most as many elements of the family $\boldsymbol{R}$ as there are non-negative integers, which was to be proved.

THEOREM 4. *The set of isolated points is countable.*

P r o o f. Since each isolated point of the space constitutes an open set (see Chapter XI, § 5, Theorem 2), it follows that the one-element sets, whose single element is an isolated point, form a family of disjoint open sets. This family is countable by virtue of Theorem 3, and hence the set of isolated points is also countable.

COROLLARY. *Let $Z \subset X$. Then the set of isolated points of $Z$ is countable.*

The set $Z$, being a subset of a space with countable base, can itself be considered as such (by virtue of Theorems 1, § 2, and 1, § 1).

THEOREM 5. *If the spaces $X$ and $Y$ have countable bases, then the space $Y^X$ (i.e. the set of continuous mappings $f$: $X \to Y$) has power $\leqslant \mathfrak{c}$.*

P r o o f. By virtue of Theorem 2, Chapter XIII, § 4, if $f \in Y^X$, then $f$ is a closed set in the space $X \times Y$; but since the latter space has a countable base (Theorem 4, § 1), the family of all its closed subsets has power $\leqslant \mathfrak{c}$ (Theorem 1).

R e m a r k. If the space $Y$ has power $\mathfrak{c}$, then the space $Y^X$ has the same power, because the set of constant functions is then of power $\mathfrak{c}$. Under the assumption that the space $X$ also has power $\mathfrak{c}$, we note that there are more discontinuous than continuous mappings, because the set of all mappings of $X$ into $Y$ has power $\mathfrak{c}^{\mathfrak{c}} > \mathfrak{c}$ (cf. Chapter VI, § 4, (45)).

## § 4. Imbedding in the Hilbert cube

URYSOHN THEOREM. *Every separable metric space $X$ is homeomorphic to a subset of the Hilbert cube $\mathscr{H}$, i.e.*

$$X \underset{\text{top}}{\subset} \mathscr{H}.$$

P r o o f. By the Theorem 8 of § 3, Chapter XII, we can assume that

$$\delta(X) \leqslant 1.$$

Let $p_1, p_2, \ldots$ be a sequence of points dense in the space $X$. To each $x \in X$ we assign the point of the Hilbert cube with "coordinates": $|x-p_1|, |x-p_2|, \ldots,$ i.e.

$$h(x) = (|x-p_1|, |x-p_2|, \ldots, |x-p_n|, \ldots). \tag{5}$$

The functions

$$h_n(x) = |x-p_n| \tag{6}$$

are continuous (Chapter XII, § 4, Theorem 5), and therefore, by Theorem 2, Chapter XIII, § 5, the function $h$ is also continuous. We shall prove that this function is a homeomorphism.

We assume that

$$\lim_{k\to\infty} h(x_k) = h(x). \tag{7}$$

We must show that

$$\lim_{k\to\infty} x_k = x. \tag{8}$$

Let $\varepsilon > 0$. Since the sequence $p_1, p_2, \ldots$ is dense in the space $X$, there exists a point $p_j$ such that

$$|x-p_j| < \varepsilon. \tag{9}$$

It follows from formulas (7) and (5) that

$$\lim_{k\to\infty} h_j(x_k) = h_j(x).$$

Because of (6) this means that

$$\lim_{k\to\infty} |x_k - p_j| = |x-p_j|;$$

therefore, there exists a $k_0$ such that

$$|x_k-p_j| < |x-p_j|+\varepsilon \tag{10}$$

provided that $k > k_0$.

By the inequalities (9) and (10), we therefore have

$$|x-x_k| \leqslant |x-p_j|+|p_j-x_k| < 3\varepsilon$$

for $k > k_0$. This means that formula (8) is valid.

R e m a r k 1. Since every subset of the Hilbert cube is a separable metric space, it follows from the above theorem that *from the topological point of view separable metric spaces are equivalent to subsets of the Hilbert cube.*

* R e m a r k 2. Instead of assuming that $X$ is metric separable, we could have assumed that $X$ is a $\mathcal{T}_1$ *normal space with a countable base.* In this case the proof would run as follows.

Let $G_1, G_2, \ldots$ be a base of $X$. Consider all pairs $\langle i,j \rangle$ such that $\overline{G_i} \subset G_j$. By the Urysohn Lemma (Chapter XII, § 5), there is a continuous function $f_{ij}\colon X \to \mathcal{I}$ such that

$$(11)\qquad f_{ij}(x) = 0 \text{ for } x \in \overline{G}_j \quad\text{ and }\quad f_{ij}(x) = 1 \text{ for } x \in X - G_j .$$

Arrange the double sequence $\{f_{ij}\}$ into a simple sequence $\{g_n\}$ and put

$$(12)\qquad h(x) = (g_1(x), g_2(x), \ldots) \in \mathcal{H} .$$

We shall show that $h$ is a homeomorphism:

As the continuity of $h$ follows from the continuity of $g_n$, it remains to show that $p \notin \overline{A}$ implies $h(p) \notin \overline{h(A)}$.

Now, as $X$ is a normal $\mathcal{T}_1$-space and $G_1, G_2, \ldots$ is its base, it is easy to see that there exists a pair of indices $\langle i,j \rangle$ such that $p \in G_i$ and $\overline{G_i} \subset G_j \subset X - \overline{A}$. Put $g_n = f_{ij}$. By (11), $g_n(p) = 0$ and $g_n(x) = 1$ for $x \in A$. According to the definition of the distance in $\mathcal{H}$ (see Chapter IX, (7)), we get for $x \in A$

$$(13)\qquad |h(x) - h(p)| \geqslant 1/2^n, \quad\text{ i.e. }\quad h(A) \cap K[h(p), 1/2^n] = \emptyset .$$

Thus $h(p) \notin \overline{h(A)}$.

Let us add that another proof of the theorem under consideration was referred to in Chapter XIII, § 6 (Remark to Theorem 4).

## *§ 5. Condensation points. The Cantor–Bendixson theorem

A point $p$ of a set $A$ is said to be a *condensation point* of $A$ if every neighbourhood of $p$ contains a non-countable set of points of the set $A$.

We denote the set of condensation points of the set $A$ by the symbol $A^0$.

Every condensation point of the set $A$ is an accumulation point of $A$, i.e.

(14) $$A^0 \subset A^d.$$

It is also easy to prove that the set $A^0$ is closed, i.e.

(15) $$A^0 = \overline{A^0},$$

and that

(16) $$(A \cup B)^0 = A^0 \cup B^0.$$

The following generalization of Theorem 4, § 3, is valid:

THEOREM 1. *In a space $X$ with a countable base the set $A-A^0$ is countable.*

Proof. Let $G_1, G_2, \ldots$ be a base of the space. Let $p \in A-A^0$. Then there exists a neighbourhood $K$ of $p$ such that $A \cap K$ is countable. There exists also an index $n(p)$ such that $p \in G_{n(p)} \subset K$, whence $A \cap G_{n(p)} \subset A \cap K$, and therefore the set $A \cap G_{n(p)}$ is countable.

Since the union of a countable number of countable sets is countable (Chapter V, § 3, Theorem 4), the set

$$S = \bigcup_p A \cap G_{n(p)},$$

where $p \in A - A^0$, is countable. Now, we have $A - A^0 \subset S$, for $p \in A \cap G_{n(p)}$. Therefore the set $A-A^0$ is countable.

Since a countable set clearly has no point of condensation, it follows from the theorem that

(17) $$(A-A^0)^0 = 0.$$

From this we deduce that

(18) $$X^0 = X^{00}.$$

In fact, the identity $X = X^0 \cup (X-X^0)$ yields, by virtue of (16) and (17), that

$$X^0 = X^{00} \cup (X-X^0)^0 = X^{00}.$$

THEOREM 2. *Every $\mathscr{T}_1$-space $X$ which contains a countable base and does not contain non-empty sets dense in themselves (i.e. a scattered space) is countable.*

Proof. By virtue of (18) and (14), we have $X^0 = X^{00} \subset X^{0d}$; i.e. $X^0 \subset X^{0d}$, which means that the set $X^0$ is dense in itself. Therefore $X^0 = \emptyset$ by hypothesis; it follows that $X = X - X^0$, and this last set is countable by virtue of Theorem 1.

THEOREM 3 (CANTOR–BENDIXSON). *Every $\mathcal{T}_1$-space with a countable base is the union of two disjoint sets, one dense in itself and closed (i.e. perfect) and the other countable.*

This is an immediate consequence of the preceding theorem and of Theorem 3, Chapter XI, § 7.

**Exercises**

1. Define a dense sequence in the Hilbert cube $\mathcal{H}$.

2. Let $X$ be the set of all real numbers having the topology induced by the family $\boldsymbol{B}$ of all half-open intervals $(a, b) = \{x: a \leqslant x < b\}$, considered as base.

Prove that

(a) The members of the base $\boldsymbol{B}$ are both open and closed.

(b) The space $X$ is separable but has no countable base.

*Hint*: Notice that for every $x \in X$ each base contains a set whose infimum is $x$.

(c) Every subspace of $X$ is separable.

3. Show that the space considered in Exercise 1, Chapter IX, is not separable.

*Hint*: Show that there exists a continuum of disjoint open sets in this space.

4. Prove that $A^0 - B^0 \subset (A-B)^0$.

5. Prove that

$$\left(\bigcap_t A_t\right)^0 \subset \bigcap_t A_t^0, \qquad \bigcup_t A_t^0 \subset \left(\bigcup_t A_t\right)^0.$$

6. Assign to every ordinal number $\xi < \alpha$ an open set $A_\xi$ lying in the space $X$ (with a countable base), so that $A_{\xi+1} \subset A_\xi$ and $A_{\xi+1} \neq A_\xi$. Prove that $\alpha < \Omega$ (i.e. that there is but a countable number of sets $A_\xi$).

*Hint*: Let $G_1, G_2, \ldots$ be a base of the space $X$. Assign to each $\xi$ (with perhaps the exception of the last one) a number $n(\xi)$ such that

$$G_{n(\xi)} \subset A \quad \text{and} \quad G_{n(\xi)} - A_{\xi+1} \neq \emptyset.$$

7. Prove the analogous theorem for closed $A_\xi$.

8. Deduce the following corollary from the above theorem: every set of real numbers which is well ordered with respect to the "less than" relation is countable.

9. The *derived sets of transfinite order* are defined inductively by means of the formulas (where the space is $\mathcal{T}_1$ with a countable base):

$$X^{(1)} = X^d, \quad X^{(\xi+1)} = (X^{(\xi)})^d, \quad X^{(\lambda)} = \bigcap_{\xi<\lambda} X^{(\xi)} \ (\lambda \text{ a limit ordinal}).$$

Prove (making use of Exercise 7) that beginning with some $\alpha < \Omega$ the derived sets of all orders are equal.

**10.** Deduce the Cantor-Bendixson Theorem from the above theorem making use of Theorem 4, § 3.

**11.** Prove that every totally bounded space is separable (comp. Chapter XII, Exercise 8).

**12.** The product $\Pi_t X_t$ has a countable base iff each $X_t$ has a countable base and all but a countable number of $X_t$ have the trivial topology (comp. § 1, Theorem 3).

**13.** We say that a topological space satisfies the *first axiom of countability* if for every point $x$ there exists a countable family of open sets such that every neighbourhood of $x$ contains a member of that family.

(i) Give an example of a space which satisfies the first axiom of countability but has no countable base.

(ii) Prove that every metric space satisfies the first axiom of countability.

(iii) Let $X_t$ be a topological space satisfying the first axiom of countability for each $t \in T$. Then the product $\Pi_t X_t$ satisfies the first axiom of countability if and only if all but a countable number of the spaces $X_t$ are trivial.

**14.** A space is said to be *locally separable* at the point $p$ if there is a separable neighbourhood of $p$. Give an example of a metric space which is locally separable at none of its points.

*Hint*: Use a construction analogous to the construction used in Exercise 1, Chapter IX.

## CHAPTER XV

# COMPLETE SPACES

### § 1. Complete spaces

D e f i n i t i o n. We say that a sequence of points $p_1, p_2, \ldots$ in a metric space is a *Cauchy sequence* if for every $\varepsilon > 0$ there exists a $k$ such that for every $n > k$ we have

$$|p_n - p_k| < \varepsilon, \tag{1}$$

i.e. if

$$\bigwedge_\varepsilon \bigvee_k \bigwedge_n [(n > k) \Rightarrow (|p_n - p_k| < \varepsilon)].$$

A metric space is said to be *complete* if every Cauchy sequence $p_1, p_2, \ldots$ is convergent, that is, there exists a point $p$ of this space such that $p = \lim\limits_{n\to\infty} p_n$.

The space of all real numbers is complete according to the known Cauchy theorem from analysis. Let us note that completeness is not a topological property of the space. The space of all real numbers is homeomorphic to the open interval $0 < x < 1$ (see Chapter XII, § 2) which is not a complete space inasmuch as the sequence $1/2, 1/3, 1/4, \ldots$ is a Cauchy sequence but is not convergent (in this space).

THEOREM. *Every convergent sequence in an arbitrary metric space is a Cauchy sequence.*

P r o o f. In fact, if the sequence $p_1, p_2, \ldots$ is convergent to the point $p$, then for every $\varepsilon > 0$ there exists a $k$ such that for every $n \geqslant k$ we have the inequality

$$|p_n - p| < \varepsilon/2. \tag{2}$$

In particular, for $n = k$ we have

$$|p_k - p| < \varepsilon/2. \tag{3}$$

For $n \geqslant k$, inequality (1) follows from the inequalities (2) and (3).

**§ 2. Cantor theorem.** *Let $\{F_n\}$ be a decreasing sequence of non-empty closed sets in a complete space*:

$$F_1 \supset F_2 \supset \ldots \supset F_n \supset F_{n+1} \supset \ldots \tag{4}$$

*If*

$$\lim_{n\to\infty} \delta(F_n) = 0, \tag{5}$$

*then*

$$\bigcap_{n=1}^{\infty} F_n \neq \emptyset. \tag{6}$$

Proof. Let $p_n \in F_n$. Then $p_1, p_2, \ldots$ is a Cauchy sequence. In fact, by virtue of (5), there exists for every $\varepsilon > 0$ a $k$ such that $\delta(F_n) < \varepsilon$ provided $n \geqslant k$.

By (4), $p_n \in F_n \subset F_k$, and hence for $n \geqslant k$

$$p_n, p_k \in F_k, \quad \text{whence} \quad |p_n - p_k| \leqslant \delta(F_k) < \varepsilon;$$

i.e. $p_1, p_2, \ldots$ is a Cauchy sequence. Since the space is complete, this sequence is convergent. Hence, let $p = \lim_{n\to\infty} p_n$.

For every $m$, the terms of the sequence $p_1, p_2, \ldots$ with the exception of at most the first $m-1$ terms belong to $F_m$, and since the set $F_m$ is closed, the limit of this sequence also belongs to $F_m$, i.e.

$$p \in F_m \text{ for } m = 1, 2, \ldots, \quad \text{i.e.} \quad p \in \bigcap_{m=1}^{\infty} F_m.$$

Remark. The set $\bigcap_{m=1}^{\infty} F_m$ consists of only one point $p$.

**§ 3. Baire theorem.** *In a non-empty complete space $X$ the union*

$$E = F_1 \cup F_2 \cup \ldots \cup F_k \cup \ldots \tag{7}$$

*of closed boundary sets cannot fill the entire space*; *furthermore, this union is a boundary set.*†

Proof. In order to prove that the set $E$ is a boundary set in the space $X$, it suffices to show that every neighbourhood $S_0$ of an arbitrary point contains points of the set $X-E$ (see Chapter XI, § 2, Theorem 2).

Since the closed set $F_1$ is a boundary set, there exists a ball $S_1$ such that $\overline{S}_1 \subset S_0$ and $\overline{S}_1 \cap F_1 = \emptyset$ (see Chapter XI, § 2, Theorem 3). Clearly, we can assume that $\delta(S_1) < 1$.

† Sets of the form (7) (where the sets $F_k$ are closed boundary sets), as well as all their subsets, are said to be sets of the *first category*.

Similarly, we find an $S_2$ such that $\bar{S}_2 \subset S_1, \bar{S}_2 \cap F_2 = \emptyset$ and $\delta(S_2) < 1/2$.

Continuing in this manner, we obtain a sequence of balls which satisfy the conditions

$$S_0 \supset \bar{S}_1 \supset \bar{S}_2 \supset \ldots \supset \bar{S}_n \supset \ldots, \tag{8}$$

$$\bar{S}_n \cap F_n = \emptyset \tag{9}$$

and

$$\delta(S_n) < 1/n, \quad \text{whence} \quad \lim_{n\to\infty} \delta(S_n) = 0. \tag{10}$$

From the Cantor theorem, we deduce by virtue of (8) and (10) that there exists a point $p$ belonging to all the sets $\bar{S}_n$. Therefore (by (9))

$$p \in \bigcap_{n=1}^{\infty} \bar{S}_n \subset \bigcap_{n=1}^{\infty} (X - F_n) = X - \bigcup_{n=1}^{\infty} F_n,$$

hence by (7) $p \in X - E$. Also $p \in S_0$.

R e m a r k s. 1. Since a subset of a boundary set is a boundary set, Baire theorem can also be stated in the following manner: *in a complete space every set of the first category is a boundary set.*

2. It follows from the Baire theorem that every nonvoid complete dense in itself space is noncountable.

In fact, if the space were countable: $X = (p_1, p_2, \ldots)$, then it would be the union of a sequence of sets each consisting of one point: $X = \{p_1\} \cup \{p_2\} \cup \ldots$ But each of these sets is a closed boundary set, inasmuch as each of the points $p_n$ is an accumulation point of the space $X$.

Since the space $\mathscr{E}$ of real numbers is complete and dense in itself, we have thus obtained another proof of the inequality $\mathfrak{c} > \mathfrak{a}$.

3. The set of irrational numbers is not an $F_\sigma$-set in the space $\mathscr{E}$ (and therefore the set of rational numbers is not a $G_\delta$-set).

For, if the opposite were true, the set of irrational numbers would be the union of a countable number of closed boundary sets (because the set of irrational numbers is itself a boundary set). But since the set of rational numbers is the union of a countable number of one-element sets—and hence of closed boundary sets—

the entire space $\mathcal{E}$ could be represented as the union of a countable number of closed boundary sets; but this contradicts Baire theorem.

## § 4. Extension of a metric space to a complete space

LEMMA. *Let $X$ be a topological space and $Y$ a metric complete space. Then the space $\Phi(X, Y)$ of all continuous bounded mappings $f\colon X \to Y$ metrized by the formula* (6) *of Chapter* IX *is complete.*

Proof. Suppose that $|f_n - f_k| < \varepsilon$ for $n > k$. Then, for each $x$, $|f_n(x) - f_k(x)| < \varepsilon$ and $f_1(x), f_2(x), \ldots$ is a Cauchy sequence, hence is convergent. Put $f(x) = \lim_{n\to\infty} f_n(x)$. The convergence is uniform. Because, for every $m > k$, we have

$$|f_m(x) - f_k(x)| < \varepsilon, \quad \text{hence} \quad |f_n(x) - f_m(x)| < 2\varepsilon.$$

Consequently $|f_n(x) - \lim_{m\to\infty} f_m(x)| \leqslant 2\varepsilon$ and finally $|f_n(x) - f(x)| \leqslant 2\varepsilon$.

It follows that $f \in \Phi(X, Y)$ since the limit of a uniformly convergent sequence of continuous bounded functions is continuous and bounded (compare Chapter XII, § 3, Theorem 3 and Chapter IX, § 7, Theorem 2). On the other hand $f = \lim_{n\to\infty} f_n$ (by Theorem 1 of Chapter IX, § 7).

THEOREM. *Each metric space $X$ is isometric with a subset of a complete space.*

More precisely: let us define $\Phi = \Phi(X, \mathcal{E})$ as in the Lemma, let $a$ be a fixed point of $X$ and let

$$h_p(x) = |x-p| - |x-a|; \tag{11}$$

then $h_p \in \Phi$ for each $p \in X$, and $h\colon X \to \Phi$ is an isometric mapping, which means (according to the general definition of isometric mappings) that

$$|h_p - h_q| = |p-q|. \tag{12}$$

Proof. $h_p$ is bounded since $|h_p(x)| \leqslant |p-a|$ by (11). According to the Lemma, we have to prove (12). Now

$$|h_p(x) - h_q(x)| = \big||x-p| - |x-q|\big| \leqslant |p-q|,$$

hence $|h_p - h_q| \leqslant |q-p|$. On the other hand

$$h_p(p) - h_q(p) = -|p-a| - |p-q| + |p-a|,$$

hence $|h_p - h_q| \geqslant |p-q|$, and (12) follows.

Remark. If $X$ is bounded, the definition of $h_p$ may be simplified. We can assume, namely, that $h_p(x) = |x-p|$.

**Exercises**

**1.** Show by means of an example that Baire theorem is not valid in arbitrary metric spaces.

**2.** The cartesian product $X \times Y$ of two complete spaces, metrized with the aid of the formula

$$|\langle x_1, y_1 \rangle - \langle x_2, y_2 \rangle| = \{|x_1 - x_2|^2 + |y_1 - y_2|^2\}^{1/2},$$

is complete.

**3.** The cartesian product $X_1 \times X_2 \times X_3 \times \ldots$ of complete spaces is complete if the distance between two points $x = (x_1, x_2, \ldots)$ and $y = (y_1, y_2, \ldots)$ is defined by the formula

$$|x-y| = \sum_{n=1}^{\infty} (1/2^n)|x_n - y_n|/(1 + |x_n - y_n|).$$

**4.** Prove that every $G_\delta$-set lying in a complete space is homeomorphic to a complete space (Alexandrov theorem).

*Hint*: Use Exercise 5 of Chapter XIII.

**5.** Let $f$ be a continuous mapping of a complete space and let the sequence $F_1, F_2, \ldots$ satisfy the assumption of Cantor theorem. Then

$$f\left(\bigcap_{n=1}^{\infty} F_n\right) = \bigcap_{n=1}^{\infty} f(F_n).$$

**6.** Prove that the set $Z$ of convergence points of a sequence $f_1, f_2, \ldots$ of mappings with values in a complete space satisfies the equivalence

$$(x \in Z) \equiv \bigwedge_k \bigvee_m \bigwedge_i |f_{m+i}(x) - f_m(x)| \leqslant 1/k.$$

Infer that if $f_i$ are continuous mappings, then $Z$ is $F_{\sigma\delta}$ (Hahn's theorem).

**7.** Prove that if $X$ is complete, then so is $B(X)$ (comp. Chapter XII, Exercise 7).

**8.** Let $X$ be metric (bounded). Denote by $\alpha(X)$ the infimum of all $\varepsilon$ such that there exists a finite cover composed of sets of diameter $< \varepsilon$ (compare the concept of a totally bounded space, Chapter XII, Exercise 8).

Prove the following generalization of Cantor theorem.

Let $\{F_t\}$, $t \in T$, be a family of closed sets such that

(i) each finite intersection of the sets $F_t$ is non-empty,

(ii) $\inf \alpha(F_t) = 0$;

then $\bigcap_t F_t \neq \emptyset$.

**9.** Define a space which is bounded, complete and separable,† but is not totally bounded.

† A complete separable space is called a *Polish space* (see Bourbaki).

CHAPTER XVI

# COMPACT SPACES

## § 1. Definition

A topological space $X$ is called *compact* if every open cover of $X$ contains a finite cover of $X$.

In other words, if $\{G_t\}$ is a family of open sets such that $\bigcup_t G_t = X$, where $t$ ranges over an arbitrary set $T$, then there is a finite system $t_1, \ldots, t_n$ such that $X = G_{t_1} \cup \ldots \cup G_{t_n}$.

It is easy to show that the above condition, called the *Borel–Lebesgue condition*, is equivalent to the following *Riesz condition*: if $\{F_t\}$ is a family of closed sets such that $\bigcap_t F_t = \emptyset$, then there is a finite system $t_1, \ldots, t_n$ such that $F_{t_1} \cap \ldots \cap F_{t_n} = \emptyset$.

EXAMPLES. The interval $\mathscr{I}$, the cube $\mathscr{I}^n$ are compact. More generally, each closed and bounded subset of the space $\mathscr{E}^n$ is compact (see § 5).

## § 2. Fundamental properties of compact spaces

THEOREM 1. *Each compact subset of a $\mathscr{T}_2$-space is closed.*

P r o o f. Let $A \subset X$ be compact. We have to show that $X - A$ is open, i.e. that given a point $b \in X - A$, there is an open $G$ such that $b \in G \subset X - A$.

Since $X$ is a $\mathscr{T}_2$-space, there is for each $x \in A$ a pair of open sets $U_x$ and $V_x$ such that

$$b \in U_x, \quad x \in V_x \quad \text{and} \quad U_x \cap V_x = 0.$$

Consequently, the family of sets $A \cap V_x$, where $x \in A$, is an open cover of $A$ (considered as a space). Since $A$ is compact, there is a finite system $x_1, \ldots, x_n$ such that

$$A = (A \cap V_{x_1}) \cup \ldots \cup (A \cap V_{x_n}), \quad \text{i.e.} \quad A \subset V_{x_1} \cup \ldots \cup V_{x_n}.$$

Put $G = U_{x_1} \cap \ldots \cap U_{x_n}$. Therefore $G$ is open and $b \in G \subset X - A$.

THEOREM 2. *Each closed subset of a compact space is compact.*

P r o o f. Let $F = \overline{F} \subset X$. Let $\{G_t\}$, $t \in T$, be a cover of $F$, where $G_t$ is open relative to $F$. Hence there is an $H_t$ open (relative to $X$) such that $G_t = F \cap H_t$. Consequently the family of sets $H_t$, where $t \in T$, augmented of the set $H = X - F$, is an open cover of $X$. Since $X$ is compact, there is a finite system $t_1, \ldots, t_n$ such that $X = H \cup H_{t_1} \cup \ldots \cup H_{t_n}$. Hence $F = G_{t_1} \cup \ldots \cup G_{t_n}$.

THEOREM 3. *The image under a continuous mapping of a compact space is compact.*

P r o o f. Let $X$ be compact and $f\colon X \to Y$ continuous and onto. Let $\{G_t\}$ be an open cover of $Y$. Hence $\{f^{-1}(G_t)\}$ is an open cover of $X$. Since $X$ is compact, there are $t_1, \ldots, t_n$ such that

$$X = f^{-1}(G_{t_1}) \cup \ldots \cup f^{-1}(G_{t_n}), \quad \text{hence} \quad Y = G_{t_1} \cup \ldots \cup G_{t_n}.$$

COROLLARY 1. *Each continuous mapping of a compact space into a $\mathcal{T}_2$-space is a closed mapping* (i.e. if $F \subset X$ is closed, so is $f(F)$).

P r o o f. Let $X$ be compact and $f\colon X \to Y$ continuous. Let $F \subset X$ be closed. By Theorem 2, $F$ is compact. Therefore, by Theorem 3, $f(F)$ is compact, hence closed by Theorem 1 (since $Y$ is a $\mathcal{T}_2$-space).

COROLLARY 2. *Each one-to-one and continuous mapping of a compact space into a $\mathcal{T}_2$-space is a homeomorphism.*

P r o o f. Put $g = f^{-1}$. Hence $g^{-1}(F) = f(F)$, and since $F$ is closed, so is $g^{-1}(F)$.

THEOREM 4. *Each compact $\mathcal{T}_2$-space is normal (hence completely regular).*

P r o o f. Let $A$ and $B$ be two closed disjoint sets. As in the proof of Theorem 1, one can assign to each $b \in B$ two open sets $G_b$ and $H_b$ such that

$$b \in G_b, \quad A \subset H_b \quad \text{and} \quad G_b \cap H_b = \emptyset$$

(in fact, it is sufficient to put $G_b = U_{x_1} \cap \ldots \cap U_{x_n}$ and $H_b = V_{x_1} \cup \ldots \cup V_{x_n}$).

Since $B$ is compact (by Theorem 2), it follows that $B \subset G_{b_1} \cup \ldots \cup G_{b_m}$. Put $G = G_{b_1} \cup \ldots \cup G_{b_m}$, $H = H_{b_1} \cap \ldots \cap H_{b_m}$. Hence $A \subset H$, $B \subset G$ and $G \cap H = \emptyset$.

Moreover, $G$ and $H$ are open. This completes the proof.

THEOREM 5. *A topological space which is the union of two compact sets is compact.*

We leave the proof to the reader.

## § 3. Cartesian products

THEOREM 1. *Let $X$ be an arbitrary topological space and let $Y$ be compact. Then the projection of $X\times Y$ on the $X$-axis is a closed mapping of $X\times Y$ onto $X$.*

Equivalently: *if $G$ is open in $X\times Y$, then the set $Q$ of all points $x$ of $X$ such that $(x)\times Y\subset G$ is open in $X$.*

P r o o f. We shall prove the theorem in the second formulation. By the definition of product topology, $G$ is of the form $G=\bigcup_t G_t\times H_t$ where $G_t$ is open in $X$ and $H_t$ in $Y$. Let $x_0$ be a fixed point of $Q$, i.e. $(x_0)\times Y\subset G$. Hence for each $y\in Y$ there is an index $t(y)$ such that $\langle x_0, y\rangle\in G_{t(y)}\times H_{t(y)}$. Thus

$$(1)\qquad x_0\in G_{t(y)},\quad y\in H_{t(y)}\quad \text{and}\quad G_{t(y)}\times H_{t(y)}\subset G.$$

Since the family $\{H_{t(y)}\}$, where $y$ ranges over $Y$, is an open cover of $Y$, there is a finite system $y_1, \ldots, y_n$ such that

$$(2)\qquad Y=H_{t(y_1)}\cup\ldots\cup H_{t(y_n)}.$$

Put $R(x_0)=G_{t(y_1)}\cap\ldots\cap G_{t(y_n)}$. Hence $R(x_0)$ is open, $x_0\in R(x_0)$ by (1), and by virtue of (2) and (1):

$$[R(x_0)\times Y]\subset[(G_{t(y_1)}\times H_{t(y_1)})\cup\ldots\cup(G_{t(y_n)}\times H_{t(y_n)})]\subset G.$$

Thus $R(x_0)\subset Q$. Since $R(x_0)$ is open and contains $x_0$, it follows that $Q$ is open. This completes the proof.

The main object of this section is to show that the cartesian product of compact spaces is a compact space. We shall start with the case of two spaces (which is simpler).

THEOREM 2. *The cartesian product $X\times Y$ of two compact spaces is compact.*

P r o o f. Let $C$ be a cover of $X\times Y$. We have to show that it contains a finite subcover. Since the sets $G\times H$, where $G$ is open in $X$ and $H$ in $Y$, form a base, we are allowed to assume that $C=\{G_t\times H_t\}$ (see the Remark of Chapter X, § 11).

Let $R$ be the family of all open sets $Q \subset X$ such that

$$(3) \qquad Q\times Y \subset (G_{t_1}\times H_{t_1}) \cup \ldots \cup (G_{t_n}\times H_{t_n})$$

for an appropriate system $t_1, \ldots, t_n$.

We shall show that $R$ is a cover of $X$.

Let $x_0 \in X$. Since $\{x_0\}\times Y$ is homeomorphic to $Y$ and $Y$ is compact, so $\{x_0\}\times Y$ is contained in a finite subcover of $C$:

$$\{x_0\}\times Y \subset (G_{t_1}\times H_{t_1}) \cup \ldots \cup (G_{t_n}\times H_{t_n}).$$

By Theorem 1, there is an open $Q$ containing $x_0$ and satisfying (3). Thus $R$ is a cover of $X$.

Since $X$ is compact, $R$ contains a finite subcover: $X = Q_1 \cup \cup \ldots \cup Q_k$. Since we have by (3), for $i \leqslant k$,

$$Q_i\times Y \subset (G_{t_{i,1}}\times H_{t_{i,1}}) \cup \ldots \cup (G_{t_{i,n(i)}}\times H_{t_{i,n(i)}}),$$

so

$$X\times Y = \bigcup_{i=1}^{k} Q_i\times Y \subset \bigcup_{i=1}^{k}\bigcup_{j=1}^{n(i)} G_{t_{i,j}}\times H_{t_{i,j}},$$

and the right-hand side is a finite subcover of $C$.

This completes the proof for the case of the product of two spaces. The proof of the general case will be based on the Alexander Lemma.

Let us call *essentially infinite* every cover which contains no finite subcover. (Thus compact spaces are spaces which have no essentially infinite subcover.)

ALEXANDER LEMMA. *Let $A$ be an open subbase of the topological space $X$. Suppose there exists an essentially infinite open cover of $X$. Then there exists such a cover contained in $A$.*

Proof. Denote by $\mathfrak{M}$ the totality of all essentially infinite open covers of $X$. By assumption $\mathfrak{M} \neq \emptyset$. We shall show first that $\mathfrak{M}$ has the following property: whenever $\{P_\alpha\}$ is a transfinite (see Chapter VIII, § 7) monotone sequence (i.e. $\alpha < \beta \Rightarrow P_\alpha \subset P_\beta$) of members of $\mathfrak{M}$, then $(\bigcup_\alpha P_\alpha) \in \mathfrak{M}$.

$\bigcup_\alpha P_\alpha$ is obviously a cover of $X$. It is essentially infinite. For otherwise, it would contain a finite cover $G_1, \ldots, G_n$ and consequently there would be a finite system $\alpha_1, \ldots, \alpha_n$ such that $G_i \in P_{\alpha_i}$. Denote by $\beta$ the greatest $\alpha_i$ $(1 \leqslant i \leqslant n)$. Hence $G_i \in P_\beta$

for each $i = 1, \ldots, n$ and it follows that $P_\beta$ is not essentially infinite.

From the property of $\mathfrak{M}$ just shown, it follows (see Chapter VIII, § 8) that $\mathfrak{M}$ contains a maximal element. Denote it by $\boldsymbol{P}$. Thus, if $H$ is open and does not belong to $\boldsymbol{P}$, then $\boldsymbol{P} \cup \{H\}$ is not essentially infinite, which means that there is a finite system $G_1, \ldots, G_n$ such that

$$(4) \quad H \cup G_1 \cup G_2 \cup \ldots \cup G_n = X \quad \text{and} \quad G_i \in \boldsymbol{P} \text{ for } i = 1, \ldots, n.$$

We shall show that the family of open sets which do not belong to $\boldsymbol{P}$ is a filter, i.e. that ($H$ and $G$ being open)

$$(5_1) \qquad H_1 \notin \boldsymbol{P} \quad \text{and} \quad H_2 \notin \boldsymbol{P} \text{ imply } (H_1 \cap H_2) \notin \boldsymbol{P}$$

and

$$(5_2) \qquad H \notin \boldsymbol{P} \quad \text{and} \quad H \subset G \text{ imply } G \notin \boldsymbol{P}.$$

The condition $H_j \notin \boldsymbol{P}$, $j = 1, 2$, implies (see (4)) the existence of sets $G_{j,1} \ldots, G_{j,n_j}$ such that

$$H_j \cup G_{j,1} \cup G_{j,2} \cup \ldots \cup G_{j,n_j} = X \quad \text{and} \quad G_{j,i} \in \boldsymbol{P}.$$

It follows that

$$(H_1 \cap H_2) \cup \bigcup_{j,i} G_{j,i} = X,$$

hence $(H_1 \cap H_2) \notin \boldsymbol{P}$, since $\boldsymbol{P}$ is an essentially infinite cover.

Thus $(5_1)$ has been established.

Now let $H \notin \boldsymbol{P}$. We may suppose that (4) is fulfilled. Therefore, if $H \subset G$, we have

$$G \cup G_1 \cup G_2 \cup \ldots \cup G_n = X$$

which yields $G \notin \boldsymbol{P}$.

We shall show that $(5_1)$ and $(5_2)$ imply that $A \cap \boldsymbol{P}$ is a cover of $X$.

Let $x_0 \in X$. Since $\boldsymbol{P}$ is a cover of $X$, there is a $G \in \boldsymbol{P}$ such that $x_0 \in G$, and since $A$ is a subbase of $X$, there is a finite system $H_1, \ldots, H_n$ of elements of $A$ such that

$$x_0 \in (H_1 \cap \ldots \cap H_n) \subset G.$$

It follows by $(5_1)$ and $(5_2)$ that there is an $i$ such that $H_i \in \boldsymbol{P}$. Hence $x_0 \in H_i \in A \cap \boldsymbol{P}$. Thus $A \cap \boldsymbol{P}$ is a cover of $X$.

Finally, since $\boldsymbol{P}$ is essentially infinite, so is $A \cap \boldsymbol{P}$.

THEOREM 3 (of Tychonov). *The product* $Z = \prod_{t\in T} X_t$ *of compact spaces* $X_t$ *is compact whatever is the set* $T$.

In particular, *the generalized cube* $\mathcal{I}^{\aleph_\alpha}$ *is compact for each* $\alpha$.

Proof. Let $A$ be an open subbase of $Z$ composed of sets (see Chapter XIII, § 5, (2)):

(6) $Q_{t,G} = \{f\colon f(t) \in G\}$ where $t \in T$ and $G \subset X_t$ is open.

Suppose that $Z$ is not compact. Then by the Alexander Lemma there is an essentially infinite cover $U \subset A$.

Denote by $V_t$ the family of sets defined by the condition

(7) $$G \in V_t \equiv Q_{t,G} \in U.$$

We shall show that $V_t$ is not a cover of $X_t$.

Suppose that the contrary is true for some $t \in T$. Since $X_t$ is compact, if follows that

(8) $X_t = G_1 \cup \ldots \cup G_n$ where $G_i \in V_t$, i.e. $Q_{t,G_i} \in U$ for $i \leqslant n$.

Therefore, by (6) and Chapter IV, (8),

$$\bigcup_i Q_{t,G_i} = \bigcup_i \{f\colon f(t) \in G_i\} = \{f\colon \bigvee_i [f(t) \in G_i]\}$$
$$= \{f\colon f(t) \in \bigcup_i G_i\} = \{f\colon f(t) \in X_t\} = Z.$$

Thus $U$ contains a finite cover of $Z$, which is impossible.

Consequently, $V_t$ is not a cover of $X_t$. This means that there is $f(t)$ in $X_t$ which belongs to no member $G$ of $V_t$. Thus

(9) $G \in V_t \Rightarrow f(t) \notin G$, i.e. $Q_{t,G} \in U \Rightarrow f(t) \notin G$

by (7).

On the other hand, since $U$ is a cover of $Z$ and $U \subset A$, there is a pair $(t, G)$ such that $f \in Q_{t,G} \in U$. But this contradicts (6) and (9) because

$$f \in Q_{t,G} \Rightarrow f(t) \in G \quad \text{while} \quad Q_{t,G} \in U \Rightarrow f(t) \notin G.$$

## § 4. Compactification of completely regular spaces

A compact space $Y$ is called a *compactification* of the space $X$ if it is compact and $X$ is homeomorphic to a dense subset of $Y$.

For example the interval $\mathcal{I}$ and the circle $\mathcal{S}_1$ are compactifications of the space $\mathcal{E}$ of reals.

It was shown in § 6 of Chapter XIII (Theorem 4) that if $X$ is a completely regular $\mathcal{T}_1$-space, the evaluation $e$ of the set $\Phi = (\mathcal{J}^X)_{\text{top}}$ ( = the set of continuous functions $f$: $X \to \mathcal{J}$) is a homeomorphism of $X$ into $Z = \mathcal{J}^\Phi$. Put

$$(1) \qquad \beta X = \overline{e(X)}.$$

Since $Z$ is compact by the Tychonov Theorem, so is $\beta X$ (by Theorem 2 of § 2). Thus *$\beta X$ is a compactification of $X$* (called the Čech–Stone compactification). It may be considered as a *maximal* compactification (see the Remark below).

FUNDAMENTAL LEMMA. *Let $f$: $X \to \mathcal{J}$ be continuous. Then the function $f \circ e^{-1}$: $e(X) \to \mathcal{J}$ has a continuous extension $f^*$: $Z \to \mathcal{J}$. Namely $f^* = \pi_f$* (compare the Remark of Chapter V, § 1).

In other words: let us identify $X$ with $e(X)$ (which is homeomorphic to $X$ by Theorem 4 of Chapter XIII, § 6); then

$$(2) \qquad f \subset f^*: Z \to \mathcal{J} \quad \text{and} \quad f^* \text{ is continuous.}$$

GENERALIZED LEMMA. *Let $T$ be an arbitrary set and let $X$ be identified with $e(X)$. If $f$: $X \to \mathcal{J}^T$ is continuous, then*

$$(3) \qquad f \subset f^*: Z \to \mathcal{J}^T, \quad \textit{where} \quad f^* \textit{ is continuous.}$$

P r o o f. Let $f_t$ denote the $t$th coordinate of $f$ and let $f_t \subset f_t^*$: $Z \to \mathcal{J}$. Then the complex mapping $f^*$ which has $f_t^*$ as its $t$th coordinate is the required mapping.

THEOREM. *Let $X$ be a completely regular $\mathcal{T}_1$-space, $Y$ a compact $\mathcal{T}_2$-space and $f$: $X \to Y$ continuous. Then, identifying $X$ with $e(X)$, we have*

$$(4) \qquad f \subset g: \beta X \to Y \quad \textit{where} \quad g \textit{ is continuous.}$$

P r o o f. Since $Y$ is a completely regular $\mathcal{T}_1$-space (by Theorem 4 of § 2), it may be assumed to be a subset of a cube $\mathcal{J}^T$ for an appropriate set $T$. Thus $f$: $X \to \mathcal{J}^T$. Applying formula (3) put $g = f^*|\beta X$. It follows that $f \subset g$. Finally $g$: $\beta X \to Y$, since the continuity of $g$ and the compactness of $Y$ imply:

$$g(\beta X) = g(\overline{X}) \subset \overline{g(X)} = \overline{f(X)} \subset \overline{Y} = Y.$$

R e m a r k. The Čech–Stone compactification is maximal in the following sense. Given a compactification $Y$ of $X$ (where $Y$

is a $\mathcal{T}_2$-space), there exists a continuous mapping of $\beta X$ into $Y$ which is the identity on $X$.

This is just another form of the preceding theorem. For, let $h$ be a topological immersion of $X$ into $Y$ and $h \subset h^*: \beta X \to Y$. By identifying $x$ with $h(x)$, one obtains the required mapping $h^*$.

## § 5. Compact metric spaces

Definition. A topological space $\mathcal{X}$ is called *countably compact*, if it satisfies the following condition (called the *Borel condition*):

*every countable open cover contains a finite subcover.*

Obviously a compact space is countably compact, while the converse is not true (as seen in the example of the space of ordinals $\alpha < \Omega$). If, however, the space is supposed to be metric, then compactness and countable compactness are equivalent (see Theorem 1).

The Borel condition is equivalent (by duality) to the following condition:

(i) if $F_1, F_2, \ldots$ is a sequence of closed sets such that

$$F_{k_1} \cap \ldots \cap F_{k_m} \neq \emptyset \tag{1}$$

for every finite system $k_1, \ldots, k_m$, then

$$\bigcap_{n=1}^{\infty} F_n \neq \emptyset. \tag{2}$$

It is also equivalent to the following condition (of *Cantor*).

(ii) *if $F_1, F_2, \ldots$ is a sequence of closed nonvoid sets such that*

$$F_1 \supset F_2 \supset \ldots \supset F_n \supset \ldots, \tag{3}$$

*then condition* (2) *is fulfilled.*

Proof. Suppose (3) is true. Then $F_{k_1} \cap \ldots \cap F_{k_m} = F_j \neq \emptyset$, where $j$ is the largest among the indices $k_1, \ldots, k_m$. Therefore, if (i) is supposed to be true, condition (2) follows. Thus (i) implies (ii).

The implication (ii) $\Rightarrow$ (i) follows directly from the formulas:

$$F_1 \cap F_2 \cap \ldots = F_1 \cap (F_1 \cap F_2) \cap \ldots \cap (F_1 \cap \ldots \cap F_n) \cap \ldots$$

and

$$F_1 \supset (F_1 \cap F_2) \supset \ldots \supset (F_1 \cap \ldots \cap F_n) \supset \ldots$$

LEMMA 1. *For metric spaces, the condition of countable compactness is equivalent to the following Bolzano–Weierstrass condition:*

(iii) *each sequence of points* $p_1, p_2, \ldots$ *contains a convergent subsequence; i.e. there is a point* $p$ *and a sequence of indices* $k_1 < k_2 < \ldots$ *such that*

$$p = \lim_{n\to\infty} p_{k_n}. \tag{4}$$

Proof. 1. Suppose that $X$ is metric and countably compact. Therefore the condition (ii) of Cantor is fulfilled. Let $p_1, p_2, \ldots$ be a sequence of points of $X$ and put $P_n = \{p_n, p_{n+1}, \ldots\}$. By (ii) there is a point $p$ such that $p \in \overline{P}_n$ for each $n$. Consequently the ball $K(p, 1/n)$ has points in common with $P_n$ and there exists an infinity of indices $m > n$ such that $|p_m - p| < 1/n$. The definition of the sequence $k_1 < k_2 < \ldots$ can easily be derived.

2. (iii) $\Rightarrow$ (ii). Choose from each $F_n$ a point $p_n$, and let (4) be fulfilled. By (3) each $F_n$ contains all points $p_1, p_2, \ldots$, except a finite number, and hence $p \in F_n$ (since $F_n$ is closed). This implies (2).

LEMMA 2. *Every countably compact metric space is separable, and hence contains a countable open base* (by Theorem 2 of Chapter XIV, § 2).

*Furthermore, for every number* $\varepsilon > 0$, *there exists a finite number of points* $A_\varepsilon = \{p_1, p_2, \ldots, p_k\}$ (called an *ε-net*) *such that*

$$\varrho(x, A_\varepsilon) < \varepsilon, \tag{4'}$$

*i.e. such that every point* $x$ *is at a distance less than* $\varepsilon$ *from some point of the set* $A_\varepsilon$.

We define the set $A_\varepsilon$ inductively. Let $p_1$ be an arbitrary point of our space. Let $p_2$ be an arbitrary point such that $|p_1 - p_2| \geqslant \varepsilon$, provided that such a point $p_2$ exists; if such a point does not exist, then we take $A_\varepsilon = \{p_1\}$.

In general, $p_n$ is a point such that

$$|p_n - p_m| \geqslant \varepsilon \quad \text{for all} \quad m < n, \tag{5}$$

provided that such a point $p_n$ exists; if such a point does not exist we take $A_\varepsilon = \{p_1, \ldots, p_{n-1}\}$.

The sequence $p_1, p_2, \ldots$ constructed in this manner must be finite; for in the contrary case, it should contain a convergent subsequence (by virtue of the compactness assumption), which however is impossible because it follows from condition (5) that no subsequence of $p_1, p_2, \ldots$ is a Cauchy sequence, and hence it cannot be convergent.

We have thus defined the set $A_\varepsilon$. It remains to show that the space is separable.

Let $B = A_1 \cup A_{1/2} \cup \ldots \cup A_{1/n} \cup \ldots$ This set is countable. It is dense in the space because for every $x$ and every $n$ we have $\varrho(x, B) \leqslant \varrho(x, A_{1/n}) < 1/n$ (by virtue of (4′) and Theorem 2 of Chapter XII, § 4); this means that there exists a point $b \in B$ such that $|x-b| < 1/n$. And therefore $x \in \overline{B}$.

THEOREM 1. *Every countably compact metric space $X$ is compact.*

Thus, *for metric spaces, compactness and countable compactness are equivalent.*

P r o o f. Let $\{G_t\}$ be an open cover of $X$. By Lemma 2, $X$ contains a countable base, and it follows (see Chapter XIV, § 1, Theorem 2) that $X$ is a Lindelöf space; consequently

$$X = \bigcup_{n=1}^{\infty} G_{t_n}.$$

Applying the Borel condition we get

$$X = G_{t_1} \cup \ldots \cup G_{t_m},$$

which completes the proof.

THEOREM 2. *Every compact metric space is complete.*

P r o o f. Let us assume that the sequence $p_1, p_2, \ldots$ is a Cauchy sequence. We shall show that it is convergent.

By assumption, for a given $\varepsilon > 0$, there exists a $j$ such that for $n > j$ we have the inequality

$$|p_n - p_j| < \varepsilon. \tag{6}$$

Since the space is compact, we can select a subsequence from the sequence $p_1, p_2, \ldots$ which satisfies condition (4).

We shall prove that

$$\lim_{n\to\infty} p_n = p. \tag{7}$$

By virtue of (4) there exists an $m > j$ such that

(8) $$|p_{k_m} - p| < \varepsilon.$$

Since $k_m \geqslant m > j$, we therefore have by (6):

(9) $$|p_{k_m} - p| < \varepsilon.$$

Adding inequalities (6), (8) and (9) memberwise, we obtain

(10) $$|p_n - p| < 3\varepsilon \quad \text{for} \quad n > j,$$

which proves (7).

THEOREM 3. *Every compact metric space is bounded.*

Put $\varepsilon = 1$ in Lemma 2. It follows that $\delta(X) < \delta(A_1)+2$.

COROLLARY 1. *For subsets of Euclidean spaces, compact sets and closed bounded sets coincide.*

Proof. If $F$ is bounded, $F$ is contained in a (sufficiently large) cube, hence in a compact space. If $F$ is supposed to be closed, $F$ is compact (by Theorem 2 of § 2).

Conversely, if $A \subset \mathscr{E}^n$ is compact, then $A$ is bounded by Theorem 3 and is closed by Theorem 1 of § 2.

COROLLARY 2 (GENERALIZED WEIERSTRASS THEOREM). *Every continuous real valued function $f$ defined on a compact space $X$ is bounded and attains its least upper and greatest lower bounds.*

Proof. The set $f(X)$ is, by virtue of Theorem 3 of § 2, a compact subset of the set of real numbers and hence (cf. Corollary 1) it is a closed and bounded set. Since the set $f(X)$ is closed, the least upper bound $m_0$ and the greatest lower bound $m_1$ of the function $f$ belong to $f(X)$. Therefore, there exist an $x_0$ such that $m_0 = f(x_0)$ and an $x_1$ such that $m_1 = f(x_1)$, which was to be proved.

We introduce the concept of *uniform continuity* in a way similar to the way it is done in analysis.

We say, namely, that the mapping $f$: $X \to Y$ where $X$ and $Y$ are metric is *uniformly continuous*, if for every $\varepsilon > 0$ there exists $\delta > 0$ (depending only on $\varepsilon$) such that the condition $|x'-x''| < \delta$ implies the inequality $|f(x')-f(x'')| < \varepsilon$ for arbitrary pairs of points $x'$, $x''$ of the space $X$; we write this condition symbolically in the form

(11) $$\bigwedge_{\varepsilon} \bigvee_{\delta} \bigwedge_{x'} \bigwedge_{x''} \{[|x'-x''| < \delta] \Rightarrow [|f(x')-f(x'')| < \varepsilon]\}.$$

Continuity in the usual sense follows from uniform continuity. The converse theorem is not true as shown by:

$$y = 1/x \ (0 < x < 1), \quad y = e^x \ (-\infty < x < +\infty).$$

On the other hand, the following theorem is valid in compact spaces:

THEOREM 4 (GENERALIZED HEINE THEOREM). *Let $f: X \to Y$ be continuous and $X$ and $Y$ metric. If $X$ is compact, $f$ is uniformly continuous.*

P r o o f. Let us assume that $f$ is not uniformly continuous. Hence there exists an $\varepsilon > 0$ such that for every $\delta > 0$ there exists a pair of points $x'$, $x''$ in the space $X$ which satisfies the conditions

$$(12) \qquad |x'-x''| < \delta \quad \text{and} \quad |f(x')-f(x'')| \geqslant \varepsilon,$$

i.e.

$$(13) \quad \bigvee_\varepsilon \bigwedge_\delta \bigvee_{x'} \bigvee_{x''} \{[|x'-x''| < \delta]\,[|f(x')-f(x'')| \geqslant \varepsilon]\}.$$

From this it follows in particular for $\delta = 1/n$ that there exists a pair of points $x_n'$, $x_n''$ such that:

$$(14) \qquad |x_n'-x_n''| < 1/n,$$

$$(15) \qquad |f(x_n')-f(x_n'')| \geqslant \varepsilon.$$

Since the space $X$ is compact, we can select a convergent subsequence $x_{k_1}', x_{k_2}', \ldots$ from the sequence $x_1', x_2', \ldots$ Let

$$(16) \qquad \lim_{n\to\infty} x_{k_n}' = x.$$

It follows from conditions (14) and (16) that

$$(17) \qquad \lim_{n\to\infty} x_{k_n}'' = x.$$

Since $f$ is continuous, we deduce from (16) and (17) that

$$\lim_{n\to\infty} f(x_{k_n}') = f(x) \quad \text{and} \quad \lim_{n\to\infty} f(x_{k_n}'') = f(x),$$

whence

$$\lim_{n\to\infty} |f(x_{k_n}')-f(x_{k_n}'')| = 0,$$

which is a contradiction of inequality (15).

Remark. Theorem 4 can be derived from the following more general statement, in which we do not assume that $X$ is metric.

* THEOREM 4'. *Let $f$: $X \to Y$ be continuous, $X$ compact and $Y$ metric. Let $\varepsilon > 0$. Then there exists an open finite cover $G_1, \ldots, G_n$ of $X$ such that*

$$\delta[f(G_i)] < \varepsilon \quad \textit{for} \quad i = 1, 2, \ldots, n. \tag{17'}$$

Proof. Let $K_y$ denote the open ball with centre $y \in Y$ and radius $\varepsilon/2$ (see Chapter IX, (10)). Obviously, the family of all $K_y$ is an open cover of $Y$, and consequently (since $f$ is continuous) the family of all $f^{-1}(K_y)$ is an open cover of $X$. Since $X$ is compact, the last cover contains a finite cover $f^{-1}(K_{y_1}), \ldots, f^{-1}(K_{y_n})$ of $X$. It remains to put $G_i = f^{-1}(K_{y_i})$ for $i = 1, \ldots, n$.

In order to derive Theorem 4 from Theorem 4', we denote by $\delta$ ($> 0$) the Lebesgue coefficient of the system of sets $X - G_1, \ldots, X - G_n$ (see Theorem 7). Then the condition $|x' - x''| < \delta$ implies that both points $x'$ and $x''$ belong to one of the sets $G_i$, and hence by (17'), $|f(x') - f(x'')| < \varepsilon$.

* THEOREM 5 (on continuous convergence). *A necessary and sufficient condition for a sequence of continuous mappings $f_1, f_2, \ldots$ defined on a compact metric space $X$ to be uniformly convergent to $f$, is that the condition*

$$\lim_{n \to \infty} x_n = x \tag{18}$$

*implies*

$$\lim_{n \to \infty} f_n(x_n) = f(x). \tag{19}$$

[We say that the sequence $f_1, f_2, \ldots$ is *continuously convergent* if condition (18) implies condition (19).]

Proof. *Necessity.* Let us assume that the sequence $f_1, f_2, \ldots$ is uniformly convergent to $f$. Let $\varepsilon > 0$. Hence, there exists a $k$ such that

$$|f_n(x) - f(x)| < \varepsilon. \tag{20}$$

for all $x$ and for $n > k$.

Let us assume (18) is satisfied. We must prove (19).

Applying (20), we have

$$(21)\qquad |f_n(x_n)-f(x_n)| < \varepsilon$$

for $n > k$.

Since the function $f$ is continuous, being the limit of a uniformly convergent sequence of continuous functions (cf. Chapter XII, § 3, Theorem 3), therefore by (18), we have

$$(22)\qquad |f(x_n)-f(x)| < \varepsilon$$

for sufficiently large $n$.

From inequalities (21) and (22) we deduce for sufficiently large $n$

$$|f_n(x_n)-f(x)| < 2\varepsilon,$$

which proves that (19) is satisfied.

*Sufficiency*. Let us assume that the sequence of continuous mappings $f_n$ is continuously but not uniformly convergent to $f$. Hence

$$\bigvee_{\varepsilon} \bigwedge_{n} \bigvee_{x} \bigvee_{k} \{(k > n)\ [|f_k(x)-f(x)| \geqslant \varepsilon]\},$$

i.e. for some $\varepsilon > 0$ and for every natural number $n$ we can choose a point $x_n$ and an index $k_n$ in such a way that

$$(23)\qquad k_1 < k_2 < \ldots < k_n < \ldots,$$

$$(24)\qquad |f_{k_n}(x_n)-f(x_n)| \geqslant \varepsilon \quad \text{for} \quad n = 1, 2, \ldots$$

The space $X$ being compact, we can choose a convergent subsequence from the sequence $x_1, x_2, \ldots$ Clearly, we can assume that the points $x_n$ are so chosen that the sequence $x_1, x_2, \ldots$ is convergent. Now, let (18) be satisfied. We shall prove that

$$(25)\qquad \lim_{n\to\infty} f_{k_n}(x_n) = f(x).$$

Let us construct the sequence $x_1', x_2' \ldots$ in the following way:

$$(26)\qquad x_m' = x_n \quad \text{for} \quad k_{n-1} < m \leqslant k_n \quad (\text{where } k_0 = 0).$$

Obviously

$$\lim_{m\to\infty} x_m' = \lim_{n\to\infty} x_n = x.$$

From this, by virtue of the continuous convergence of the sequence $f_1, f_2, \ldots$, we have

$$\lim_{m\to\infty} f_m(x_m') = f(x),$$

and hence

(27) $$\lim_{n\to\infty} f_{k_n}(x'_{k_n}) = f(x).$$

But since, by virtue of (26), $x'_{k_n} = x_n$, (27) yields (25).

Since the sequence $\{f_m\}$ is continuously convergent, we have

$$\lim_{m\to\infty} f_m(x_0) = f(x_0)$$

for fixed $x_0$. Therefore, for every $n$ we have

$$\lim_{m\to\infty} f_m(x_n) = f(x_n),$$

whence we deduce that the inequality

(28) $$|f_{m_n}(x_n) - f(x_n)| < 1/n$$

holds for some increasing sequence of indices

(29) $$m_1 < m_2 < \ldots < m_n < \ldots$$

We showed above that conditions (18) and (23) imply (25). Therefore, taking (28) into account, we have

(30) $$\lim_{n\to\infty} f_{m_n}(x_n) = f(x).$$

Formulas (25) and (30) yield

$$\lim_{n\to\infty} |f_{k_n}(x_n) - f_{m_n}(x_n)| = 0,$$

whence by virtue of (28) we have

$$\lim_{n\to\infty} |f_{k_n}(x_n) - f(x_n)| = 0;$$

but this contradicts inequality (24).

This also concludes the proof of the theorem.

THEOREM 6. *In a compact metric space $X$ the family of sets which are simultaneously closed and open is countable.*

Proof. By Lemma 2, $X$ contains a countable open base $G_1, G_2, \ldots$ Thus every open $H$ is the union of a number of sets $G_n$; if moreover $X$ is compact, this number may be assumed to be finite. Thus to every open-closed set $H$ we can assign a finite system $k_1, k_2, \ldots, k_n$ in such a way that

$$H = G_{k_1} \cup G_{k_2} \cup \ldots \cup G_{k_n}.$$

To distinct sets $H$ there obviously correspond distinct systems of natural numbers. Hence, there are at most as many open-closed sets as there are finite systems of natural numbers, and the number of the latter is countable (see Chapter V, § 3, Theorem 5).

THEOREM 7. *Let* $\{F_t\}$ *be a family of closed subsets of a compact metric space such that* $\bigcap_t F_t = \emptyset$. *Then there is* $\varepsilon > 0$ *(called the Lebesgue coefficient of the system* $\{F_t\}$*) such that every set of diameter* $< \varepsilon$ *is disjoint from at least one of the sets* $F_t$.

Proof. Note that (since the space is compact) we have $F_{t_1} \cap \cap \ldots \cap F_{t_n} = \emptyset$ for an appropriate system $t_1, \ldots, t_n$ of indices. Put

$$f(x_1, \ldots, x_n) = \max_{i,j \leqslant n} |x_i - x_j| \quad \text{where} \quad x_k \in F_{t_k},$$

and denote by $\varepsilon$ its lower bound (observe that $f\colon F_{t_1} \times \ldots \times F_{t_n} \to \mathcal{E}$ is continuous). As $F_{t_1} \cap \ldots \cap F_{t_n} = \emptyset$, we have $\varepsilon > 0$.

Suppose that $A \cap F_t \neq \emptyset$ for each $t \in T$. Put $x_{t_k} \in A \cap F_{t_k}$. Then $\delta(A) \geqslant \delta(x_1, \ldots, x_n) \geqslant \varepsilon$.

The following statement is dual to Theorem 7.

COROLLARY. *Let* $\boldsymbol{C}$ *be an open cover of the compact metric space* $X$. *Then there is* $\varepsilon > 0$ *such that every cover of* $X$ *composed of sets of diameter* $< \varepsilon$ *is a refinement of* $\boldsymbol{C}$.

Remarks. In many cases the assumption that the compact space is metric leads to simpler proofs. Such is the case of Theorems 1–3 of § 2. Here we shall give a simple proof of the invariance of compactness under finite or countable cartesian multiplication.

1. *The product* $X \times Y$ *of two compact metric spaces is compact.*

Proof. Let $z_n = \langle x_n, y_n \rangle \in X \times Y$, i.e. $x_n \in X, y_n \in Y$. We must show that the sequence $z_1, z_2, \ldots$ contains a convergent subsequence.

We can choose a convergent subsequence from the sequence $x_1, x_2, \ldots$ since the space $X$ is compact. Hence, let

$$\lim_{n \to \infty} x_{k_n} = x. \tag{31}$$

Similarly, since the space $Y$ is compact we can select a convergent subsequence from the sequence $y_{k_1}, y_{k_2}, \ldots$ Let

$$\lim_{n \to \infty} y_{r_{k_n}} = y. \tag{32}$$

By (31) we have

$$\lim_{n\to\infty} x_{r_{k_n}} = x. \tag{33}$$

Because of (32) and (33) we get

$$\lim_{n\to\infty} \langle x_{r_{k_n}}, y_{r_{k_n}} \rangle = \langle x, y \rangle, \quad \text{i.e.} \quad \lim_{n\to\infty} z_{r_{k_n}} = z.$$

We have thus selected a convergent subsequence from the sequence $z_1, z_2, \ldots$; this completes the proof.

In an analogous manner it can be proved that the cartesian product of an arbitrary finite number of compact spaces is a compact space.

2. *If the metric spaces $X_1, X_2, \ldots$ are compact, then the space $X_1 \times X_2 \times \ldots$ is also compact.*

P r o o f. Let $p_1, p_2, \ldots$ be a sequence of points belonging to the space $X_1 \times X_2 \times \ldots$ Hence

$$p_n = (x_n^1, x_n^2, \ldots, x_n^m, \ldots), \quad \text{where} \quad x_n^m \in X_m \text{ for } n, m = 1, 2, \ldots$$

Since the space $X_1$ is compact, there exists a sequence

$$1 < k_1 < k_2 < \ldots \tag{34}$$

such that the sequence $x_{k_1}^1, x_{k_2}^2, \ldots$ is convergent. Let

$$\lim_{n\to\infty} x_{k_n}^1 = x^1. \tag{35}$$

Similarly, there exists a sequence

$$1 < j_1 < j_2 < \ldots \tag{36}$$

such that the sequence $x_{k_{j_1}}^2, x_{k_{j_2}}^2, \ldots$ is convergent. Let

$$\lim_{n\to\infty} x_{k_{j_n}}^2 = x^2. \tag{37}$$

Continuing this process, we define an infinite sequence $x^1, x^2, x^3, \ldots$ Let us set

$$q = (x^1, x^2, x^3, \ldots).$$

Hence we have $q \in X_1 \times X_2 \times \ldots$ We shall prove that $q$ is the limit of the sequence

$$p_1, p_{k_1}, p_{k_{j_1}}, \ldots \tag{38}$$

In fact, making use of formulas (34) and (36) we verify that

$$1 < k_1 < k_{j_1} < k_{j_{i_1}} < \ldots,$$

and therefore the sequence (38) is a subsequence of the sequence $p_1, p_2, p_3, \ldots$

The sequence

$$x^1_{k_1}, x^1_{k_{j_1}}, \ldots$$

is therefore a subsequence of the sequence $x^1_{k_1}, x^1_{k_2}, x^1_{k_3}, \ldots$; hence, by virtue of (35) it is convergent to $x^1$. Similarly the sequence

$$x^2_{k_{j_1}}, x^2_{k_{j_{i_1}}}, \ldots$$

converges to $x^2$ by virtue of (37).

In general, the sequence

$$x^n_1, x^n_{k_1}, x^n_{k_{j_1}}, \ldots$$

converges to $x^n$.

Thus, we have proved that the sequence (38), which forms a subsequence of the sequence $p_1, p_2, \ldots$, is convergent to $q$. This means that the space $X_1 \times X_2 \times \ldots$ is compact.

## § 6. The topology of uniform convergence of $Y^X$

Let $X$ be compact and $Y$ metric. The set $Y^X$ of all continuous mappings $f\colon X \to Y$ can be considered as a metric space if the distance between its elements is defined by formula (6) of Chapter IX (the topology induced by the distance thus defined is called *topology of uniform convergence*).

We refer here to the fact that the mappings $f$ are bounded (because of the compactness of $X$, comp. § 5, Theorem 3) and hence $Y^X = \Phi(X, Y)$ (the latter symbol denoting the space of continuous bounded mappings $f\colon X \to Y$).

This identity implies the following theorem:

THEOREM 1. *If $X$ is compact and $Y$ complete, then $Y^X$ is complete.*

Because according to the Auxiliary Theorem of Chapter XV, § 4, if $Y$ is complete, so is $\Phi(X, Y)$.

R e m a r k s. In particular, the space $\mathscr{E}^{\mathscr{I}}$ is complete; this space is not compact, as is shown by the example $f_n(x) = x^n$. This same remark applies to the space $\mathscr{I}^{\mathscr{I}}$.

Theorem 1 allows us to apply the Baire theorem of Chapter XV, § 3, to function spaces (in the case where the space $X$ is compact and the space $Y$ is complete) for the purpose of proving existence theorems.

As an example of the numerous applications to analysis let us quote the following theorem:

BANACH THEOREM. *In the space* $\mathscr{E}^{\mathscr{I}}$ *the set of functions which possess a derivative in one point at least forms a boundary set.*

Banach theorem is a remarkable sharpening of Weierstrass theorem on the existence of continuous functions which do not possess a derivative at any point.

THEOREM 2. *Let $X$ be compact metric and $Y$ metric. Then the set of all one-to-one mappings $f\colon X \to Y$ (i.e. the set of homeomorphisms) is a $G_\delta$-set in the space $Y^X$.*

Proof. Let $\Gamma$ denote the set of mappings $f$ which are not one-to-one. Thus $f \in \Gamma$ means that there are $x_1 \neq x_2$ such that $f(x_1) = f(x_2)$. Let $f \in \Gamma_n$ mean that there are $x_1$ and $x_2$ such that $|x_1 - x_2| \geqslant 1/n$ and $f(x_1) = f(x_2)$.

Hence $\Gamma = \Gamma_1 \cup \Gamma_2 \cup \ldots$ Obviously the limit of a uniformly convergent sequence of members of $\Gamma_n$ belongs to $\Gamma_n$; and hence $\Gamma_n$ is closed in $Y^X$. This completes the proof.

## § 7. The compact-open topology of $Y^X_{\mathscr{A}}$

Let $X$ and $Y$ be arbitrary topological spaces. We introduce in $Y^X$ topology, called *compact-open*, in the following way.

Definition. For $C \subset X$ and $H \subset Y$, put

$$\Gamma(C, H) = \{f\colon f(C) \subset H\}, \quad \text{where} \quad f \in Y^X.$$

The compact-open topology of $Y^X$ is defined by assuming that the totality of sets $\Gamma(C, H)$, where $C$ is compact and $H$ open, is an open subbase of $Y^X$.

THEOREM 1. *If $X$ is compact and $Y$ is metric, then the uniform convergence topology of $Y^X$ coincides with its compact-open topology.*

Proof. 1. Let $G$ be open in the compact-open topology of $Y^X$. We must show that it is open in the uniform convergence topology. Without loss of generality we may assume that $G = \Gamma(C, H)$.

Let $f_0 \in \Gamma(C, H)$. We have to define $\varepsilon > 0$ such that

(1) $$|f - f_0| < \varepsilon \quad \text{implies} \quad f \in \Gamma(C, H).$$

Put

(2) $$\varepsilon = \inf \varrho[f_0(x), Y-H] \quad \text{where} \quad x \in C.$$

Since $f_0(x) \in H$ for each $x \in C$, then $\varrho[f_0(x), Y-H] > 0$ (comp. Chapter XII, § 4), and since $C$ is compact and $\varrho$ a continuous function of $x$ (ibid., Theorem 5), $\varrho$ attains its lower bound on $C$ (by the Corollary of § 5). Thus $\varepsilon > 0$.

Let $|f - f_0| < \varepsilon$. Suppose, contrary to (1), that $f \notin \Gamma(C, H)$, i.e. that $f(x_0) \in Y-H$ for some $x_0 \in C$. Hence

$$\varrho[f_0(x_0), Y-H] \leqslant |f_0(x_0) - f(x_0)| < \varepsilon.$$

But this contradicts (2).

2. We must show that each open set in the uniform convergence topology is open in the compact-open topology. Clearly, it suffices to prove that, for each $f_0 \in Y^X$ and $\varepsilon > 0$, there are two finite systems $C_1, \ldots, C_n$ and $H_1, \ldots, H_n$ (where $C_i \subset X$ is compact and $H_i \subset Y$ is open) such that

(3) $$f_0 \in \Gamma(C_1, H_1) \cap \ldots \cap \Gamma(C_n, H_n) \subset K(f_0, \varepsilon).$$

Since $X$ is compact and $f_0$ continuous, there is a finite open cover $X = G_0 \cup \ldots \cup G_n$ such that $\delta[f_0(G_i)] < \varepsilon/3$ and $G_i \neq \emptyset$ for $i = 1, \ldots, n$. Choose $x_i \in G_i$ and put

(4) $$C_i = \bar{G}_i \quad \text{and} \quad H_i = K[f_0(x_i), \varepsilon/2] = \{y: |y - f_0(x_i)| < \varepsilon/2\}.$$

Now, since $\delta[f_0(C_i)] < \varepsilon/2$, we have for $x \in C_i$ the inequality

$$|f_0(x) - f_0(x_i)| < \varepsilon/2,$$

and hence

$$f_0(x) \in H_i, \quad \text{i.e.} \quad f_0 \in \Gamma(C_i, H_i).$$

This is true for each $i = 1, \ldots, n$. Thus the proof of the first part of (3) is completed.

To prove the second part of (3), put $f \in \Gamma(C_i, H_i)$. Then for each $x \in C_i$ we have $f(x) \in H_i$, and by (4),

$$|f(x) - f_0(x_i)| < \varepsilon/2.$$

Since $\delta[f_0(C_i)] < \varepsilon/2$, it follows that

$$|f(x)-f_0(x)| \leqslant |f(x)-f_0(x_i)|+|f_0(x_i)-f_0(x)| < \varepsilon.$$

Therefore $f \in K(f_0, \varepsilon)$.

Remark. It follows from Theorem 1 that the uniform convergence topology for $X$ compact and $Y$ metric is a *topological invariant* (it does not depend upon the metric of $Y$); for $X$ compact metric this follows also from Theorem 5 of § 5.

THEOREM 2. *Let $X$ be a compact $\mathcal{T}_2$-space and $Y$ arbitrary. Given a continuous $f$: $X \to Y$, put $\varphi(f, x) = f(x)$. Then the mapping $\varphi$: $Y^X \times X \to Y$ is continuous.*

Proof. Let $H \subset Y$ be open. We have to show that the set

$$\varphi^{-1}(H) = \{\langle f, x\rangle : f(x) \in H\}$$

is open in $Y^X \times X$. In other words, that for each $f_0(x_0) \in H$ there is $Q$ open in $Y^X \times X$ such that

$$\langle f_0, x_0\rangle \in Q \tag{5}$$

and

$$Q \subset \varphi^{-1}(H), \quad \text{i.e.} \quad [\langle f, x\rangle \in Q] \Rightarrow [f(x) \in H]. \tag{6}$$

Since $f_0$ is continuous and $X$ regular (cf. § 2, Theorem 4), there is $G$ open in $X$ such that

$$x_0 \in G \quad \text{and} \quad f_0(\overline{G}) \subset H, \quad \text{i.e.} \quad f_0 \in \Gamma(\overline{G}, H).$$

Put

$$Q = \Gamma(\overline{G}, H) \times G. \tag{7}$$

It follows that $Q$ is open in $Y^X \times X$ and that formula (5) is fulfilled.

Formula (6) is also true because the condition $(f, x) \in Q$ means (by (7)) that $f(\overline{G}) \subset H$ and $x \in G$, hence $f(x) \in H$.

Remark. Instead of assuming that the compact space $X$ is a $\mathcal{T}_2$-space one could assume that $Y$ is regular. The proof would be similar.

## § 8. The Cantor discontinuum

The *Cantor discontinuum* is the set $\mathscr{C}$ of all numbers $t$ of the form

$$t = t_1/3 + t_2/9 + \ldots + t_n/3^n + \ldots, \tag{1}$$

where $t_n$ assumes one of the values 0 or 2.

They are therefore numbers of the interval [0, 1] which can be written in the ternary system of calculation without using the digit 1.

For example, 1/3 belongs to $\mathscr{C}$ because

$$1/3 = 0/3+2/9+2/27+ \dots +2/3^n+ \dots = (0.0222\dots)_3,$$

but 1/2 does not belong to $\mathscr{C}$.

We can also define the set $\mathscr{C}$ geometrically as follows.

Let us divide the closed interval [0, 1] into 3 equal parts and let us remove the middle open interval. We divide the remaining two intervals (0, 1/3) and (2/3, 1) into three equal parts and remove their (open) middle parts. Continuing in this way we obtain an infinite sequence of deleted intervals

$$(1/3, 2/3), (1/9, 2/9), (7/9, 8/9), (1/27, 2/27), \dots$$

Deleting from the interval [0, 1] the union of the removed intervals we obtain the set $\mathscr{C}$ which was defined previously arithmetically.



FIG. 7

It is therefore a *closed set* and—as is easy to see—it is dense in itself (and hence perfect), and also a *boundary set* in the interval [0, 1] (it does not contain any interval).

Next, let us note that every number of the set $\mathscr{C}$ possesses only one development of the form (1), where $t_n$ is either 0 or 2 (without this last assumption this uniqueness would not hold). It follows easily that a necessary and sufficient condition for the sequence of numbers of the Cantor set $t^{(1)}, t^{(2)}, \dots, t^{(n)}, \dots$ to converge to $t$, is that the $k$th digits in the development of these numbers converge to the $k$th digit in the development of the number $t$ (for $k = 1, 2, \dots$), i.e.

$$(2) \qquad (t = \lim_{n\to\infty} t^{(n)}) \equiv \bigwedge_k (t_k = \lim_{n\to\infty} t_k^{(n)}).$$

This means that the following theorem holds (cf. Chapter IX, § 6, Theorem 2):

THEOREM 1. *The Cantor discontinuum is homeomorphic to the infinite power of the set consisting of two elements*:

$$\mathscr{C} \underset{\text{top}}{=} \{0,2\}\times\{0,2\}\times\{0,2\}\times \ldots$$

Hence, we may identify the points of the Cantor discontinuum with sequences of zeros and twos; in other words, we identify a number belonging to $\mathscr{C}$ with the sequence of its digits in the ternary expansion (of type (1)).

We deduce from this the following theorem:

THEOREM 2. $\mathscr{C}^2 \underset{\text{top}}{=} \mathscr{C}$.

In fact, every point $p$ of the set $\mathscr{C}^2$ can be represented in the form $p = \langle x, y\rangle$ where $x$ and $y$ are sequences of zeros and twos:

$$x = (x_1, x_2, \ldots) \quad \text{and} \quad y = (y_1, y_2, \ldots).$$

From these two sequences we form one: $x_1, y_1, x_2, y_2, \ldots$ and we denote this sequence by $f(p)$.

It is easy to verify that $f$ is a homeomorphic transformation of the set $\mathscr{C}^2$ onto the set $f(\mathscr{C}^2) = \mathscr{C}$.

We could prove similarly that $\mathscr{C}^n \underset{\text{top}}{=} \mathscr{C}$ for arbitrary $n$. Moreover, the following theorem holds:

THEOREM 3. $\mathscr{C}\times\mathscr{C}\times\mathscr{C}\times \ldots \underset{\text{top}}{=} \mathscr{C}$.

The points $p$ of the set $\mathscr{C}\times\mathscr{C}\times \ldots$ are sequences of points belonging to $\mathscr{C}$:

(3) $$p = [p^{(1)}, p^{(2)}, \ldots, p^{(n)}, \ldots], \quad p^{(n)} \in \mathscr{C}.$$

In turn, $p^{(n)}$ being a point of the Cantor set, can be considered as a sequence of zeros and twos:

$$p^{(n)} = [p_1^{(n)}, p_2^{(n)}, \ldots, p_m^{(n)}, \ldots].$$

The double sequence $\{p_m^{(n)}\}$, where $n = 1, 2, \ldots$ and $m = 1, 2, \ldots$, can, by a known method (cf. Chapter V, § 3, (13) and (14)), be transformed into a single sequence

$$p_1^{(1)}, p_2^{(1)}, p_1^{(2)}, p_3^{(1)}, p_2^{(2)}, p_1^{(3)}, \ldots$$

Denoting this last sequence by $f(p)$, we obtain—as is easily proved—a homeomorphism mapping $\mathscr{C}\times\mathscr{C}\times \ldots$ onto $\mathscr{C}$.

R e m a r k. Let us consider the (closed) "non-deleted" intervals which appear in the construction of the Cantor discontinuum, i.e.

$$(0, 1/3), (2/3, 1),$$
$$(0, 1/9), (2/9, 1/3), (2/3, 7/9), (8/9, 1),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

The intersections of these intervals with the set $\mathscr{C}$ we denote successively by $P_1, P_2, P_3, \ldots$ The following theorem holds:

THEOREM 4. *The sets $P_1, P_2, \ldots$ are open-closed in the space $\mathscr{C}$ and form a base of the space. Furthermore*

$$\lim_{n\to\infty} \delta(P_n) = 0.$$

The proof that the sets $P_n$ are open-closed does not offer any difficulty. In order to prove that these sets form a base of the space $\mathscr{C}$, it suffices to note that the intervals of the first row have length 1/3, those of the second 1/9, of the $n$th $1/3^n$; furthermore the intervals of each row form a covering of the set $\mathscr{C}$.

## § 9. Continuous mappings of the Cantor discontinuum

THEOREM 1. *The interval $\mathscr{I}$ is a continuous image of $\mathscr{C}$.*

P r o o f. We define a so-called *step function* which maps the Cantor discontinuum onto the interval [0, 1]. Namely, the number $t \in \mathscr{C}$ being represented in the form (1), we set

$$\varphi(t) = \tfrac{1}{2}(t_1/2+t_2/4+ \ldots +t_n/2^n+ \ldots). \tag{4}$$

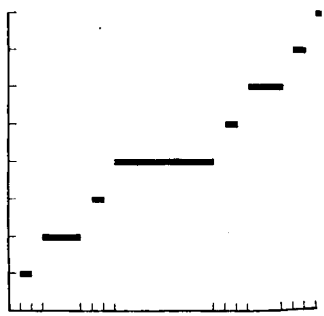


FIG. 8

It is easy to verify that the function $\varphi$ has the same value at both endpoints of each deleted interval; we take this value as a constant value of the function $f$ in this interval; otherwise, i.e. for $t \in \mathscr{C}$, we set $f(t) = \varphi(t)$. Figure 8 is the graph of this "step" function.

THEOREM 2. *The Hilbert cube $\mathscr{H}$ is a continuous image of $\mathscr{C}$.*

P r o o f. Since, by virtue of Theorem 3, § 8, the set $\mathscr{C}\times\mathscr{C}\times\times\mathscr{C}\times$ ... is a continuous image of the set $\mathscr{C}$, it suffices to prove that the space $\mathscr{H} = \mathscr{I}\times\mathscr{I}\times\mathscr{I}\times$ ... is a continuous image of the space $\mathscr{C}\times\mathscr{C}\times\mathscr{C}\times$ ... Thus, if we represent the point $p$ of this last space in the form (3) we set

$$(5) \qquad f(p) = [\varphi(p^{(1)}), \varphi(p^{(2)}), \ldots, \varphi(p^{(n)}), \ldots],$$

where $\varphi$ is the step function defined by formula (4).

The function $f$ is continuous, as can easily be seen (cf. Chapter XIII, § 5, Theorem 2). Its values are sequences of numbers belonging to the interval $[0, 1]$, i.e. they are points of the space $\mathscr{H}$. Every point $x = (x_1, x_2, \ldots, x_n, \ldots)$ of this space is a value of the function $f$, for, by virtue of Theorem 1, for every $n$ there exists a point $p^{(n)} \in \mathscr{C}$ such that $x_n = \varphi(p^{(n)})$; hence it suffices to define $p$ by formula (3) in order to obtain $x = f(p)$.

THEOREM 3. *Every compact metric space is the continuous image of some closed subset of $\mathscr{C}$.*

In fact, by virtue of the Urysohn theorem (Chapter XIV, § 4) each compact metric space $X$ can be regarded as a subset $F$ of the Hilbert cube $\mathscr{H}$. Here, $F = F$ because the space $X$ is compact (cf. Theorem 1, § 2).

Let $f$ be a function which maps $\mathscr{C}$ continuously onto the space $\mathscr{H}$. Let $A = f^{-1}(F)$.

Because of the continuity of the function $f$, the set $A$ is closed. At the same time (cf. Chapter IV, § 4, (18)): $f(A) = ff^{-1}(F) = F$.

*R e m a r k. Theorem 3 can be sharpened as follows.

THEOREM 4. *Every non-empty compact metric space is a continuous image of $\mathscr{C}$.*

Because of Theorem 3 it suffices, for this purpose, to prove the following lemma:

LEMMA. *Every non-empty closed subset $F$ of the Cantor discontinuum $\mathscr{C}$ is a continuous image of $\mathscr{C}$.*

Proof. Since the sequence $P_1, P_2, \ldots$ forms a base of the space $\mathscr{C}$ (see Theorem 4, § 8), the open set $\mathscr{C}-F$ is the union of a certain number of terms of this sequence. Hence, let

$$\mathscr{C}-F = G_1 \cup G_2 \cup \ldots, \tag{6}$$

where the sets $G_n$ belong to the sequence $P_1, P_2, \ldots$ Since we have either $P_i \cap P_j = \emptyset$ or $P_j \subset P_i$ for $i < j$, we can assume that the sets $G_n$ are disjoint (for we can omit terms in the series (6) which are contained in the earlier terms).

We denote by $p_n$ the point of the set $F$ which lies nearest the set $G_n$, i.e. the point in which the function $\varrho(x, G_n)$ defined on the set $F$ attains its greatest lower bound (cf. Chapter XII, § 4, Theorem 5 and Corollary 2 of § 5); if there is more than one such point, then we denote by $p_n$ any one of them.

We define the function $f$ as the retraction of the set $\mathscr{C}$ to $F$, namely:

$$f(x) = \begin{cases} x & \text{for} \quad x \in F, \\ p_n & \text{for} \quad x \in G_n. \end{cases}$$

Hence we have $f(\mathscr{C}) = F$. We must prove that the function $f$ is continuous.

The sets $G_n$ being open, the function $f$ is obviously continuous on their union. It remains to prove that if

$$\lim_{k\to\infty} x_k = x, \quad \text{where} \quad x_k \in \mathscr{C}-F \text{ and } x \in F, \tag{7}$$

then

$$\lim_{k\to\infty} f(x_k) = f(x), \quad \text{i.e.} \quad \lim_{k\to\infty} f(x_k) = x. \tag{8}$$

We denote by $n(k)$ an index such that

$$x_k \in G_{n(k)}. \tag{9}$$

Since to a given $G_n$ there can belong only a finite number of points of the sequence $x_1, x_2, \ldots$ (for $x \notin G_n$) and since (cf. Theorem 4, § 8) we have

$$\lim_{n\to\infty} \delta(P_n) = 0, \quad \text{and hence} \quad \lim_{n\to\infty} \delta(G_n) = 0,$$

we deduce that

$$\lim_{n\to\infty} \delta(G_{n(k)}) = 0. \tag{10}$$

Let $q_n$ denote the point of the (closed) set $G_n$ lying nearest the point $p_n$. Hence we have by virtue of the definition of the points $p_n$ and $q_n$

$$|p_n - q_n| = \varrho(p_n, G_n) \leqslant \varrho(x, G_n),$$

and therefore

$$|p_{n(k)} - q_{n(k)}| \leqslant \varrho(x, G_{n(k)}) \leqslant |x - x_k|$$

according to (9); whence

$$|p_{n(k)} - x_k| \leqslant |p_{n(k)} - q_{n(k)}| + |q_{n(k)} - x_k| \leqslant |x - x_k| + \delta(G_{n(k)}).$$

And therefore by virtue of (7) and (10) we have

$$\lim_{k\to\infty} p_{n(k)} = x. \tag{11}$$

At the same time, by virtue of the definition of the function $f$ and by the formula (9) we have

$$f(x_k) = p_{n(k)}, \tag{12}$$

and hence (8).

**Exercises**

**1.** Prove the following theorem:

If $X$ is a completely regular space, $A$ is a compact subset and $U$ is a neighbourhood of $A$, then there is a continuous function $f$: $X \to \mathscr{I}$ such that $f$ is 1 on $A$ and 0 on $X - U$.

*Hint*: For each $x$ in $A$ consider a function $g$ which is 1 at $x$ and 0 on $X - U$. Put $h(y) = \min [2g(y), 1]$ and construct a finite family $h_0, \ldots, h_n$ of continuous functions on $X$ to $\mathscr{I}$ such that $A \subset \bigcup_{i=0}^{n} \{h_i^{-1}(1)\}$ and each $h_i$ is 0 on $X - U$.

**2.** Prove the following theorem (of Wallace):

If $X$ and $Y$ are topological spaces, $A$ and $B$ are compact subsets of $X$ and $Y$ respectively, and $W$ is a neighbourhood of $A \times B$ in the product space $X \times Y$, then there are neighbourhoods $U$ of $A$ and $V$ of $B$ such that $U \times V \subset W$.

**3.** Let $X$ be the square $\mathscr{I}^2$ linearly ordered by the relation:

$$[(a, b) \prec (c, d)] \equiv [(a < c) \text{ or } (a = c \text{ and } b < d)].$$

Let the topology for $X$ be generated by the subbase $\boldsymbol{S}$ which consists of sets of the form:

$$\{x \in X: x \prec p\} \quad \text{or} \quad \{x \in X: p \prec x\} \quad \text{for some } p \text{ in } X.$$

Prove that $X$ is a compact $\mathcal{T}_2$-space and that it satisfies the first countability axiom, but is not separable.

**4.** Prove that a $\mathcal{T}_1$-space $X$ is countably compact iff the derived set of each infinite subset of $X$ is nonvoid.

**5.** Prove that if $f$: $X \to Y$ is continuous and the sequence $A_1 \supset A_2 \supset \dots$ is a decreasing sequence of countably compact subsets of the space $X$, then

$$f(\bigcap_{n=1}^{\infty} A_n) = \bigcap_{n=1}^{\infty} f(A_n).$$

**6.** Let $f$: $X \to Y$ be continuous and $X$ compact and $X$ and $Y$ $\mathcal{T}_2$-spaces. Define the equivalence relation $\varrho$ as follows

$$(x_1 \varrho x_2) \equiv (f(x_1) = f(x_2)),$$

and consider in the *quotient-space* $X/\varrho$ (cf. Exercise 9 of Chapter V) the topology defined by the condition: a set $\boldsymbol{R} \subset X/\varrho$ is open iff the union of its members is open (in $X$). Prove that $X/\varrho \underset{\text{top}}{=} Y$.

**7.** Let $f$ be a mapping defined on a $\mathcal{T}_2$-space $X$ satisfying the first axiom of countability (comp. Chapter XIV, Exercise 13). Prove that if for every compact subset $F \subset X$ the partial function $f|F$ is continuous, then $f$ is continuous (on $X$).

**8.** Let $X$, $Y$ and $T$ be metric spaces. Prove that

1°. $Y^{X \cup T} \underset{\text{top}}{=} Y^X \times Y^T$ if $X \cup T$ is compact and $X$ and $T$ are two disjoint closed sets;

2°. $(X \times Y)^T \underset{\text{top}}{=} X^T \times Y^T$ if $T$ is compact;

3°. $(Y^X)^T \underset{\text{top}}{=} Y^{X \times T}$ if $X$ and $T$ are compact.

*Hint*: Proceed as in the proofs of formulas (11)–(13), Chapter VI, § 2. In particular, using formula (17), we have

$$f \in Y^{X \times T} \equiv g \in (Y^X)^T.$$

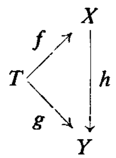
**9.** Prove that if the space under consideration in Exercise 7, of Chapter XII is compact, then the implication can be replaced by equivalence.

**10.** Let $f$: $X \to Y$ be continuous and $Y$ a compact $\mathcal{T}_2$-space. Prove that if the set $\{\langle x, y\rangle: y = f(x)\}$ is closed (in $X \times Y$), then $f$ is continuous (the converse theorem is true without the assumption of compactness of $Y$, comp. Chapter XIII, § 4, Theorem 2).

**11.** Deduce from Theorem 1 of § 3 (under the same assumption on $X$ and $Y$), that if $\varphi(x, y)$ is a propositional function of two variables such that the set $\{\langle x, y\rangle: \varphi(x, y)\}$ is closed, then so is the set $\{x: \bigvee_y \varphi(x, y)\}$; if the set $\{\langle x, y\rangle: \varphi(x, y)\}$ is open, then so is the set $\{x: \bigwedge_y \varphi(x, y)\}$.

**12.** Let $T, X$ and $Y$ be compact $\mathcal{T}_2$-spaces and $f\colon T \to X$ and $g\colon T \to Y$ continuous. Let $X = f(T)$ and suppose that $g$ is constant on the set $f^{-1}(x)$ for every $x \in X$. Let $h(x) = gf^{-1}(x)$ for $x \in X$.

Prove that the following diagram is commutative and that $h$ is continuous




*Hint*: Show that

$$(y = h(x)) \equiv \bigvee_t (y = g(t))\,(x = f(t))$$

and use Exercises 10 and 11.

**13.** Let $\varphi_1(x), \varphi_2(x), \ldots$ be a sequence of propositional functions defined on a countably compact space $X$. Show that if $\varphi_n(x) \Rightarrow \varphi_{n-1}(x)$ and if the sets $\{x\colon \varphi_n(x)\}$ are closed, then

$$\bigwedge_n \bigvee_x \varphi_n(x) = \bigvee_x \bigwedge_n \varphi_n(x).$$

Similarly: the preceding equivalence holds if $\varphi_n(x) \Rightarrow \varphi_{n+1}(x)$ and the sets $\{x\colon \varphi_n(x)\}$ are open.

**14.** A metric space is compact iff it is complete and totally bounded (comp. Chapter XII, Exercise 8).

**15.** Prove the following generalized Cantor condition: if the subsets $F_1, F_2, \ldots$ of a compact metric space are closed and non-empty, then (comp. Chapter X, Exercise 8)

$$\operatorname*{Ls}_{n\to\infty} F_n \neq \emptyset.$$

**16.** Prove that for each metric non-compact space there is a real-valued bounded continuous function whose least upper bound is not attained.

*Hint*: Use Tietze Extension Theorem.

**17.** Prove that a compact metric space cannot be isometric to a proper subset of itself.

**18.** Prove: A necessary and sufficient condition for the function $f$ defined on a metric space $X$ (compact or not) to be uniformly continuous, is that the condition

$$\lim_{n\to\infty} |x_n - x_n'| = 0$$

implies the condition

$$\lim_{n\to\infty} |f(x_n) - f(x_n')| = 0$$

for every pair of sequences $x_1, x_2, \ldots$ and $x_1', x_2', \ldots$ of points belonging to the space $X$.

**19.** Theorem 4′ of § 5 gives rise to the concept of *equicontinuity* of a family of continuous mappings. We shall restrict ourselves to the case of $X$ compact and $Y$ metric.† Then we call a set $\Phi \subset Y^X$ *equicontinuous* if for each $\varepsilon > 0$ there is an open cover $G_1, \ldots, G_n$ of $X$ such that

$$\delta\,[f(G_i)] < \varepsilon \quad \text{for each} \quad i = 1, \ldots, n \quad \text{and each} \quad f \in \Phi.$$

Show that

(i) If $\Phi_1$ and $\Phi_2$ are equicontinuous, then so is $\Phi_1 \cup \Phi_2$. In particular, each finite $\Phi$ is equicontinuous.

(ii) If $\Phi$ is countable, $\Phi = (f_1, f_2, \ldots)$, and $\lim f_n(x) = f(x)$ for each $x$, then $\Phi$ is equicontinuous iff the convergence $f_n \to f$ is uniform.

(iii) (Generalized Ascoli-Arzelà Theorem). If $\Phi$ is closed and $Y$ compact, then $\Phi$ is equicontinuous iff $\Phi$ is compact.

*Hint*: Suppose that $\Phi$ is equicontinuous and let $f_1, f_2, \ldots$ belong to $\Phi$. Choose $p_i \in G_i$ for each $i \leqslant n$ and put

$$p = \langle p_1, \ldots, p_n \rangle \in X^n \quad \text{and} \quad g_k(p) = \langle f_k(p_1), \ldots, f_k(p_n) \rangle \in Y^n.$$

Since $Y^n$ is compact metric, there exist $k_1 < k_2 < \ldots$ such that $g_{k_1}(p), g_{k_2}(p), \ldots$ is convergent. It remains to show that the sequence $f_{k_1}, f_{k_2}, \ldots$ is uniformly convergent.

Conversely, suppose that $\Phi$ is compact. Let $f_1, \ldots, f_m$ be an $\varepsilon$-net in $\Phi$ (see § 5, Lemma 2) and let, according to (i), $G_1, \ldots, G_n$ be an open cover of $X$ such that $\delta[f_j(G_i)] < \varepsilon/3$ for each $i \leqslant n$ and $j \leqslant m$.

**20.** Prove the following *Banach Fixed Point Theorem* (which holds in an arbitrary complete space):

If $f$ is a continuous mapping of the complete space $X$ into itself, and if for every pair of points $x_1, x_2 \in X$ the inequality

$$|f(x_1) - f(x_2)| \leqslant k|x_1 - x_2|$$

holds, where $k$ is a constant satisfying the condition $0 < k < 1$, then there exists exactly one point $x_0 \in X$ such that $f(x_0) = x_0$.

*Hint*: Construct inductively a sequence of points $x_1, x_2, \ldots$ in the following way: let $x_1$ be an arbitrary point of the space $X$ and let $x_n = f(x_{n-1})$. Show that a sequence constructed this way is a Cauchy sequence, and then, setting $x_0 = \lim_{n\to\infty} x_n$, prove that $f(x_0) = x_0$.

**21.** Prove using Banach theorem the following theorem on the existence of a solution of a differential equation:

Given the differential equation

$$dy/dx = f(x, y), \tag{i}$$

where the function $f$ is continuous in some plane region $G$ and satisfies in this region the Lipschitz condition with respect to $y$, i.e. there exists a constant $M$ such that the inequality

$$|f(x, y_1) - f(x, y_2)| < M|y_1 - y_2| \tag{ii}$$

† For a more general approach, see J. L. Kelley, *General Topology*, Chapter 7, p. 234; J. D. Weston, A Generalisation of Ascoli's Theorem, *Mathematika* **6** (1959), pp. 19–24; and H. Poppe, Stetige Konvergenz und der Satz von Ascoli und Arzelà, *Proceedings Japan Acad.* **44** (1968), where further references are given.

holds for every pair of points $\langle x, y_1\rangle, \langle x, y_2\rangle \in G$. Furthermore, let $\langle x_0, y_0\rangle \in G$ be a given point. Then there exists a number $\delta > 0$ such that in the interval $x_0-\delta, x_0+\delta$ there exists exactly one function $g$ satisfying equation (i), i.e.

(iii) $$dg(x)/dx = f(x, g(x)),$$

and satisfying the initial condition

(iv) $$y_0 = g(x_0).$$

*Hint*: Instead of the differential equation (i) we consider the equivalent integral equation

(v) $$y = y_0 + \int_{x_0}^{x} f(x, y)\,dx.$$

To each element $g$ of the space of continuous functions $\mathscr{E}^{\mathscr{I}}$, where $\mathscr{I}$ denotes the closed interval $x_0-\delta, x_0+\delta$, we assign the function $h_g$ of the variable $x$ defined as follows:

(vi) $$h_g(x) = y_0 + \int_{x_0}^{x} f(t, g(t))\,dt.$$

Making use of (ii) we prove that for sufficiently small $\delta > 0$ the inequality

$$|h_{g_1} - h_{g_2}| \leqslant k|g_1 - g_2|, \quad \text{where} \quad 0 < k < 1,$$

holds.

Then applying Banach theorem (Exercise 20) to the space $\mathscr{E}^{\mathscr{I}}$ we deduce that there exists exactly one function $g$ such that $h_g = g$; it is a solution of equation (v), and consequently also of equation (i), and satisfies condition (iv).

**22.** *Theorem on implicit definitions.* Let $g$ be a continuous function of two variables $x$ and $y$ with a continuous partial derivative with respect to $y$ in some square with centre $\langle x_0, y_0\rangle$; let, also,

$$g(x_0, y_0) = 0 \quad \text{and} \quad g'_y(x_0, y_0) \neq 0.$$

Then there exists one and only one function $f$, continuous in a neighbourhood of the point $x_0$, such that

$$g(x, f(x)) = 0 \quad \text{and} \quad f(x_0) = y_0;$$

in other words, the curve $\{\langle x, y\rangle : g(x, y) = 0\}$ is locally, at the point $\langle x_0, y_0\rangle$, the graph of a function.

Reduce the proof by means of the substitution

$$h(x, y) = y - y_0 - g(x, y)/g'_y(x_0, y_0)$$

to the following theorem:

Let $h$ be a function of the variables $x$ and $y$, which is continuous and has a continuous partial derivative with respect to $y$ in a square $K$ with centre $\langle x_0, y_0\rangle$ and with side $2d$; let, also,

$$h(x_0, y_0) = 0 = h'_y(x_0, y_0).$$

Then, there exists one and only one function $f$ continuous in a neighbourhood of the point $x_0$, such that

(vii) $$f(x) = h(x, f(x)) + y_0 \quad \text{and} \quad f(x_0) = y_0.$$

Sketch of the proof. We can assume that the number $d$ is so small that

$$|h'_y(x, y)| < \tfrac{1}{2} \quad \text{for} \quad \langle x, y\rangle \in K.$$

Let $I_1$ denote a closed interval with centre $x_0$ so small that

$$|h(x, y_0)| < \tfrac{1}{2}d \quad \text{for} \quad x \in I_1.$$

Let $I_2 = \{y: |y-y_0| < d\}$.

Let us assign to each function $f \in I_2^{I_1}$, satisfying the condition $f(x_0) = y_0$, the function $F_f$ of the variable $x$ defined as follows:

$$F_f(x) = y_0 + h(x, f(x)) \quad \text{for} \quad x \in I_1.$$

We obtain

$$|F_{f_1}(x) - F_{f_2}(x)| = |h(x, f_1(x)) - h(x, f_2(x))|$$
$$= |f_1(x) - f_2(x)| \cdot |h'_y(x, z_x)| \leqslant \tfrac{1}{2}|f_1(x) - f_2(x)|,$$

where $f_1(x) < z_x < f_2(x)$.

We deduce from this that

$$|F_{f_1} - F_{f_2}| \leqslant \tfrac{1}{2}|f_1 - f_2|.$$

At the same time $F_f \in I_2^{I_1}$, which we prove easily by using the inequality $|h(x, y)| \leqslant |h(x, y) - h(x, y_0)| + |h(x, y_0)|$. Finally, $F_f(x_0) = y_0$.

Hence we may apply Banach theorem. It follows that there exists a function $f$ such that $F_f = f$, i.e. satisfying conditions (vii).

**23.** Let $X$ and $Y$ be metric spaces. In the set $Y^X$ of all continuous $f$: $X \to Y$ introduce the topology as follows:

(*) for $\Phi \subset Y^X$ let $(f \in \overline{\Phi}) \equiv [(f|F) \in \overline{\Phi|F}$ for every compact $F \subset X]$,

where $\Phi|F$ denotes the set of all mappings of the form $f|F$, $f \in \Phi$, and the topology in the space $Y^F$ is defined as in § 6.

Prove that:

1°. If $X$ is a compact space, then the topology introduced by the formula (*) coincides with the topology considered in § 6.

2°. The topology introduced by (*) is compact-open.

3°. If $X$ is an open subset of a compact space, then $f$ belongs to $\overline{\Phi}$ iff there exists in $\Phi$ a sequence $f_1, f_2, \ldots$ uniformly converging to $f$ on every compact subset of $X$.

4°. The space $Y^X$ with the topology defined by (*) is completely regular.

5°. Let $T$ denote the family of all compact subsets of $X$ considered as a directed family with respect to the relation $F_0 \subset F_1$; let $\Psi$ be the function of the variable $F \in T$ defined by the condition $\Psi_F = Y^F$; let $\Gamma$ be the function of two variables $F_0$ and $F_1$ defined by the condition $\Gamma_{F_0F_1}(f) = f|F_0$ for $f \in Y^{F_1}$ and $F_0 \subset F_1$.

Prove that $\{T, \Psi, \Gamma\}$ is an inverse system (comp. Chapter VII, § 5).

6°. To every $f \in Y^X$ assign the element $\Psi(f) \in \Pi_{F \in T} Y^F$ defined by the condition:

$$\Psi_F(f) = f|F;$$

prove that $\Psi$: $Y^X \to \underleftarrow{\mathrm{Lim}}\, \{T, \Psi, \Gamma\}$, i.e. $\Psi(f) \in \underleftarrow{\mathrm{Lim}}\, \{T, \Psi, \Gamma\}$ and that $\Psi$ is a homeomorphism.

7°. Prove that the function $g$ defined by the condition $g(f, x) = f(x)$ is continuous on $Y^X \times X$. More generally: condition (*) implies the compact-open topology of $Y^X$.

**24.** Let $\{T, F, g\}$ be an inverse system, let $F_t$ be compact $\mathcal{T}_2$-spaces and $g_{t_0 t_t}$ continuous. Prove that the limit of this system is a compact space.

**25.** Let $X$ be a compact metric space and let $2^X$ denote the family of all closed subsets of $X$. Prove that if the distance in $2^X$ is defined as in Exercise 7 of Chapter XII, then $2^X$ is compact (*Hausdorff topology*).

**26.** Let $X$ be a topological space and $2^X$ be the family of all closed subsets of $X$.

Let the sets of the form $2^G \cap 2^X$ and those of the form $2^X - 2^{X-G}$ (where $G$ is open) be taken as a subbase of the topology of $2^X$ (called *Vietoris topology*).

Prove that:

(i) A base of the space $2^X$ is composed of sets of the form:

$$B(U_0, U_1, \dots, U_n) = \{A \in 2^X\colon\ A \subset U_0 \text{ and } A \cap U_i \neq \emptyset \text{ for } i = 1, \dots, n\},$$

where $U_0, \dots, U_n$ are open subsets of $X$.

(ii) If $X$ is $\mathcal{T}_1$ then so is $2^X$; $X$ is regular iff $2^X$ is a $\mathcal{T}_2$-space.

(iii) If $X$ is a compact metric space, then Vietoris topology coincides with Hausdorff topology.

**27.** Prove the following *Dini theorem.* Let $X$ be compact, $f_n \in \mathcal{E}^X$ and $f_n \leqslant f_{n+1}$ for $n = 1, 2, \dots$ If $\lim f_n(x) = f(x)$ for each $x$ and $f$ is continuous then the convergence is uniform.

*Hint*: Suppose that the convergence is not uniform. Then there is $\varepsilon > 0$ such that for each $n = 1, 2, \dots$ there are $k > n$ and $x_0$ such that $f(x_0) - f_k(x_0) \geqslant \varepsilon$. Consider the set $F_n = \{x\colon f(x) - f_n(x) \geqslant \varepsilon\}$ and apply the Cantor condition (§ 5, (ii)) to the sequence $F_1, F_2, \dots$

**28.** A space is called *locally compact* if for each point $x$ there is an open $G$ such that $x \in G$ and $\overline{G}$ is compact.

Prove the *Alexandrov One-point Compactification Theorem*: Each locally compact $\mathcal{T}_2$-space $X$ is homeomorphic to a subset $X_0$ of a compact $\mathcal{T}_2$-space $X_1$ such that $X_1 - X_0$ consists of a single point; i.e. by adjoining a single point to a locally compact space (like the point "at infinity" if $X = \mathcal{E}^2$) one obtains a compact space.

*Hint*: Let $p$ be a point not belonging to $X$. Put $X_1 = X \cup \{p\}$ and define the topology in $X_1$ by taking as members of its open base the open subsets of $X$ and the sets of the form $\{p\} \cup (X - C)$ where $C \subset X$ is compact.

CHAPTER XVII

# CONNECTED SPACES

## § 1. Definition. Separated sets

A topological space $X$ is said to be *connected* if it is not the union of two disjoint, closed and non-empty sets.

In other words, if the conditions

(1) $$X = A \cup B, \quad A = \bar{A}, \quad B = \bar{B}, \quad A \neq \emptyset \neq B$$

imply

(2) $$A \cap B \neq \emptyset.$$

EXAMPLE. *The space $\mathscr{E}$ of reals is connected.*

Suppose it is not connected. Then $\mathscr{E}$ can be decomposed in two disjoint, closed and non-empty sets $A$ and $B$. Let $a \in A$ and $b \in B$. Since $A \cap B = \emptyset$, we may suppose that $a < b$. Let $c$ be the last point of the interval $a \leqslant x \leqslant b$ which belongs to $A$ (such a point exists since $A$ is closed). Hence for each $x$ such that $c < x \leqslant b$ we have $x \notin A$ and therefore $x \in B$. Since $B$ is closed, it follows that $c \in B$. But then $c \in A \cap B$, which contradicts our hypothesis.

In a similar way one can show that *each interval (closed or open) is connected.*

THEOREM 1. *$X$ is connected iff $X$ contains no set $A$ such that*

(3) $$\emptyset \neq A \neq X$$

*and*

(4) $$\bar{A} \cap \overline{X-A} = \emptyset, \quad \textit{i.e.} \quad \mathrm{Fr}(A) = \emptyset.$$

Proof. 1. Let the set $A$ satisfy conditions (3) and (4).

The sets $A$ and $B = X-A$ are then non-empty and closed, and satisfy condition (1), but do not satisfy condition (2). Hence $X$ is not connected.

2. $X$ is not connected. Let us assume that condition (1) is satisfied but condition (2) is not, i.e.

(5) $$A \cap B = \emptyset.$$

It follows from (1) and (5) that $X-A = B$, and hence conditions (3) and (4) are fulfilled.

R e m a r k 1. It follows from Theorem 1 that a space $X$ is connected iff it contains only two closed-open subsets: $X$ and $\emptyset$.

R e m a r k 2. The condition given in Theorem 1 can be formulated in the following manner: a space is connected if for each of its decompositions into two non-empty sets $A$ and $B$ at least one of these sets contains a point which belongs to the closure of the other set (i.e. in the case of metric spaces, if there exists a point $p$ of the form $p = \lim p_n$, where $p \in A$ and $p_n \in B$, or $p \in B$ and $p_n \in A$).

This condition leads to the following formulation of the definition of a *connected set.*

A set is said to be connected if this set treated as a space forms a connected space. Therefore, *a set $C$ is connected iff for each of its decompositions into two nonvoid sets $A$ and $B$*:

$$C = A \cup B, \tag{6}$$

*we have*

$$(\bar{A} \cap B) \cup (A \cap \bar{B}) \neq \emptyset. \tag{7}$$

In other words, if two sets $A$ and $B$ are said to be *separated* provided that

$$(\bar{A} \cap B) \cup (A \cap \bar{B}) = \emptyset, \tag{8}$$

a set $C$ is connected if it cannot be decomposed into two nonvoid separated sets.

We shall prove several properties of separated sets which will be useful in the sequel.

THEOREM 2. *If the sets $A$ and $B$ are separated and $A_1 \subset A$ and $B_1 \subset B$, then the sets $A_1$ and $B_1$ are separated.*

This is true because

$$(\bar{A}_1 \cap B_1) \cup (A_1 \cap \bar{B}_1) \subset (\bar{A} \cap B) \cup (A \cap \bar{B}) = \emptyset.$$

THEOREM 3. *If the sets $A$ and $B$ are separated and the sets $A$ and $C$ are separated, then the sets $A$ and $B \cup C$ are separated.*

This follows from the formula

$$[\bar{A} \cap (B \cup C)] \cup [A \cap \overline{B \cup C}]$$
$$= (\bar{A} \cap B) \cup (\bar{A} \cap C) \cup (A \cap \bar{B}) \cup (A \cap \bar{C}) = \emptyset.$$

THEOREM 4. *If the sets A and B are both closed or both open, then the sets A—B and B—A are separated.*

P r o o f. We have

$$\overline{A-B} \cap (B-A) = \overline{A \cap (X-B)} \cap B \cap (X-A)$$
$$\subset \bar{A} \cap \overline{X-B} \cap B \cap (X-A).$$

If $\bar{A} = A$, then

$$\bar{A} \cap \overline{X-B} \cap B \cap (X-A) \subset A \cap (X-A) = \emptyset.$$

If the set $B$ is open, i.e. if the set $X-B$ is closed, then

$$\bar{A} \cap \overline{X-B} \cap B \cap (X-A) \subset (X-B) \cap B = \emptyset.$$

In an analogous way we prove that under our assumptions

$$(A-B) \cap \overline{B-A} = \emptyset,$$

and hence the sets $A-B$ and $B-A$ are separated.

## § 2. Properties of connected spaces

THEOREM 1. *The image under a continuous mapping of a connected space is a connected space; in other words, connectedness is an invariant of continuous mappings.*

P r o o f. Let $f$ be a continuous mapping of the space $X$ and let $f(X) = Y$. Let us assume that the space $Y$ is not connected. We shall then prove that the space $X$ is not connected.

Hence, let $A$ and $B$ be nonvoid closed sets such that

$$A \cup B = Y \tag{9}$$

and

$$A \cap B = \emptyset. \tag{10}$$

Then by virtue of (9) (cf. Chapter IV, § 4, (16)):

$$f^{-1}(A) \cup f^{-1}(B) = f^{-1}(Y) = X.$$

The sets $f^{-1}(A)$ and $f^{-1}(B)$ are non-empty and, since $f$ is continuous, they are also closed; making use of (10) (cf. Chapter IV, § 4, (17)), we have

$$f^{-1}(A) \cap f^{-1}(B) = f^{-1}(A \cap B) = \emptyset.$$

Thus, the space $X$ has been decomposed into two nonvoid disjoint closed sets. Hence, the space $X$ is not connected.

R e m a r k 1. The only connected subsets of the space of real numbers (other than the entire space, the void set and single points) are closed or open rays, i.e. sets of the form

$$\{x:\ x \leqslant a\}, \quad \{x:\ x < a\}, \quad \{x:\ x \geqslant a\}, \quad \{x:\ x > a\},$$

closed or open intervals, and, finally, sets of the form

$$\{x:\ a < x \leqslant b\}, \quad \{x:\ a \leqslant x < b\}.$$

For, if the set $A$ is not of one of these forms, then there exists a number $d \notin A$ and numbers $x_1, x_2 \in A$ such that $x_1 < d < x_2$. The set $A$ is then the union of two non-empty sets $M$ and $N$ contained in the separated sets

$$\{x:\ x < d\} \quad \text{and} \quad \{x:\ x > d\},$$

respectively, and hence $A$ is the union of two non-empty separated sets, i.e. it is not a connected set.

Now let $f$ be a real valued continuous function defined on the connected space $X$. The set $f(X)$ is then, by Theorem 1, a connected subset of the set of real numbers and hence it is one of the sets we indicated above.

If follows that if $y_1 \in f(X)$, $y_2 \in f(X)$ and $y_1 < y_2$, then the entire interval $y_1 \leqslant y \leqslant y_2$ is contained in the set $f(X)$, or in other words, if $y_1 \leqslant y \leqslant y_2$, then $y \in f(X)$. This means that the function $f$ has the *Darboux property*, i.e. it assumes all intermediate values in passing from one value to another. We have thus proved the following property of connected spaces:

THEOREM 2. *Every real valued continuous function defined on a connected space has the Darboux property.*

We note further that this property is characteristic of a connected space. For if a space $X$ is not connected and $A$ and $B$ are non-empty disjoint closed sets such that $A \cup B = X$, then the characteristic function of the set $A$, i.e. the function defined by the conditions

$$f(x) = \begin{cases} 1 & \text{for} \quad x \in A, \\ 0 & \text{for} \quad x \in B, \end{cases}$$

is a real valued continuous function defined on the space $X$ and not having the Darboux property.

THEOREM 3. *If $C$ is connected and $C \cap A \neq \emptyset \neq C-A$, then*

$$C \cap \mathrm{Fr}(A) \neq \emptyset.$$

In other words, *if a connected set $C$ has points in common with the set $A$ and also with its complement, then it also has points in common with the boundary of the set $A$.*

P r o o f. By virtue of the connectedness of the set $C$ and the identity $C = (C \cap A) \cup (C-A)$, the sets $C \cap A$ and $C-A$ are not separated, i.e.

$$(11) \qquad [\overline{C \cap A} \cap (C-A)] \cup [\overline{C-A} \cap C \cap A] \neq \emptyset,$$

hence

$$C \cap [(\overline{C \cap A} \cap (X-A)) \cup (\overline{C-A} \cap A)] \neq \emptyset.$$

We also have

$$\overline{C \cap A} \subset \bar{A}, \quad X-A \subset \overline{X-A}, \quad \overline{C-A} \subset \overline{X-A}, \quad A \subset \bar{A}.$$

Therefore, by (11), we have

$$\emptyset \neq C \cap \bar{A} \cap \overline{X-A} = C \cap \mathrm{Fr}(A).$$

THEOREM 4. *If the set $C$ is connected, and $C \subset M \cup N$ and the sets $M$ and $N$ are separated, then $C \subset M$ or $C \subset N$.*

P r o o f. The sets $C \cap M$ and $C \cap N$ are separated (see Theorem 2, § 1) and $(C \cap M) \cup (C \cap N) = C$. Hence, because of the connectedness of the set $C$, one of these two sets is void. If $C \cap N = \emptyset$, then $C = C \cap M$, i.e. $C \subset M$. Similarly, if $C \cap M = \emptyset$, then $C \subset N$.

THEOREM 5. *If the sets $C$ and $D$ are connected and are not separated, then their union is connected.*

P r o o f. Let $C \cup D = M \cup N$, where the sets $M$ and $N$ are separated. We have to prove that one of them is void. By Theorem 4 we can assume that $C \subset M$. Similarly, $D \subset M$ or $D \subset N$. The inclusion $D \subset N$ does not hold, because the sets $C$ and $D$ would then be separated (by Theorem 2 of § 1), contrary to assumption. Therefore $D \subset M$, whence $C \cup D \subset M$ and hence $N = \emptyset$.

Theorem 5 can be generalized as follows.

THEOREM 6. *If $\{C_t\}$ is a family of connected sets and if one of them, $C_{t_0}$, is not separated from any of the remaining sets, then the union $S = \bigcup_t C_t$ is a connected set.*

Proof. Let $S = M \cup N$, where the sets $M$ and $N$ are separated. We shall show that $M = \emptyset$ or $N = \emptyset$.

By virtue of Theorem 4, we can assume that $C_{t_0} \subset M$. Since the sets $C_{t_0}$ and $C_t$ are not separated for any $t$, we deduce from Theorem 5 that the sets $C_{t_0} \cup C_t$ are connected, and hence $C_{t_0} \cup \cup C_t \subset M$ for all $t$, whence $S \subset M$ and therefore $N = \emptyset$.

Remark 2. It follows immediately from Theorem 6 that *if $\{C_t\}$ is a family of connected sets and $\bigcap_t C_t \neq \emptyset$, then the set $\bigcup_t C_t$ is connected.*

Theorem 6 can also be derived from the following:

* THEOREM 6'. *Let $\{C_t\}$ be a directed family of connected sets, i.e. for each pair $t_1, t_2$ there is $t_3$ such that $C_{t_1} \subset C_{t_3}$ and $C_{t_2} \subset C_{t_3}$. Then the union $S = \bigcup_t C_t$ is also connected.*

Proof. Let as before $S = M \cup N$ where $M$ and $N$ are separated sets. By Theorem 4, we have for each $t$ either $C_t \subset M$ or $C_t \subset N$. Consider a $C_{t_0} \neq \emptyset$. We may assume that $C_{t_0} \subset M$. We shall show that $N = \emptyset$, which will complete the proof.

Let $t$ be arbitrary and $t'$ such that $C_{t_0} \subset C_{t'}$ and $C_t \subset C_{t'}$. The first inclusion implies that $C_{t'} \not\subset N$ (since $C_{t_0} \not\subset N$) and hence $C_{t'} \subset M$ and consequently $C_t \subset M$ (since $C_t \subset C_{t'}$). It follows that $S \subset M$ and thus $N = \emptyset$.

THEOREM 7. *If $C \subset A \subset \overline{C}$ and $C$ is connected, then so is $A$.*

In particular, *the closure of a connected set is connected.*

Proof. Let $A = M \cup N$ where $M$ and $N$ are separated. We may assume, according to Theorem 4, that $C \subset M$. Hence $\overline{C} \subset \overline{M}$ and consequently $\overline{C} \cap N = \emptyset$, whence $A \cap N = \emptyset$ (since $A \subset \overline{C}$). Finally $N = \emptyset$.

* THEOREM 8. *If $C$ is a connected subset of the connected space $X$ and*

$$X - C = M \cup N, \tag{12}$$

*where the sets $M$ and $N$ are separated, then the sets $C \cup M$ and $C \cup N$ are connected.*

*Furthermore, if the set $C$ is closed, the sets $C \cup M$ and $C \cup N$ are also closed.*

P r o o f. Let us assume that

$$C \cup M = A \cup B, \tag{13}$$

where the sets $A$ and $B$ are separated. We have to show that $A = \emptyset$ or $B = \emptyset$.

Since $C \subset A \cup B$ (by virtue of (13)), we can therefore assume, by Theorem 4, that $C \subset B$. If follows (see Theorem 2 of § 1), that the sets $A$ and $C$ are separated and in particular $A \cap C = \emptyset$. But since $A \subset C \cup M$, hence $A \subset M$, and since the sets $M$ and $N$ are separated, it follows that the sets $A$ and $N$ are separated. The set $A$ is therefore separated from $B$ as well as from $N$; it is therefore separated from $B \cup N$ (see Theorem 3 of § 1).

On the other hand, by (12) and (13) we have

$$X = C \cup M \cup N = A \cup B \cup N = A \cup (B \cup N). \tag{14}$$

The space $X$ is therefore the union of two separated sets $A$ and $B \cup N$. Since the space is connected, one of these two sets must be void. Hence, either $A = \emptyset$ or else $B \cup N = \emptyset$, whence $B = \emptyset$.

If, moreover, $C = \overline{C}$, then by (14):

$$\overline{C \cup M} = C \cup \overline{M} = C \cup [\overline{M} \cap (C \cup M \cup N)]$$
$$= C \cup M \cup (\overline{M} \cap N) = C \cup M,$$

since $\overline{M} \cap N = \emptyset$ ($M$ and $N$ being separated).

Hence the set $C \cup M$ is closed.

The same argument proves that the set $C \cup N$ is connected and closed.

* THEOREM 9. *Let $\mathbf{C}$ be a countably infinite open cover of a connected space $X$. Then $\mathbf{C}$ can be represented as an infinite sequence $G_1, G_2, \ldots$ so that $(G_1 \cup \ldots \cup G_n) \cap G_{n+1} \neq \emptyset$ for each $n$ (provided the members of $\mathbf{C}$ are nonvoid).*

The easy proof is left to the reader.

## § 3. Components

*The component of the point $p$* is the union of all connected sets which contain this point.

THEOREM 1. *Each component is a connected set.*

Moreover, *a component $S$ is a maximal connected set, i.e. if $C$ is a connected set then*

$$(S \subset C) \Rightarrow (C = S). \tag{15}$$

Proof. Let $S$ be the component of the point $p$. Therefore, $S$ is of the form

$$S = \bigcup_t C_t,$$

where $C_t$ is a connected set containing the point $p$. By virtue of Remark 2 following Theorem 6 of § 2, $S$ is a connected set.

Moreover, if $S \subset C$, then $p \in C$, and hence $C$ is of the form $C = C_t$, whence $C \subset S$. Thus $C = S$.

THEOREM 2. *Each component $S$ is a closed set.*

Proof. By Theorem 7, § 2, the set $\overline{S}$ is connected. But since $S \subset \overline{S}$, we have, making use of (15), $\overline{S} = S$.

THEOREM 3. *Two distinct components are always separated.*

Proof. If the components $S_1$ and $S_2$ are not separated, then the set $S_1 \cup S_2$ is connected (see Theorem 5 of § 2), and hence $S_1 \cup S_2 \subset S_1$ and $S_1 \cup S_2 \subset S_2$, that is $S_1 = S_2$.

EXAMPLE. Let $I_n$ denote the segment (situated in the plane) consisting of points $\langle x, y\rangle$ such that $x = 1/n$, $0 \leqslant y \leqslant 1$ for $n = 1, 2, \ldots$ Let $I_0$ denote the segment $x = 0$, $0 \leqslant y \leqslant 1$. Let $A = I_0 \cup I_1 \cup I_2 \cup \ldots$ The components of the space $A$ are segments $I_m$ ($m \geqslant 0$). Let us note that the component $I_0$ is not an open set in the space under consideration.

THEOREM 4. *If $A$ is a connected subset of a connected space $X$ and $C$ is a component of the set $X-A$, then the set $X-C$ is connected.*

Proof. Let $X-C = M \cup N$, where the sets $M$ and $N$ are separated. We shall show that $M = \emptyset$ or $N = \emptyset$.

By assumption, we have $C \subset X-A$ and hence

$$A \subset X-C = M \cup N. \tag{16}$$

We can assume (see Theorem 4 of § 2) that $A \subset M$, whence $A \cap N = \emptyset$. Since

$$A \cap (C \cup N) = (A \cap C) \cup (A \cap N) = \emptyset,$$

then $C \cup N \subset X - A$, whence

$$C \subset C \cup N \subset X - A. \tag{17}$$

Since $C$ is a component of the set $X-A$, and the set $C \cup N$ is connected (by Theorem 8 of § 2), formula (17) yields $C = C \cup N$ (cf. (15)). It follows that $N \subset C$. Since, by (16), we have $N \subset X - C$, hence $N = \emptyset$.

## § 4. Cartesian products of connected spaces

THEOREM 1. *The product $X \times Y$ of two connected spaces is connected.*

Proof. It is sufficient to show that each pair $\langle x_1, y_1\rangle$, $\langle x_2, y_2\rangle$ of points of $X \times Y$ belongs to a connected set $\subset X \times Y$. Such is in fact the set

$$(\{x_1\} \times Y) \cup (X \times \{y_2\}), \tag{1}$$

because it is the union of two connected sets which have the point $\langle x_1, y_2\rangle$ in common.

Theorem 1 can be easily extended to a finite number of connected spaces. We are going to show that it can also be extended to an arbitrary family of connected sets.

THEOREM 1'. *If $X_t$ is connected for each $t \in T$, then so is $\prod_t X_t$.*

Proof. We may, of course, assume that $X_t \neq \emptyset$ for each $t \in T$. Let $f_0$ be a fixed element of $\prod_t X_t$ (call it "the origin of the axes"; so we could choose, for example, the point $0, 0, \ldots$ in the case of the Hilbert cube). Let us assign to each finite system $\alpha = (t_1, \ldots, t_n)$ of elements of $T$ the product $C_\alpha$ of the spaces $X_t$ for $t \in \alpha$ and of the one-element sets $\{f_0(t)\}$ for $t \notin \alpha$.

$C_\alpha$ is connected because it can be obtained from the set $X_{t_1} \times \ldots \times X_{t_n}$ (which is connected by Theorem 1) by means of the continuous mapping $h$ defined as follows:

$$\pi_t[h(t_1, \ldots, t_n)] = \begin{cases} t_i & \text{for} \quad t = t_i, \\ f_0(t) & \text{for} \quad t \notin \alpha. \end{cases}$$

One shows easily that $(\alpha \subset \beta) \Rightarrow (C_\alpha \subset C_\beta)$, and it follows that the family $\{C_\alpha\}$ is directed, i.e. for each pair $\alpha_1$ and $\alpha_2$ there is $\alpha_3$ such that $C_{\alpha_1} \cup C_{\alpha_2} \subset C_{\alpha_3}$ (e.g. $\alpha_3 = \alpha_1 \cup \alpha_2$).

It follows by Theorem 6′ of § 2 that the set $S = \bigcup_\alpha C_\alpha$ is connected, and so is $\bar{S}$. It remains to be shown that $\bar{S} = \prod_t X_t$; in other words, that if $Q \subset \prod_t X_t$ is open ($\neq \emptyset$), then $S \cap Q \neq \emptyset$.

We can, of course, assume that $Q$ belongs to a base of $\prod_t X_t$. Thus we may suppose that there are a system $\alpha = (t_1, \ldots, t_n)$ and a system of sets $G_{t_i}$ open in $X_{t_i}$ such that $Q$ is the product of these sets and of the axes $X_t$ for $t \notin \alpha$, i.e.

$$Q = \prod_t G_t \qquad \text{where} \qquad G_t = X_t \text{ for } t \notin \alpha.$$

Let $f$ be a point of $Q$ such that $f(t_i) \in G_i$ for $i = 1, \ldots, n$, and $f(t) = f_0(t)$ for $t \notin \alpha$. Therefore $f \in C_\alpha$ and hence $f \in S \cap Q$.

## § 5. Continua

A *continuum* is a compact connected space.

For example, a closed interval is a continuum. Other examples of continua are a circular disk together with its boundary and the closed $n$-dimensional cube.

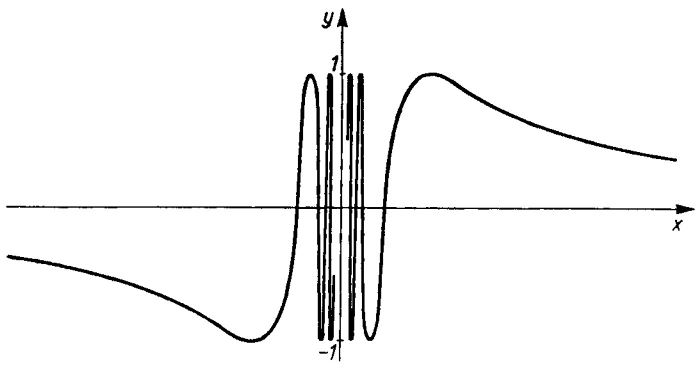



FIG. 9

The set $S$ of points in the plane defined by the following equations:

$$(1) \qquad \begin{cases} y = \sin(1/x) & \text{for} \quad 0 < |x| \leqslant 1, \\ -1 \leqslant y \leqslant 1 & \text{for} \quad x = 0 \end{cases}$$

is a continuum (see Fig. 9).

The set consisting of a single point and the void set are obviously continua; closed intervals are the only other sets which are continua on the real line.

## § 6. Properties of continua

The following five theorems are immediate consequences of the corresponding theorems in Chapters XVI and § 2 of this chapter (these are specified precisely in parentheses).

THEOREM 1. *The union of two continua which have a common point is a continuum* (cf. § 2, Theorem 5).

THEOREM 2. *If the space $X$ is a continuum, $C$ is a continuum contained in $X$, and $X-C$ is the union of two disjoint open sets $M$ and $N$, then the sets $C \cup M$ and $C \cup N$ are continua* (cf. § 2, Theorem 8).

THEOREM 3. *A continuous image of a continuum is a continuum* (Chapter XVI, § 2, Theorem 3 and Chapter XVII, § 2, Theorem 1).

In particular, *if $C$ is a non-empty continuum and $f$ is a continuous real valued function defined on $C$, then $f(C)$ is either a single point or a closed interval.*

This is a generalization of the known theorem from analysis, stating that a continuous function defined on a closed interval attains its bounds and passes through all intermediate points.

THEOREM 4. *The cartesian product $\prod_t X_t$ of continua $X_t$ is a continuum* (cf. Chapter XVII, § 4, Theorem 1 and Chapter XVI, § 3, Tychonov Theorem).

In particular, *the cube $\mathcal{I}^n$ and the Hilbert cube $\mathcal{H}$ are continua.*

THEOREM 5. *Every component of a compact space is a continuum* (cf. § 3, Theorems 1 and 2).

* THEOREM 6. *If $A$ and $B$ are two distinct components of a compact $\mathcal{T}_2$-space $X$, then $X$ can be decomposed into two disjoint closed sets $F$ and $K$ which contain the sets $A$ and $B$, respectively:*

$$(1)\quad X = F \cup K, \quad F \cap K = \emptyset, \quad A \subset F \quad \textit{and} \quad B \subset K.$$

In other words, *there exists a closed-open set $F$ which satisfies the conditions $A \subset F$ and $F \cap B = \emptyset$* (we can, of course, take $K = X-F$).

LEMMA. *The intersection $C$ of all closed-open subsets of a compact $\mathscr{T}_2$-space, which contain a given point $p$, is connected.*

In other words, *the quasi-components* (comp. Exercise 7) *of a compact $\mathscr{T}_2$-space are identical with its components.*

Proof. Let us assume the contrary. Then let $P$ and $Q$ be two closed sets such that

(2) $$C = P \cup Q,$$

(3) $$P \cap Q = \emptyset,$$

(4) $$P \neq \emptyset \neq Q,$$

(5) $$p \in P.$$

By virtue of (3) and of the normality of the space (cf. Chapter XVI, § 2, Theorem 4), there exist two open sets $G$ and $H$ such that

(6) $$P \subset G, \quad Q \subset H \quad \text{and} \quad G \cap H = \emptyset.$$

Therefore, setting $G^c = X - G$ and $H^c = X - H$, we have

(7) $$P \cap G^c = \emptyset,$$

(8) $$Q \cap H^c = \emptyset,$$

(9) $$X = G^c \cup H^c,$$

and the sets $G^c$ and $H^c$ are closed.

Let

(10) $$\{D_t\}, \quad \text{where} \quad t \in T,$$

be the family of all closed-open sets which contain the point $p$. By the definition of $C$ we have

(11) $$C = \bigcap_t D_t.$$

Let

(12) $$F_t = D_t \cap G^c \cap H^c.$$

By (11) we have

(13) $$\bigcap_t F_t = C \cap G^c \cap H^c = \emptyset,$$

because by (2) and (6), $C = (P \cup Q) \subset G \cup H$.

Since the space is compact and the sets $F_t$ are closed, it follows by (13) (comp. the Riesz condition, Chapter XVI, § 1) that there exists a finite system $t_1, \ldots, t_n$ such that

(14) $F_{t_1} \cap \ldots \cap F_{t_n} = \emptyset$, i.e. $D_{t_1} \cap \ldots \cap D_{t_n} \cap G^c \cap H^c = \emptyset$.

Since the set $D_{t_1} \cap \ldots \cap D_{t_n}$ is closed-open and contains $p$, there is $t_0 \in T$ such that

(15) $$D_{t_0} = D_{t_1} \cap \ldots \cap D_{t_n}.$$

It follows that

(16) $$D_{t_0} \cap G^c \cap H^c = \emptyset, \quad \text{i.e.} \quad D_{t_0} \cap G^c \subset H.$$

The set $D_{t_0} \cap G$ is closed-open. It is obviously open since it is the intersection of two open sets. It is also closed because, by (9) and (16), we have

(17) $$D_{t_0} \cap G = D_{t_0} \cap G \cap (G^c \cup H^c) = D_{t_0} \cap H^c,$$

and the set $D_{t_0} \cap H^c$ is the intersection of two closed sets.

As the closed-open set $D_{t_0} \cap G$ contains the point $p$ (cf. (5) and (6)), it is therefore one of the terms of the family (10): $D_{t_0} \cap G = D_{t'}$. Hence, by (11) and (17), we have

$$C \subset D_{t'} = D_{t_0} \cap G = D_{t_0} \cap H^c \subset H^c,$$

whence by (2):

$$Q \subset C \subset H^c, \quad \text{i.e.} \quad Q = Q \cap H^c = \emptyset$$

by (8). But this contradicts the inequality (4).

Proof of Theorem 6. Let $p \in A$ and let $C$ (as in the lemma) be the intersection of all closed-open sets which contain the point $p$. Each of these closed-open sets obviously contains the set $A$, since $A$ is connected (cf. § 2, Theorem 4); and therefore

(18) $$A \subset C.$$

Since $C$ is connected and $A$ is a component of the space, inclusion (18) yields

(19) $$C = A$$

(cf. § 3, (15)).

If every closed-open set containing $A$ also contained $B$ (contrary to the hypothesis of Theorem 6), then we should have $B \subset C$, whence $B \subset A$ (cf. (19)). But this is impossible, because the components are disjoint (see § 3, Theorem 3). Therefore, there exists a closed set $F$ such that $A \subset F$ and $B - F \neq \emptyset$. Since the set $B$ is connected, the last inequality yields $F \cap B = \emptyset$.

COROLLARY. *For every compact metric space there exists a continuous mapping of this space into the Cantor discontinuum $\mathscr{C}$ which maps distinct components into distinct points.*

P r o o f. Let $D_1, D_2, \ldots$ be the sequence of all closed-open subsets of the given space (comp. Chapter XVI, § 5, Theorem 6). We shall define the function $f$ as follows:

$$f(x) = t_1/3 + t_2/9 + \ldots + t_n/3^n + \ldots,$$

where $t_n = 2$ if $x \in D_n$, and $t_n = 0$ if $x \notin D_n$. (This is the *characteristic function* of the sequence $D_1, D_2, \ldots$)

Hence the values of $f$ are points of $\mathscr{C}$.

Since the set $D_n$ is closed-open a function assuming the value 2 on it and the value 0 on its complement is continuous. It easily follows that $f$ is continuous.

Finally, if $A$ and $B$ are two distinct components, then by virtue of Theorem 6 there exists an $n$ such that $A \subset D_n$ and $B \cap D_n = \emptyset$; and hence we have $t_n = 2$ for $x \in A$ and $t_n = 0$ for $x \in B$. Therefore the values of the function $f$ on the sets $A$ and $B$ are distinct.

Let us add that every component is mapped under this mapping into some point (and distinct components map onto distinct points); this follows from the fact that the continuous image of a connected set is connected, and the Cantor discontinuum does not contain nonvoid connected sets other than sets consisting of single points.

*THEOREM 7. *The intersection of a decreasing sequence of $\mathscr{T}_2$-continua is a continuum.*

P r o o f. Let $C_n$ $(n = 1, 2, \ldots)$ be $\mathscr{T}_2$-continua and let

$$C_1 \supset C_2 \supset \ldots \supset C_n \supset \ldots \tag{20}$$

and

$$C = \bigcap_{n=1}^{\infty} C_n. \tag{21}$$

Let us assume that $C$ is not a continuum. Then there exist two closed sets $P$ and $Q$ which satisfy conditions (2)–(4). Let $G$ and

$H$ be two open sets which satisfy conditions (6) and hence also conditions (7)–(9) (with $X = C_1$). Let us set

$$F_n = C_n \cap G^c \cap H^c. \tag{22}$$

Then, by (22) and (21), we have

$$\bigcap_{n=1}^{\infty} F_n = \left(\bigcap_{n=1}^{\infty} C_n\right) \cap G^c \cap H^c = C \cap G^c \cap H^c = \emptyset$$

because formulas (2) and (6) yield $C = (P \cup Q) \subset (G \cup H)$.

Since the sets $F_n$ are closed and form a decreasing sequence (because of (20)), it follows (by the Cantor condition) that not all these sets are non-empty; thus $F_n = \emptyset$ for some $n$, i.e.

$$C_n \cap G^c \cap H^c = \emptyset. \tag{23}$$

At the same time, by (9), we have

$$C_n \subset G^c \cup H^c, \quad \text{i.e.} \quad C_n = (C_n \cap G^c) \cup (C_n \cap H^c). \tag{24}$$

It follows from formulas (23) and (24) that $C_n$ is the union of two disjoint closed sets $C_n \cap G^c$ and $C_n \cap H^c$. Since $C_n$ is a continuum, one of these two sets is void. Let, for instance, $C_n \cap G^c = \emptyset$, i.e. $C_n \subset G$, and therefore, because of (2) and (21), $Q \subset C \subset C_n \subset G$, i.e. $Q \subset G$, hence by (6), $Q \subset G \cap H = \emptyset$. Thus $Q = \emptyset$, contradicting formula (4).

**Exercises**

**1.** Prove that every connected completely regular space which contains more than one point has at least the power of the continuum.

**2.** Show that the Euclidean space $\mathscr{E}^n$ $(n > 1)$ remains connected after the removal of a countable number of points.

*Hint*: Let $N$ be a countable subset of $\mathscr{E}^n$ and let $p, q \in \mathscr{E}^n - N$. Further, let $L$ be a straight line which does not pass through the points $p$ and $q$. Notice that on the line $L$ there exists a point $x$ such that the segments $px$ and $xq$ are disjoint from the set $N$.

**3.** Let the sets $A$ and $B$ be either both closed or both open. Show that if the sets $A \cup B$ and $A \cap B$ are connected, then so are the sets $A$ and $B$.

*Hint*: Make use of Theorem 8 of § 2, setting $X = A \cup B$, $C = A \cap B$, $M = A - B$, $N = B - A$, and of Theorem 4 of § 1.

**4.** Let

$$X = \bigcup_t G_t$$

be a given open cover of the connected space $X$.

Prove that every pair of points $a, b$ of the space $X$ can be joined by a chain consisting of sets $G_t$, i.e. that there exists a finite system of indices $t_1, \ldots, t_n$ such that

$$a \in G_{t_1}, \quad G_{t_1} \cap G_{t_2} \neq \emptyset, \quad \ldots, \quad G_{t_{n-1}} \cap G_{t_n} \neq \emptyset, \quad b \in G_{t_n}.$$

*Hint*: Let $Z$ be the set of all points which can be joined by a chain with the point $a$. Prove that the set $Z$ is closed-open.

**4a.** Prove the following corollary to the theorem of Exercise 4.

If the family $\{G_t\}$ is countable and $G_t \neq \emptyset$, it can be represented as an infinite sequence $G_1, G_2, \ldots$ (possibly with repetitions) so that $G_n \cap G_{n+1} \neq \emptyset$ for each $n$.

**5.** We say that the space $X$ is *connected between the sets $A$ and $B$*, if the space cannot be decomposed into two disjoint closed sets one of which contains $A$ and the other contains $B$. Prove that if there is a system of sets $A_0, \ldots, A_n$ such that the space is connected between no pair $A_i, A_j$ (for $i \neq j$), then there exists a system of disjoint closed sets $F_0, \ldots, F_n$ satisfying the conditions

$$X = F_0 \cup \ldots \cup F_n, \quad A_i \subset F_i \quad \text{for} \quad i = 0, \ldots, n.$$

**6.** Show that the relation

$p \varrho q \equiv$ (*the space $X$ is connected between the points $p$ and $q$*)

is an equivalence relation (cf. Chapter V, Exercise 9).

**7.** The equivalence sets determined by the above considered relation are called *quasi-components* of the space.

Show that

(1) every quasi-component is the intersection of all closed-open sets containing a given point;

(2) every component of the space is contained in a quasi-component but the converse is not true;

(3) if $X$ is connected between $x_1$ and $x_2$, and $Y$ is connected between $y_1$ and $y_2$, then $X \times Y$ is connected between $\langle x_1, y_1 \rangle$ and $\langle x_2, y_2 \rangle$;

(4) generalize the last statement to the case of a cartesian product of $n$ factors,

**8.** Let $A$ be a subset of a metric space. Establish the equivalence:

(*$A$ is connected between $p$ and $q$*)
$\equiv$ (*each open $G$ containing $A$ is connected between $p$ and $q$*)

*Hint*: Use the theorem stated in Chapter XII, Exercise 11.

**9.** Prove that the relation $\varrho$ defined in Exercise 6 is *closed*, i.e. the set of points $\langle x, y \rangle$ of $X \times X$ such that $x \varrho y$ is closed.

Show that the above theorem is not true for the relation "$x$ and $y$ belong to a connected subset of the space" (construct the space having the required property in the plane).

**10.** Show that a connected, metric and locally separable space is separable (cf. Chapter XIV, Exercise 14).

**11.** Show that the Corollary to Theorem 6 of § 6 remains true for separable spaces without the compactness hypothesis when components are replaced by quasi-components.

**12.** Prove that for every two points $a$ and $b$ of the metric continuum $C$ and for every $\varepsilon > 0$, there exists in $C$ a finite sequence of points

$$a = p_0, p_1, \dots, p_n = b$$

such that $|p_{i-1} - p_i| < \varepsilon$ for $i = 1, 2, \dots, n$. Show that this property is a characterization of continua among compact spaces (Cantor's definition).

**13.** Show by means of an example that in Theorem 7 of § 6 the compactness assumption is essential: the intersection of a decreasing sequence of closed connected sets does not need to be connected.

CHAPTER XVIII

# LOCALLY CONNECTED SPACES

## § 1. Definitions and examples

A topological space $X$ is said to be *locally connected at the point* $p$ if for each open set $G$ containing $p$, $p$ is an interior point of its component in $G$.

$X$ is said to be *locally connected* if it is connected at each of its points.

EXAMPLES. 1. The set of all real numbers, the Euclidean $n$-space, and the $n$-dimensional cube are locally connected spaces.

2. The set $S$ defined in Chapter XVII, § 5, (1), is not locally connected at the points of this set which are situated on the $y$-axis.

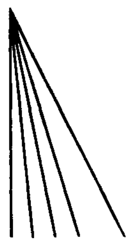


FIG. 10

3. The so-called "whisk-broom" set, shown in Fig. 10, is not locally connected.

We obtain this set by joining the point $\langle 0, 1\rangle$ with segments to the point $\langle 0, 0\rangle$ and to the points $\langle 1/n, 0\rangle$ for $n = 1, 2, \ldots$

This set is not locally connected at the points on the segment on the $y$-axis, except at the point $\langle 0, 1\rangle$.

## § 2. Properties of locally connected spaces

THEOREM 1. *In a locally connected space every component $C$ is an open set.*

For, we have, by definition, $p \in \operatorname{Int}(C)$ for every $p \in C$.

The definition implies also the following:

THEOREM 2. *A space is locally connected iff each component of an open set is open.*

THEOREM 3. *If $X$ is locally connected at $p$, then so is every open $G$ which contains $p$.*

P r o o f. Let $p \in H$ where $H$ is open relative $G$; this means that $H$ is open and $H \subset G$. Let $C$ be the component of $p$ in $H$. Since $X$ is locally connected at $p$, we have $p \in \mathrm{Int}(C)$, i.e. $p \notin \overline{X-C}$ and hence $p \in G-\overline{G-C}$, which means that $p$ is an interior point of $C$ relative $G$. Thus, $G$ is locally connected at $p$.

THEOREM 4. *$X$ is locally connected iff it has a base composed of connected open sets.*

P r o o f. 1. $X$ is locally connected. Let $\{G_t\}$ be its base. Let $\{S_{t,v}\}$ be the family of components of $G_t$. By Theorem 2 the sets $S_{t,v}$ are open and hence they form the required base.

2. $X$ has an open base composed of connected sets. Then $X$ is locally connected by definition.

R e m a r k s. A similar argument shows that, if a locally connected space $X$ has a countable base, then it has a *countable base composed of connected sets.* It follows that, in this case, every open subset of $X$ has a *countable* number of components.

COROLLARY 4′. *Every open cover* **C** *of a locally connected space $X$ contains a refinement composed of open connected sets.*

*If, moreover, the space $X$ is regular, then there is a cover of $X$ composed of open connected sets whose closures form a refinement of* **C**.

This follows from Theorem 4 by virtue of the Theorem of Chapter X, § 11 and the Remark 2 of Chapter XI, § 4.

THEOREM 5. *If $S$ is a component of an open set $G$ in a locally connected space, then*

$$\mathrm{Fr}(S) \cap G = \emptyset.$$

P r o o f. $\mathrm{Fr}(S) = \bar{S}-S$ because $S$ is open. But since $G-S$ is the union of open sets, hence $G-S$ is open and therefore $\bar{S} \cap (G-S) = S \cap (G-S) = \emptyset$, which completes the proof.

In locally connected spaces the condition of normality can be strengthened as follows.

THEOREM 6. *In a locally connected normal space $X$, let $A$ and $B$ be two closed disjoint sets. If $A$ is connected, then there is a connected open set $R$ such that $A \subset R$ and $\overline{R} \cap B = 0$.*

*If both $A$ and $B$ are connected, then there are two open connected sets $R$ and $S$ such that $A \subset R$, $B \subset S$ and $R \cap S = \emptyset$.*

P r o o f. Since $X$ is normal, there is an open $G$ such that $A \subset G$ and $\overline{G} \cap B = 0$. The component of $G$ which contains $A$ is the required open connected set $R$ (comp. Theorem 2).

If $B$ is connected, we denote by $S$ its component in $X - \overline{G}$.

THEOREM 7 (of Wilder). *Let $X$ be a metric separable and locally connected space. Let $G$ be an open connected subset of $X$. Then there exist open connected sets $R_1, R_2, \ldots$ such that*

$$(1) \qquad G = R_1 \cup R_2 \cup \ldots \quad \textit{and} \quad \overline{R}_n \subset R_{n+1}.\dagger$$

P r o o f. By Theorem 4 (comp. Remark), $X$ contains a countable base composed of open connected sets. Thus $G$ is the union of a countable family $\boldsymbol{C}$ of open connected sets whose closures are contained in $G$ (according to Remark 2 of Chapter XI, § 4). Since $G$ is connected, $\boldsymbol{C}$ can be represented (by Theorem 9 of Chapter XVII, § 2) in the form

$$(2) \qquad G = Q_1 \cup Q_2 \cup \ldots, \quad (Q_1 \cup \ldots \cup Q_n) \cap Q_{n+1} \neq \emptyset,$$
$$\overline{Q}_n \subset G, \quad Q_n \in \boldsymbol{C}.$$

We define $R_n$ by induction. Since the sets $\overline{Q}_1$ and $X - G$ are closed and disjoint, there exists by Theorem 6 an open connected set $R_1$ such that

$$(3_1) \qquad \overline{Q}_1 \subset R_1 \quad \text{and} \quad \overline{R}_1 \subset G.$$

Now, let for a given $n \geqslant 1$, $R_n$ be open connected and

$$(3_n) \qquad \overline{Q_1 \cup \ldots \cup Q_n} \subset R_n \quad \text{and} \quad \overline{R}_n \subset G.$$

Since the set $\overline{R_n \cup Q_{n+1}}$ is connected (according to the inequality (2)) and disjoint from $X - G$, therefore there is by Theorem 6 an open connected $R_{n+1}$ such that $\overline{R_n \cup Q_{n+1}} \subset R_{n+1}$ and $\overline{R}_{n+1} \subset G$.

Thus condition $(3_{n+1})$ is fulfilled and condition (1) follows.

† As shown by T. Przymusiński, "metric separable" can be replaced by "perfectly normal" (to appear in *Colloquium Math.*).

## § 3. Locally connected continua

THEOREM 1. *A continuum $X$ is locally connected iff for each open cover $\boldsymbol{C}$ there is a finite refinement composed of continua*:

$$X = C_1 \cup \ldots \cup C_n. \tag{1}$$

P r o o f. 1. *Necessity*. By Corollary 4′ of § 2, there is a cover $\boldsymbol{R}$ of $X$ composed of open connected sets whose closures form a refinement of $\boldsymbol{C}$. Since $X$ is compact we may assume that $\boldsymbol{R}$ is finite: $\boldsymbol{R} = (Q_1, \ldots, Q_n)$. We put $C_i = \overline{Q}_i$.

2. *Sufficiency*. Suppose that our condition is fulfilled. Let $p \in G$ where $G$ is open. We have to show that there exists a connected neighbourhood $E \subset G$ of $p$.

Consider the (open) cover $\boldsymbol{C}$ composed of two sets: $G$ and $H = X - \{p\}$. By assumption, (1) is fulfilled and $C_i$ is a continuum contained either in $G$ or in $H$.

Let us denote by $C_{k_1}, C_{k_2}, \ldots, C_{k_r}$ continua which contain the point $p$ and all the remaining ones by $C_{m_1}, C_{m_2}, \ldots, C_{m_s}$.

Let

$$E = C_{k_1} \cup C_{k_2} \cup \ldots \cup C_{k_r}. \tag{2}$$

We therefore have

$$X - E \subset C_{m_1} \cup C_{m_2} \cup \ldots \cup C_{m_s}, \tag{3}$$

whence

$$\overline{X - E} \subset C_{m_1} \cup C_{m_2} \cup \ldots \cup C_{m_s}. \tag{4}$$

Thus $p \in X - \overline{X - E}$, i.e. $p \in \operatorname{Int}(E)$. The set $E$ is therefore a neighbourhood of the point $p$. It is a connected set, since it is the union of connected sets which contain $p$. Finally $E \subset G$ because for $i = 1, \ldots, r$, $p \in C_{k_i}$ and hence $C_{k_i} \not\subset H$, i.e. $C_{k_i} \subset G$.

COROLLARY (Sierpiński Theorem). *A metric continuum $X$ is locally connected iff, for each $\varepsilon > 0$, $X$ can be represented as the union of a finite number of continua each of diameter less than $\varepsilon$.*

P r o o f. 1. Let $X$ be a metric locally connected continuum. Let $\boldsymbol{C}$ be a cover of $X$ composed of open sets (for instance, of balls) of diameters less than $\varepsilon$. By Theorem 1, $X$ can be represented in the form (1), where $C_1, \ldots, C_n$ are continua, each contained in a member of $\boldsymbol{C}$. Hence $\delta(C_i) < \varepsilon$ for $i = 1, \ldots, n$.

2. Now suppose that our condition is fulfilled. Let $C$ be an open cover of $X$. According to Theorem 1, we have to show that there is a finite refinement of $C$ into continua whose union is $X$. Since $X$ is compact, it is legitimate to assume that $C$ is finite, and it only remains to denote by $\varepsilon$ its coefficient (comp. Chapter XVI, § 5, Theorem 7), because formula (1) represents by virtue of the inequality $\delta(C_i) < \varepsilon$ a refinement of $C$.

THEOREM 2. *Let $X$ and $Y$ be two $\mathcal{T}_2$-spaces and $f$ a continuous mapping of $X$ onto $Y$. If $X$ is a locally connected continuum, then so is $Y$.*

P r o o f. Let $\{H_t\}$ be an open cover of $Y$. According to Theorem 1, we have to define a finite cover of $Y$ which is a refinement of $\{H_t\}$ and is composed of continua.

Now, $Y = \bigcup_t H_t$ implies $X = \bigcup_t f^{-1}(H_t)$. Since $\{f^{-1}(H_t)\}$ is an open cover of $X$, there is (by Theorem 1) a finite refinement of this cover into continua:

$$X = C_1 \cup \ldots \cup C_n, \quad \text{where} \quad C_i \subset f^{-1}(H_{t_i}). \tag{5}$$

It follows that

$$Y = f(C_1) \cup \ldots \cup f(C_n) \quad \text{and} \quad f(C_i) \subset f[f^{-1}(H_{t_i})] \subset H_{t_i}. \tag{6}$$

This completes the proof.

R e m a r k 1. A continuous image of a locally connected space which is not compact is not necessarily a locally connected space.

Let us consider the example of the space $S$ in Chapter XVII, § 5, (1) defined for $x \geqslant 0$, and let us join the point $\langle 1, \sin 1\rangle$ with the point $\langle 0, 1\rangle$ by means of an arc in such a manner that the arc does not cut the set $S$ at any point. The set thus obtained is, as can easily be seen, a continuous image of the half-ray $0 \leqslant x < +\infty$, but is not locally connected.

R e m a r k 2. From Theorem 2 it follows in particular that a continuous image of a closed segment or of a rectangle (together with boundary) is a locally connected continuum. Therefore the curves possessing continuous parametric representations on an interval of the form

$$x = x(t), \quad y = y(t), \quad z = z(t), \quad \text{where} \quad a \leqslant t \leqslant b, \tag{7}$$

are locally connected continua, as well as surfaces of the form

$$(8) \qquad x = x(u, v), \quad y = y(u, v), \quad z = z(u, v),$$

where $a \leqslant u \leqslant b$, $c \leqslant v \leqslant d$.

Thus, the geometric figures which appear most frequently in analysis are locally connected.

## § 4. Arcs. Arcwise connectedness

Definition 1. An *arc* is a set which is homeomorphic to the closed interval $0 \leqslant t \leqslant 1$.

Every arc is a locally connected continuum.

An arc with endpoints $x$ and $y$ is usually denoted by the symbol $xy$ (or $yx$).

THEOREM 1. *If $ab \cap bc = \{b\}$, then the union $ab \cup bc$ is an arc $ac$.*

For, we can define a continuous one-to-one mapping of the closed interval $[0, \frac{1}{2}]$ onto the arc $ab$ and a continuous one-to-one mapping of the closed interval $[\frac{1}{2}, 1]$ onto the arc $bc$ in such a way that both of these mappings map the point $\frac{1}{2}$ onto the point $b$. In this manner we obtain a homeomorphic mapping of the closed interval $[0, 1]$ onto the set $ab \cup bc$.

THEOREM 2. *If $ab \cap bc \neq \emptyset$, then the union $ab \cup bc$ contains an arc which connects $a$ with $c$.*

For, let $d$ be the first point on the arc $ab$ (ordered from $a$ to $b$) which lies on the arc $bc$. Let $ad$ denote the arc contained in $ab$, and let $dc$ be the arc contained in $bc$. We therefore have $ad \cap dc = \{d\}$. By Theorem 1 the set $ad \cup dc$ is an arc $ac$.

Definition 2. A space is (integrally) *arcwise connected* if every pair of its points belong to an arc. The space is said to be *locally arcwise connected,* if for every point $p$ and every open $G$ containing $p$ there is an open $H$ containing $p$ such that every point of $H$ can be joined to $p$ by an arc contained in $G$.

If the space is metric, this means that for every $p$ and every $\varepsilon > 0$, there exists an $\eta > 0$ such that if $|x-p| < \eta$, then $x$ can be connected to $p$ by an arc of diameter $< \varepsilon$.

THEOREM 3. *A space which is locally arcwise connected at the point $p$ is locally connected at $p$.*

For, if $E$ denotes the union of all arcs containing $p$ and contained in $G$, $E$ is a connected neighbourhood of $p$.

THEOREM 4. *A connected space which is locally arcwise connected is (integrally) arcwise connected.*

Proof. Let $p$ be a given point of the space $X$. Let us denote by $F$ the set of all points $x$ which can be connected to $p$ by an arc. We have to prove that $F=X$ or, equivalently (since the space is connected), that the set $F$ is closed and open.

Let $q \in X$. There exists by assumption an open $G$ containing $q$ such that for each $x \in G$ there is an arc $xq \subset X$.

Now, if $q \in \overline{F}$, there is $x \in F \cap G$ and hence an arc $px$ and (by Theorem 2) an arc $pq \subset px \cup xq$. Therefore $q \in F$. Thus $F$ is closed.

On the other hand, if $q \in F$, then each $x \in G$ can be joined to $p$ by an arc contained in $xq \cup qp$. Thus $q \in G \subset F$ and hence $q \in \mathrm{Int}(F)$. Therefore $F$ is open.

THEOREM 5. *If a compact metric space is locally arcwise connected, then for each $\varepsilon > 0$ there exists an $\eta > 0$ such that if $|x-x'| < \eta$ then the points $x$ and $x'$ can be connected by means of an arc $xx'$ of diameter $< \varepsilon$.*

Thus uniformity holds for the choice of $\eta$ corresponding to $\varepsilon$ (independently of $p$). The proof is entirely analogous to the proof of Theorem 4 in Chapter XVI, § 5.

## § 5. Continuous images of intervals

LEMMA 1 (of A. Lelek). *Consider a partition of the interval $\mathscr{I}$ into closed intervals and single points. Let $\boldsymbol{R}$ denote the family of members of this partition. Then there is a continuous function $g\colon \mathscr{I} \to \mathscr{I}$ such that the family of all inverse images $g^{-1}(t)$, $t \in \mathscr{I}$, coincides with $\boldsymbol{R}$.*

Moreover, *if $0 \in F_0 \in \boldsymbol{R}$, $1 \in F_1 \in \boldsymbol{R}$ and $F_0 \neq F_1$, then $g^{-1}(0) = F_0$ and $g^{-1}(1) = F_1$.*

Proof. Consider a countable subfamily of $\boldsymbol{R}$: $F_0, F_1, \ldots$ such that the set $F = F_0 \cup F_1 \cup \ldots$ is dense in $\mathscr{I}$. Obviously, if an interval belongs to $\boldsymbol{R}$, it is a term of the sequence $\{F_n\}$.

First we define $g(x)$ for $x \in F$. The definition is by induction. Let $g(x) = 0$ for $x \in F_0$ and $g(x) = 1$ for $x \in F_1$. Assume that, for a given $n > 1$, $g(F_i)$ is a single point for each $i < n$. Since the sets $F_0, \ldots, F_n$ are disjoint and connected, they are ordered in the interval $\mathscr{I}$ in a natural way; so denote by $F_j$ and $F_k$ the sets just preceding and just following $F_n$ (among the sets $F_0, \ldots \ldots, F_{n-1}$). Put

$$(1) \qquad g(x) = \frac{g(F_j)+g(F_k)}{2} \quad \text{for each } x \in F_n.$$

Thus $g$ is defined (by induction) for each point of $F$. In order to extend its definition to the whole interval $\mathscr{I}$, it is sufficient to show that $g$ is uniformly continuous on $F$.

Let $\varepsilon > 0$ be given. Let $1/2^{m-1} < \varepsilon$. Since $\bar{F} = \mathscr{I}$, then for each pair $F_j, F_k$ there is an $F_i$ lying between $F_j$ and $F_k$. Hence, for sufficiently large $n$, the set $g(F_0 \cup \ldots \cup F_n)$ contains all points $k/2^m, k = 0, \ldots, 2^m$. Denote by $\eta$ the length of the smallest interval contained in $\mathscr{I} - (F_0 \cup \ldots \cup F_n)$. We have

$$(2) \quad |x_1 - x_2| < \eta \Rightarrow |g(x_1) - g(x_2)| < \varepsilon \quad \text{for each pair } x_1, x_2 \in F.$$

Because the condition $|x_1 - x_2| < \eta$ implies that no pair $F_i, F_j$, where $i \leqslant n$ and $j \leqslant n$ $(i \neq j)$, lies between $x_1$ and $x_2$, and consequently no pair of points of the form $k/2^m$ can lie between $g(x_1)$ and $g(x_2)$; this means that $|g(x_1) - g(x_2)| \leqslant 2/2^m < \varepsilon$.

This completes the proof of the uniform continuity of $g$ on $F$. Hence $g$ can be considered as being defined on the whole interval $\mathscr{I}$. Obviously $g(\mathscr{I}) = \mathscr{I}$.

It remains to show that the family $\{g^{-1}(t)\}$ where $t \in \mathscr{I}$ coincides with $\boldsymbol{R}$.

Obviously $g|F$ is monotonic (not decreasing), and so is $g$. Moreover, if $x_1 < x_2$ are two points in $F$, then $g(x_1) = g(x_2)$ if and only if they belong to the same set $F_n$. This holds in general (without assuming that $x_1, x_2 \in F$). For, suppose that $x_1 \in \mathscr{I} - F$ (the case $x_2 \in \mathscr{I} - F$ is analogous). Then there are two sets $F_i \neq F_j$ between $x_1$ and $x_2$, and hence

$$(3) \qquad g(x_1) \leqslant g(F_i) < g(F_j) \leqslant g(x_2).$$

Thus, the single-element set $\{x_1\}$, which is a member of $\boldsymbol{R}$, is an inverse image of an element of $\mathscr{I}$ (namely of $g(x_1)$); since the same is true for all the sets $F_0, F_1, \ldots$, this completes the proof of the Lemma.

THEOREM 1 (of G. T. Whyburn). *Let* $f\colon \mathcal{J} \to Y$ *be continuous, let* $f(0) = a$, $f(1) = b$ *and* $a \neq b$. *Suppose that the inverse image* $f^{-1}(y)$ *is, for each* $y \in Y$, *either a closed interval or a single point of* $\mathcal{J}$. *Then* $Y$ *(which is supposed to be* $\mathcal{T}_2$*) is an arc* $ab$.

P r o o f. Denote by $\boldsymbol{R}$ the family of all inverse images $f^{-1}(y)$ where $y \in Y$ and consider the function $g$ satisfying the preceding Lemma (we denote: $F_0 = f^{-1}(a)$ and $F_1 = f^{-1}(b)$). Since for each $t \in \mathcal{J}$, $g^{-1}(t) \in \boldsymbol{R}$, then $f[g^{-1}(t)]$ is a single point of $Y$. Put

$$h(t) = f[g^{-1}(t)] \quad \text{for} \quad t \in \mathcal{J}. \tag{4}$$

$h$ is the required homeomorphism of $\mathcal{J}$ onto $Y$.

First, $h$ is continuous. For let $F \subset Y$ be closed. Then $h^{-1}(F) = g[f^{-1}(F)]$ and since $f^{-1}(F)$ is closed, so also is $g[f^{-1}(F)]$ (because $g$ is a continuous mapping of a compact space).

Next, $h$ is one-to-one. For let $t_1 \neq t_2$. Then $g^{-1}(t_1) \neq g^{-1}(t_2)$ and hence there are $y_1 \neq y_2$ such that $g^{-1}(t_1) = f^{-1}(y_1)$ and $g^{-1}(t_2) = f^{-1}(y_2)$. Therefore $h(t_1) = y_1 \neq y_2 = h(t_2)$.

Finally $h(0) = a$ and $h(1) = b$. For

$$h(0) = f[g^{-1}(0)] = f(F_0) = f[f^{-1}(a)] = a.$$

Similarly $h(1) = b$.

LEMMA 2. *Let* $A \subset \mathcal{J}$ *be a closed subset containing the points* 0 *and* 1, *and let* $f\colon A \to Y$ *(which is* $\mathcal{T}_2$*) be continuous and onto,* $f(0) = a$, $f(1) = b$ *and* $a \neq b$. *Suppose that* $A$ *has irreducibly the following property*:

$$\begin{cases} \textit{the points } 0 \textit{ and } 1 \textit{ belong to } A, \\ \textit{if } uv \textit{ is a component of } \mathcal{J} - A, \textit{ then } f(u) = f(v) \end{cases} \tag{5}$$

*(the irreducibility means that no proper closed subset of* $A$ *containing* 0 *and* 1 *has property* (5)*).*

*Then* $Y$ *is an arc* $ab$.

P r o o f.† Let us extend $f$ to a mapping $g\colon \mathcal{J} \to Y$ so that, for each component $uv$ of $\mathcal{J} - A$, $g(t) = f(u)$ for $u \leqslant t \leqslant v$. Then $g$ is obviously continuous and $g(\mathcal{J}) = Y$. Moreover each inverse

† See an argument of J. L. Kelley mentioned by G. T. Whyburn, *loc. cit.*

image $g^{-1}(y)$ for $y \in Y$ is either a closed interval or a single point. For, suppose that $g^{-1}(y)$ is not a single point, and let $t_1$ and $t_2$ be its first and last elements. Obviously $t_1$ and $t_2$ belong to $A$ and the open interval $t_1t_2$ is disjoint from $A$. Because, otherwise, the set $A - t_1t_2$ would have property (5).

Thus $g^{-1}(y) = \overline{t_1t_2}$ and it follows by Theorem 1 that $Y$ is an arc $ab$.

THEOREM 2. *A continuous image of an interval is arcwise connected.*

More precisely, *if $f: \mathscr{I} \to Y$ is continuous, $Y$ is $\mathscr{T}_2$ and $f(0) = a$, $f(1) = b$ and $a \neq b$, then there is a closed subset $A$ of $\mathscr{I}$, containing the points* 0 *and* 1 *and having irreducibly the property* (5).

P r o o f. According to the Brouwer reduction theorem (Chapter XIV, § 1, Theorem 4) we have to show that the property (5) is inducible; i.e. that if $A = A_0 \cap A_1 \cap \ldots$ where each $A_n$ has property (5) and $A_0 \supset A_1 \supset \ldots$, then $A$ also has property (5). Let $uv$ be a component of $\mathscr{I} - A$. Since $\mathscr{I} - A = \bigcup_n (\mathscr{I} - A_n)$, so $uv$ is the union of components $u_nv_n$ of $\mathscr{I} - A_n$, and since $A_n$ has property (5), so $f(u_n) = f(v_n)$ and consequently $f(u) = f(v)$.

COROLLARY. *A metric continuous image of an interval is locally arcwise connected.*

For let $f: \mathscr{I} \to Y$ be continuous and onto. Since $f$ is uniformly connected, there are intervals $I_1, \ldots, I_n$ such that $\mathscr{I} = I_1 \cup \ldots \cup I_n$ and $f(I_k) < \varepsilon$ for $k = 1, \ldots, n$. This clearly implies the required conclusion.

THEOREM 3. *Every locally arcwise connected metric continuum $C$ ($\neq \emptyset$) is a continuous image of an interval.*

P r o o f. By Theorem 3 of Chapter XVI, § 8, there exists a continuous mapping $f$ defined on some closed subset $H$ of the Cantor set and such that $f(H) = C$. Let $\alpha$ and $\beta$ denote the initial and terminal points of the set $H$. We shall extend $f$ to the entire segment $\alpha\beta$. The set $\alpha\beta - H$, being open in $\alpha\beta$, is the union of a sequence of open intervals $(a_1b_1), (a_2b_2), \ldots$

Obviously,

$$\lim_{n \to \infty} (b_n - a_n) = 0,$$

whence

$$\lim_{n\to\infty} |f(b_n)-f(a_n)| = 0, \tag{6}$$

because of the uniform continuity of $f$.

According to Theorem 5 of § 4, there exists a sequence of numbers $\eta_k$, such that each two points $p$ and $q$ of the continuum $C$ satisfying the inequality $|p-q| < \eta_k$ can be joined by an arc with diameter $< 1/k$. Therefore by virtue of (6) there exists a sequence of arcs $L_n$ with endpoints $f(a_n)$ and $f(b_n)$ such that

$$\lim_{n\to\infty} \delta(L_n) = 0. \tag{7}$$

Let $f_n$ denote a homeomorphism of the (closed) segment $a_n b_n$ onto the arc $L_n$, such that $f_n(a_n) = f(a_n)$ and $f_n(b_n) = f(b_n)$. Finally, let

$$g(t) = \begin{cases} f(t) & \text{for} \quad t \in H, \\ f_n(t) & \text{for} \quad a_n \leqslant t \leqslant b_n,\ n = 1, 2, \ldots \end{cases}$$

Hence $g$ maps the segment $\alpha\beta$ onto the continuum $C$. It is a continuous function, as seen from formula (7).

R e m a r k 1. The conditions of being:

(i) a locally connected continuum,

(ii) a locally arcwise connected continuum,

(iii) a continuous image of an interval,

are equivalent (for metric spaces).

Because every locally connected metric continuum is also locally arcwise connected (this is the Mazurkiewicz–Moore theorem which we state here without proof).

R e m a r k 2. It follows from Theorem 3 in particular that *$\mathscr{I}^2$ is a continuous image of a segment*; the same is true of the $n$-dimensional cube $\mathscr{I}^n$, and even of the Hilbert cube $\mathscr{H}$.

This discovery made by Peano (in 1890) was considered to be very paradoxical. For it means that the square $\mathscr{I}^2$ has a continuous parametric representation over a closed interval, contrary to the opinion that this property applies only to curves. It follows from this that the hypothesis of d i f f e r e n t i a b i l i t y usually made in analysis for parametric representations is essential from this point of view.

The following is a direct proof of the Peano theorem (given by Sierpiński).

We divide the square into 9 equal squares and draw in each of them the diagonal as shown in Fig. 11. We divide the segment

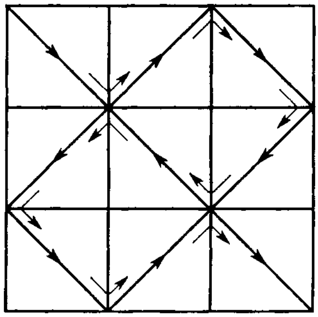


FIG. 11

[0, 1] into 9 equal segments and we transform (linearly) each of them into the corresponding diagonal in the order given in Fig. 11. We denote by $f_1$ the function thus defined, mapping the segment [0, 1] continuously into the polygonal line consisting

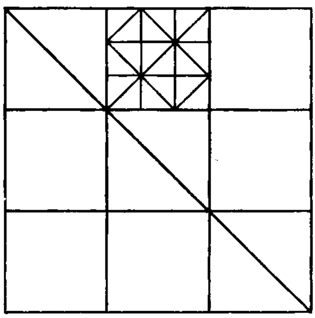


FIG. 12

of 9 diagonals. We call the squares considered squares of first approximation.

Next, we divide each of the 9 squares into 9 equal squares; they are the second approximation squares. We draw a diagonal $D$ in each of them; here in second approximation squares lying on a diagonal of a first approximation square we draw the diagonal lying on the diagonal $D$. Thus the first square of the first approximation appears as in Fig. 11 after the corresponding reduction; the second square of the first approximation is given in Fig. 12.

We divide each of the intervals $(n-1)/9, n/9$, where $n = 1, 2, \ldots, 9$, onto 9 equal parts and we map each of these parts into the diagonal of the corresponding square of the second subdivision. This defines the function $f_2$ which maps the interval $[0, 1]$ continuously onto the polygonal arc made up of $9^2$ intervals.

Continuing thus, we define an infinite sequence of continuous functions $f_1, f_2, \ldots, f_n, \ldots$ It is easy to prove that this sequence is uniformly convergent; and therefore its limit function $f$ is continuous (see Chapter XII, § 3, Theorem 3). Furthermore, every point of the square is a value of the function $f$; in fact, in each square of the $n$th approximation there are values of the function $f_n$ and consequently

$$\overline{\bigcup_n f_n(\mathscr{I})} = \mathscr{I}^2 \quad \text{whence} \quad f(\mathscr{I}) = \mathscr{I}^2.$$

R e m a r k 3. Let us notice that the proof of Theorem 3 in the case where $C = \mathscr{I}^n$ can be somewhat simplified. Namely, in this case we can take the interval with endpoints $f(a_n)$ and $f(b_n)$ for the arc $L_n$; hence, we can define the function $f_n$ as the linear transformation of the interval $a_n b_n$ into the interval $f(a_n)f(b_n)$.

This theorem can also be deduced directly from Theorem 3, Chapter XVI, § 8, and Tietze theorem (Chapter XII, § 5, Corollary 1).

**Exercises**

**1.** Let $E$ be an open subset of the interval $a < x < b$. Prove that the components of the set $E$ are open intervals. Moreover, if there are an infinite number of these components then their diameters tend to zero.

**2.** Let $p \in A \cap B$. If the sets $A$ and $B$ are locally connected at the point $p$, then the set $A \cup B$ is also locally connected at this point.

**3.** If the spaces $X$ and $Y$ are locally connected at the points $a$ and $b$ respectively, then the cartesian product $X \times Y$ is locally connected at the point $\langle a, b \rangle$.

**4.** Let $E$ be an arbitrary subset of a locally connected space. If $C$ is a connected subset of $E$ and is open in $E$, then there exists an open connected set $H$ such that $C = E \cap H$.

*Hint*: Use Theorem 2, § 2.

**5.** If a locally connected space can be represented as the union of two closed sets $A$ and $B$ with locally connected intersection, then the sets $A$ and $B$ are locally connected.

*Hint*: Use Exercises 2 and 4, above, and Exercise 3 of Chapter XVII.

**6.** Let $X$ be a locally connected space. If $F$ is a closed locally connected set and $C$ is a component of the set $X-F$, then the sets $X-C$ and $C \cup F$ are locally connected.

*Hint*: Use Exercise 5.

**7.** Let $E$ be an arbitrary subset of a locally connected space and let $E = S_1 \cup S_2 \cup \ldots$ be the decomposition of $E$ into components. Then

$$\mathrm{Int}\,(E) = \bigcup_n \mathrm{Int}\,(S_n).$$

**8.** Let $E_t$ be an arbitrary subset of a locally connected space. Prove that

$$\mathrm{Fr}(\bigcup_t E_t) \subset \overline{\bigcup_t \mathrm{Fr}(E_t)}.$$

*Hint*: Use Theorem 3 of Chapter XVII, § 2.

**9.** Let $E$ be an arbitrary subset of a locally connected space and let $S$ be a component of $E$. Prove that $\mathrm{Fr}(S) \subset \mathrm{Fr}(E)$.

**10.** Let $E$ be an arbitrary subset of a locally connected space. If the set $\mathrm{Fr}(E)$ is locally connected, then $\overline{E}$ is locally connected.

*Hint*: Use Exercise 5.

**11.** Let $G$ be an open connected subset of a locally connected regular space. Then every pair of points of $G$ can be joined by an open connected set $R$ such that $\overline{R} \subset G$.

**12.** Let $X$ be a locally connected metric continuum. Prove that each of its subcontinua $C$ is the intersection of a decreasing sequence of locally connected continua:

$$C = \bigcap_{n=1}^{\infty} C_n, \quad C_1 \supset C_2 \supset \ldots$$

CHAPTER XIX

# THE CONCEPT OF DIMENSION

The space under consideration in this chapter is metric separable (however, many definitions and theorems can be applied to more general spaces†).

## § 1. 0-dimensional spaces

Definition. A nonvoid space $X$ is said to be *0-dimensional*, i.e. $\dim X = 0$, iff for every finite open cover

(1) $$X = G_0 \cup \ldots \cup G_m$$

there is a refinement in closed disjoint sets:

(2) $$X = F_0 \cup \ldots \cup F_m,$$

(3) $$F_i \subset G_i \quad \text{for} \quad i = 0, \ldots, m,$$

($4_0$) $$F_i \cap F_j = \emptyset \quad \text{for} \quad i \neq j$$

(thus the sets $F_i$ are closed-open).

Furthermore, we agree that $\dim \emptyset = -1$.

EXAMPLES. The space of integers, the space of irrational numbers, the Cantor discontinuum are 0-dimensional.

An interval, as well as any connected space (which does not reduce to a single point), is not 0-dimensional; for it does not contain non-empty closed-open sets which are distinct from the entire space.

## § 2. Properties of 0-dimensional metric separable spaces

We state here without proof the most important properties of 0-dimensional metric separable spaces. We could already have observed some of these properties in the Cantor set $\mathscr{C}$.

THEOREM 1. *Every 0-dimensional space has a countable base consisting of open-closed sets.*

† See J. Nagata, *Modern Dimension Theory*, Noordhoff, 1965.

THEOREM 2. *Every* 0-*dimensional space is topologically contained in* $\mathscr{C}$ *(i.e. it is homeomorphic to some subset of* $\mathscr{C}$*).*

THEOREM 3. *Every* 0-*dimensional compact space can be decomposed into disjoint closed sets of diameter* $< \varepsilon$ *for each* $\varepsilon > 0$.

THEOREM 4 (Sharpened normality property). *For every pair of disjoint closed sets* $A$ *and* $B$, *there exists an closed-open set* $G$ *such that* $A \subset G$ *and* $G \cap B = \emptyset$.

THEOREM 5. *In a metric separable space the union of a finite or infinite sequence of* 0-*dimensional closed sets is a* 0-*dimensional set.*

## § 3. *n*-dimensional spaces

Definition. $\dim X \leqslant n$ iff for every finite open cover (see (1)) there is a refinement in closed sets satisfying (besides (2) and (3)) the following condition

$$(4_n) \quad F_{i_0} \cap F_{i_1} \cap \ldots \cap F_{i_n} = \emptyset \quad \text{for} \quad i_0 < i_1 < \ldots < i_{n+1},$$

which means that no point of $X$ belongs to $n+2$ sets $F_i$.

The definition of 0-dimensional spaces given in § 1 is obviously a particular case of the above definition.

THEOREM. *Let* $X$ *be a compact metric space.* $X$ *is of dimension* $\leqslant n$ *iff there is for each* $\varepsilon > 0$ *a closed cover satisfying conditions* (2), $(4_n)$ *and*

$$(5) \qquad \delta(F_i) < \varepsilon \quad \textit{for} \quad i = 0, 1, \ldots, m.$$

Proof. 1. *Necessity.* Let $\dim X \leqslant n$ and $\varepsilon > 0$. Let $G_0, \ldots, G_m$ be an open cover of $X$ such that $\delta(G_i) < \varepsilon$. By definition there is a closed refinement satisfying conditions (2), (3) and $(4_n)$. Condition (5) follows from (3).

2. *Sufficiency.* Let $G_0, \ldots, G_m$ be an open cover of $X$. Let $\varepsilon$ be its coefficient (see Chapter XVI, § 5, Corollary to Theorem 7), and let $F_1^*, \ldots, F_r^*$ be a closed cover of $X$ satisfying conditions (2), $(4_n)$ and (5) (where $F_i$ has to be replaced by $F_j^*$). Since the cover $F_1^*, \ldots, F_r^*$ is a refinement of the cover $G_0, \ldots, G_m$, the sets $F_j^*$ $(j = 1, \ldots, r)$ can be distributed in $m+1$ disjoint families $C_0, \ldots, C_m$ in the following manner: to $C_0$ belong all $F_j^*$ contained in $G_0$, to $C_1$ belong all $F_j^*$ which do not belong to $C_0$ and

are contained in $G_1$ and so on; finally $C_m$ is composed of those $F_j^*$ which are contained in $G_m$ and do not belong to any $C_i$ with $i < m$. Denote by $F_i$ the union of members of $C_i$. It follows that conditions (2), (3) and ($4_n$) are fulfilled, and hence $\dim X \leqslant n$.

Theorem 3 of § 2 is obviously a particular case of the preceding one.

COROLLARY. $\dim \mathscr{I}^n \leqslant n$.

In particular, $\mathscr{I}$ can be decomposed into arbitrarily small segments with the aid of a finite system of points; thus no point belongs to any three of these segments.

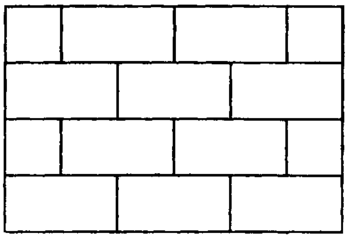


FIG. 13

A rectangle can be decomposed into small rectangles by a system of "bricks" as shown in Fig. 13 (no point belongs to 4 "bricks"). Similarly, the cube $\mathscr{I}^n$ can be decomposed into "bricks" satisfying formulas ($4_n$) and (5), and hence $\dim \mathscr{I}^n \leqslant n$.

The proof of the formula $\dim \mathscr{I}^n \geqslant n$ is less elementary (see Chapter XX).

Remark. The dimension can also be defined inductively in the following manner (which is equivalent for metric separable spaces):

1) the dimension of the void set is $-1$;

2) the dimension of a set $X$ at the point $p$ is $\leqslant n$, i.e.

$$(6) \qquad \dim_p X \leqslant n,$$

if there exists in every neighbourhood of $p$ open sets containing $p$ and having boundaries which are at most $(n-1)$-dimensional;

3) a set $X$ which has dimension $\leqslant n$ at every point is at most of dimension $n$.

Furthermore, we assume that $\dim_p X = \infty$ if formula (6) does not hold for any natural $n$, and that the dimension of $X$ is $\infty$ if it is not finite.

It follows, in particular that $\dim_p X = 0$ iff every neighbourhood of $p$ contains closed-open subsets containing $p$. For example the set consisting of the intervals

$$(1/3, 1/2),\ (1/5, 1/4),\ \ldots,\ (1/(2n+1), 1/2n),\ \ldots$$

and of the point 0, is 0-dimensional at the point 0 and only at that point.

## § 4. Properties of $n$-dimensional metric separable spaces

The following theorems on the dimension of metric separable spaces will be stated without proof.

THEOREM 1. *Every n-dimensional space has a base consisting of open sets with boundary of dimension at most $n-1$.*

THEOREM 2. *Every n-dimensional space is topologically contained in the cube $\mathscr{I}^{2n+1}$.*

In particular, every 1-dimensional set (and hence every curve) is contained topologically in the cube $\mathscr{I}^3$ and every 2-dimensional

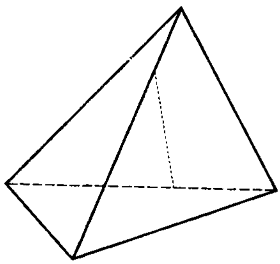


FIG. 14

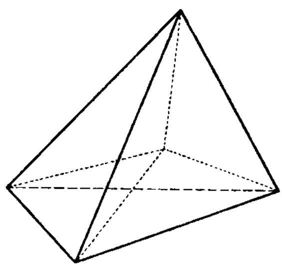


FIG. 15

set (in particular the surfaces considered in analysis) are contained in the cube $\mathscr{I}^5$.

These exponents cannot be made smaller, i.e. for every $n$ there exists an $n$-dimensional set which is not contained topologically in the cube $\mathscr{I}^{2n}$. For example, a polygonal line consisting of the edges of a tetrahedron and the segment connecting two disjoint edges (see Fig. 14) is not contained topologically in the plane (this follows easily from the Jordan theorem given in Chapter XXI, § 8).

The polygonal line shown in Fig. 15 has the same property. It consists of 6 edges of a tetrahedron and of 4 segments connecting the centre of gravity of the tetrahedron with the vertices.

Remark 1. Every polygonal line which cannot be embedded topologically in the plane contains topologically one of the two polygonal lines shown in Figs. 14 and 15.

Remark 2. If $\dim X \leqslant n$, then the set of homeomorphisms is dense in the function space $(\mathscr{I}^{2n+1})^X$.

THEOREM 3 (Sharpened normality property). *For every pair of disjoint closed sets A and B, there exists an open set G such that*

$$A \subset G, \quad \overline{G} \cap B = \emptyset, \quad \dim \mathrm{Fr}(G) \leqslant n-1.$$

THEOREM 4. *The union of a (finite or infinite) sequence of n-dimensional closed sets is an n-dimensional set.*

THEOREM 5. *For every compact n-dimensional space X, there exists a closed subset T of the Cantor discontinuum $\mathscr{C}$ and a continuous mapping of T onto X which does not assume any value more than $n+1$ times.*

For example, a closed interval can be obtained from $\mathscr{C}$ with the aid of a continuous function which does not assume any value more than twice (such a function is the step function defined in Chapter XVI, § 9, Fig. 8).

Remark. The existence of a set $T$ and of a mapping having the properties stated in Theorem 5 forms a condition which is not only necessary but also sufficient in order that $\dim X \leqslant n$.

**Exercises**

**1.** Prove that every set of real numbers which contains no interval is 0-dimensional.

**2.** Prove that the set of points in the plane, one coordinate of which is rational and the other irrational, is 0-dimensional.

**3.** Prove that the set of points in Euclidean space $\mathscr{E}^n$ all of whose coordinates are irrational is 0-dimensional.

**4.** Hint to the proof of Theorem 1 of § 2. Consider, for given $n$, all the closed-open sets with diameter $< 1/n$ and apply the Lindelöf property (Chapter XIV, § 1, Theorem 2).

**5.** Hint to the proof of Theorem 2 of § 2. Consider the characteristic function of a base consisting of closed-open sets.

CHAPTER XX

# SIMPLEXES AND THEIR PROPERTIES

## § 1. Simplexes

Definition. Let $p_0, \ldots, p_n$ be a given system of $n+1$ points in Euclidean space. By the *simplex* $p_0, \ldots p_n$ we mean the set of all points $p$ of the form

$$p = \lambda_0 p_0 + \ldots + \lambda_n p_n, \tag{1}$$

where

$$\lambda_0 + \ldots + \lambda_n = 1, \tag{2}$$

$$\lambda_i > 0, \tag{3}$$

and where the multiplication of the point by a scalar and the addition of points is to be understood as in the algebra of points (or vectors), i.e.

$$\lambda \cdot (x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n),$$
$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n).†$$

In this connection we shall always assume that the points $p_0, \ldots, p_n$ are linearly independent, i.e. that the conditions $\lambda_0 p_0 + \ldots + \lambda_n p_n = 0$ (= the origin of the axes) and $\lambda_0 + \ldots + \lambda_n = 0$ imply that $\lambda_0 = \ldots = \lambda_n = 0$ (whatever are $\lambda_0, \ldots, \lambda_n$); in other words, the points $p_0, \ldots, p_n$ do not lie in the same $(n-1)$-dimensional hyperplane. This means, in the case $n = 2$, that the points $p_0, p_1, p_2$ do not lie on a line, or that $p_0 p_1 p_2$ is a triangle (without boundary); similarly, when $n = 3$, $p_0 p_1 p_2 p_3$ is the interior of a nondegenerate tetrahedron (i.e. the points $p_0, p_1, p_2$ and $p_3$ do not lie in one plane).

† Of course, if $p = (x_1, \ldots, x_n)$ and $q = (y_1, \ldots, y_n)$, then the distance of $p$ from $q$ equals the *norm* $|p-q|$ of the vector $p-q$; let us recall that $|p| = \sqrt{x_1^2 + \ldots + x_n^2}$; this fact motivates the use of the symbol $|u-v|$ for denoting the distance between two arbitrary points $u$ and $v$.

Each of the points $p_0, \ldots, p_n$ is said to be a *vertex* of the simplex $p_0 \ldots p_n$; each of the simplexes $p_{i_0} \ldots p_{i_k}$, where $i_0 < \ldots < i_k \leqslant n$, is said to be a *face* (or edge) of the simplex.

We include the vertices as well as the entire simplex $S$ in the faces of the simplex $S = p_0 \ldots p_n$ (for $k$ assumes the values from 0 to $n$). $S_1 \leqslant S_2$ means that $S_1$ is a face of $S_2$.

Let us note that

$$\overline{S} = \bigcup p_{i_0} \ldots p_{i_k}, \tag{4}$$

for all possible systems of numbers $i_0, \ldots, i_k$, where $k$ assumes all integral values from 0 to $n$.

Finally, let us note that:

1. *the simplexes* $p_{i_0} \ldots p_{i_k}$ *in* (4) *are disjoint,*

2. *the point* $p$ *belongs to* $\overline{S}$ *when and only when it fulfils conditions* (1), (2) *and*

$$\lambda_i \geqslant 0. \tag{5}$$

The coefficients $\lambda_0, \ldots, \lambda_n$ are the *barycentric coordinates* of the point $p \in \overline{S}$; they can be interpreted as masses which must be distributed at the points $p_0, \ldots, p_n$, respectively (retaining conditions (2) and (3)), in order that the point $p$ be the centre of mass.

THEOREM 1. *Each barycentric coordinate is a continuous function of the point* $p$.

Proof. Let $T$ denote the set of points $\lambda = (\lambda_0, \ldots, \lambda_n)$ of $\mathcal{E}^{n+1}$ which satisfy conditions (2) and (5). The mapping $f$: $T \to \overline{S}$ defined by the condition $f(\lambda) = p$, where $p$ satisfies (1), is continuous, one-to-one and onto. Since $T$ is compact (because it is closed and bounded), the mapping $f$ is a homeomorphism (by Corollary 2 of Chapter XVI, § 2), and thus the mapping $f^{-1}$: $\overline{S} \to T$ is continuous. Since a barycentric coordinate is the composition of $f^{-1}$ and a projection of $\mathcal{E}^{n+1}$ on an axis, it is also continuous.

According to the above argument the following is true.

THEOREM 2. *The closure of a simplex is compact.*

## § 2. Simplicial subdivision

Let

$$S = p_0 \dots p_n.$$

By a *simplicial subdivision* of $\bar{S}$ is understood its subdivision into simplexes such that the intersection of the closures of each pair of simplexes is the closure of their common face or is empty. Figures 16 and 16a show simplicial subdivisions of a triangle.

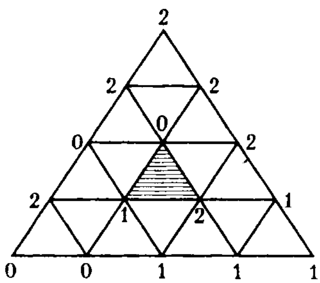



FIG. 16

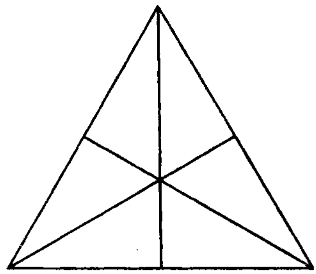


FIG. 16a

If in Figure 16 the edges of the shaded triangle were omitted, then the figure would no longer represent a simplicial subdivision.

The *centre of gravity of S* is the point

$$b(S) = \frac{1}{n+1} p_0 + \dots + \frac{1}{n+1} p_n. \tag{6}$$

Obviously $b(S)$ belongs to $S$ but belongs to no face of $S$ of dimension $< n$.

THEOREM 1. *The family **B** of all simplexes of the form $b(S_0) \dots$ $\dots b(S_k)$ where $S \geqslant S_0 > \dots > S_k$ yields a simplicial subdivision of $\bar{S}$.*

P r o o f. Each of the above considered sequences of faces can be extended to a sequence $S = S_0 > S_1 > \dots > S_n$, where

$$S_0 = p_{i_0} \dots p_{i_n}, \quad S_1 = p_{i_1} \dots p_{i_n}, \quad \dots, \quad S_n = p_{i_n} \tag{7}$$

and where $i_0, \dots, i_n$ is a permutation of the set $0, \dots, n$.

First, we shall show that the points $b(S_0), \dots, b(S_n)$ are linearly independent. Let us consider the linear combination

$$\mu_0 b(S_0) + \dots + \mu_n b(S_n).$$

Put

$$(8)\qquad \lambda_{i_j}=\sum_{k=0}^{j}\frac{\mu_k}{(n+1)-k}.$$

It follows by (6) that

$$(9)\qquad \mu_0 b(S_0)+\ldots+\mu_n b(S_n)=\lambda_{i_0}p_{i_0}+\ldots+\lambda_{i_n}p_{i_n}.$$

Furthermore

$$(10)\qquad \sum_{j=0}^{n}\lambda_{i_j}=\sum_{j=0}^{n}\sum_{k=0}^{j}\frac{\mu_k}{(n+1)-k}$$
$$=\sum_{i=0}^{n}\sum_{j=i}^{n}\frac{\mu_i}{(n+1)-i}=\sum_{i=0}^{n}\mu_i.$$

If $\mu_0 b(S_0)+\ldots+\mu_n b(S_n)=0$ and $\mu_0+\ldots+\mu_n=0$, then it follows by (9) and (10) and by virtue of the linear independence of $p_{i_0},\ldots,p_{i_n}$ that $\lambda_{i_j}=0$ for $j=0,\ldots,n$. Consequently (by (8)) $\mu_i=0$ for $i=0,\ldots,n$. Thus the points $b(S_0),\ldots,b(S_n)$ are linearly independent.

If follows from (10) that the closure of the simplex $b(S_0)\ldots b(S_n)$ is contained in $\bar{S}$. We shall show that it is identical to the set

$$(11)\qquad \{x\in\bar{S}\colon\ \lambda_{i_0}(x)\leqslant\ldots\leqslant\lambda_{i_n}(x)\}.$$

According to (8) it is sufficient to show that if $\lambda_{i_0}(x)\leqslant\ldots\leqslant\lambda_{i_n}(x)$, then there are $\mu_0,\ldots,\mu_n$ satisfying conditions

$$x=\mu_0 b(S_0)+\ldots+\mu_n b(S_n),\quad \mu_0+\ldots+\mu_n=1\quad\text{and}\quad \mu_j\geqslant 0.$$

In fact we may take

$$(12)\quad \mu_0=(n+1)\lambda_{i_0}(x)\quad\text{and}\quad \mu_j=(n+1-j)\big(\lambda_{i_j}(x)-\lambda_{i_{j-1}}(x)\big)$$

for $j=1,\ldots,n$.

It follows from (12) that the closures of faces of the simplex $b(S_0)\ldots b(S_n)$ are obtained by adding to the condition $\lambda_{i_0}(x)\leqslant\ldots\leqslant\lambda_{i_n}(x)$ a number of conditions of the form $\lambda_{i_j}(x)=\lambda_{i_{j-1}}(x)$ and perhaps of the form $\lambda_i(x)=0$. Since the intersection of a set of that kind with a similar set in the simplex $b(S_0')\ldots b(S_n')$, i.e. in a simplex corresponding to another permutation $j_0,\ldots,j_n$

of the system $0, \ldots, n$, is also of the same form (it may be void), so the intersection of the closures of any two simplexes belonging to $\boldsymbol{B}$ is the closure of their common face or is empty. Finally, since each point of $\bar{S}$ belongs to some set of the form (11), it follows that $\boldsymbol{B}$ represents a simplicial subdivision of $\bar{S}$.

The above simplicial subdivision is called the (*first*) *barycentric subdivision* of $\bar{S}$ (such is the barycentric subdivision of the triangle represented in Fig. 16a). The further barycentric subdivisions are defined inductively. If $\{S_1, \ldots, S_m\}$ is the $(k-1)$th subdivision, the $k$th subdivision is obtained by means of barycentric subdivision of each of the sets $\bar{S}_1, \ldots, \bar{S}_m$. Since the barycentric subdivision of $\bar{S}_i$ yields a barycentric subdivision of the closure of each of the faces of the simplex $S_i$, the final effect is a simplicial subdivision of $\bar{S}$.

LEMMA 1. *The diameter of the simplex $p_0 \ldots p_n$ equals the diameter of the set of its vertices, i.e.*

$$\delta(p_0 \ldots p_n) = \delta\{p_0, \ldots, p_n\}. \tag{13}$$

P r o o f. Let $p$ and $p'$ be two points of $S$. We shall first show that

$$|p-p'| \leqslant \max_i |p_i - p'|. \tag{14}$$

Let us represent $p$ in the form (1), where (2) and (3) are fulfilled. We get

$$|p-p'| = \Big|\sum_{i=0}^{n} \lambda_i p_i - \sum_{i=0}^{n} \lambda_i p'\Big| = \Big|\sum_{i=0}^{n} \lambda_i (p_i - p')\Big|$$

$$\leqslant \sum_{i=0}^{n} \lambda_i |p_i - p'| \leqslant \max_i |p_i - p'|.$$

It follows from (14) that $|p'-p_i| \leqslant \max_j |p_j - p_i|$, and hence

$$|p-p'| \leqslant \max_{i,j} |p_i - p_j|.$$

LEMMA 2. *The diameters of the simplexes of the barycentric subdivision of $\bar{S}$ are less than or equal to* $\dfrac{n}{n+1}\,\delta(S)$.

P r o o f. It is sufficient to prove that for an arbitrary permutation $i_0, \ldots, i_n$ of the system $0, \ldots, n$ and for $j < k \leqslant n$ we have

$$|b(S_j)-b(S_k)| \leqslant \frac{n}{n+1}\delta(S), \tag{15}$$

where $b(S_j) = \frac{1}{j+1}(p_{i_0}+ \ldots +p_{i_j})$ and $b(S_k) = \frac{1}{k+1}(p_{i_0}+ \ldots$ $\ldots +p_{i_k})$.

By (14) we have $|b(S_j)-b(S_k)| \leqslant |p_{i_l}-b(S_k)|$ for some $l \leqslant j < k$, and hence

$$\begin{aligned}|b(S_j)-b(S_k)| \leqslant |p_{i_l}-b(S_k)| &= \left|p_{i_l}-\frac{1}{k+1}(p_{i_0}+ \ldots + p_{i_k})\right| \\ &= \frac{1}{k+1}\left|\sum_{m=0}^{k}(p_{i_l}-p_{i_m})\right| \leqslant \frac{1}{k+1}\sum_{m=0}^{k}|p_{i_l}-p_{i_m}| \\ &\leqslant \frac{k}{k+1}\delta(S) \leqslant \frac{n}{n+1}\delta(S).\end{aligned}$$

This implies at once the following theorem.

THEOREM 2. *For each $\varepsilon > 0$ there is a simplicial subdivision of $\overline{S}$ into simplexes with diameter less than $\varepsilon$, namely the nth barycentric subdivision.*

THEOREM 3 (of Sperner). *Let $\overline{S}$ be subdivided simplicially and let the function $m(s)$ assign to each vertex of the simplexes of this subdivision an integer $m(s)$ which satisfies the following condition:*

(16) *if $s \in p_{i_0} \ldots p_{i_k}$, then $m(s)$ is one of the integers $i_0, \ldots, i_k$.*

*Then there exists among the simplexes of the subdivision under consideration at least one simplex on whose vertices the function $m(s)$ assumes all the values from 0 to $n$.*

(The shaded simplex in Fig. 16 is such a simplex.)

We shall carry out the proof by induction. We shall prove a stronger assertion, namely that the number $r$ of simplexes on whose vertices the function $m(s)$ assumes all the values from 0 to $n$, is odd.

For $n = 0$ this is obvious; for then $S = \{p_0\}$ and $r = 1$.

Let us assume that the theorem (in the stronger formulation) is valid for $n-1$. We shall prove that it is valid for $n$.

We take into consideration the family of all simplexes of $(n-1)$-dimension which appear in the given simplicial subdivision (for the subdivision represented in the figure this is the family of all sides of triangles). Among them we distinguish those simplexes on whose vertices the function $m(s)$ assumes all the values from 0 to $n-1$. We denote by $\boldsymbol{R}$ the family of these distinguished simplexes. Finally, in the family $\boldsymbol{R}$ we consider those simplexes which lie on the face $p_0 \ldots p_{n-1}$ (in the figure this is the segment [0, 1] lying at the base of the triangle); we denote by $u$ the number of these simplexes. By our assumption, $u$ is an odd number.

Let us write down the sequence

$$A_1, A_2, \ldots, A_t, A_{t+1}, \ldots, A_w$$

of all the simplexes appearing in the simplicial subdivision under consideration; let the simplexes $A_1, \ldots, A_t$ have the dimension $n$ and let the remaining have dimension $< n$.

We denote by $v_j$ for $j \leqslant t$ the number of faces of the simplex $A_j$ belonging to $\boldsymbol{R}$. Denoting by $W_j$ the set of values which the function $m(s)$ assumes on the vertices of the simplex $A_j$, we easily prove that

1. if $W_j = (0, 1, \ldots, n)$, then $v_j = 1$,
2. if $(0, 1, \ldots, n-1) \subset W_j \neq (0, 1, \ldots, n)$, then $v_j = 2$,
3. if $(0, 1, \ldots, n-1) \not\subset W_j$, then $v_j = 0$.

Therefore

$$r \equiv (v_1+v_2+ \ldots +v_t) \bmod 2.$$

On the other hand, if to each $j \leqslant t$ we assign the faces of the simplex $A_j$ belonging to $\boldsymbol{R}$ (provided that such faces exist), then every simplex belonging to $\boldsymbol{R}$ will be assigned to one or two indices $j$ depending on whether or not it lies on the face $p_0 \ldots p_{n-1}$. Hence, we have

$$v_1+v_2+ \ldots +v_t \equiv u \bmod 2, \qquad \text{whence} \qquad r \equiv u \bmod 2,$$

and therefore $r$ is an odd number (because $u$ is odd).

## § 3. Dimension of a simplex

THEOREM 1. *If the system of compact sets $F_0, \ldots, F_n$ satisfies the condition*

(17) $$p_{i_0} \ldots p_{i_k} \subset F_{i_0} \cup \ldots \cup F_{i_k}$$

*for each face of the simplex* $S = p_0 \ldots p_n$, *then*

(18) $$F_0 \cap \ldots \cap F_n \neq \emptyset.$$

P r o o f. Let us assume the contrary, i.e. that $F_0 \cap \ldots \cap F_n = \emptyset$, and denote by $\varepsilon$ the Lebesgue coefficient of the cover $G_0, \ldots, G_n$ where $G_i = \overline{S} - F_i$ (comp. Corollary to Theorem 7 of Chapter XVI, § 5).

Let there be given a simplicial subdivision of $\overline{S}$ into simplexes of diameter $< \varepsilon$. Let $s$ be a vertex of some simplex of this subdivision. By virtue of formula (4) and because of the fact that the faces of the simplex $S$ are disjoint (see § 1, 1), there exists only one face $p_{i_0} \ldots p_{i_k}$ which contains $s$; and therefore, by (17), there exists an index $i_j$ such that $s \in F_{i_j}$.

Let us set

$$m(s) = i_j, \quad \text{i.e.} \quad s \in F_{m(s)}.$$

The function $m(s)$ thus defined satisfies condition (16). Hence there exists, by virtue of Sperner theorem, a simplex $s_0 \ldots s_n$ such that for $i = 0, \ldots, n$,

$$m(s_i) = i, \text{ and hence } s_i \in F_i, \text{ i.e. } \overline{s_0 \ldots s_n} \cap F_i \neq \emptyset,$$

contrary to the formula $\delta(s_0 \ldots s_n) < \varepsilon$.

THEOREM 2. *Let* $P_i$ *be the union of all the faces of the simplex* $S$ *having* $p_i$ *for a vertex* (in other words, $P_i$ *is the set of all the points of* $\overline{S}$ *for which* $\lambda_i > 0$). *Let the system of closed sets* $F_0, \ldots, F_n$ *satisfy the conditions*:

(19) $$\overline{S} = F_0 \cup \ldots \cup F_n,$$

(20) $$F_i \subset P_i.$$

*Then condition* (17) *is satisfied and hence* (by virtue of Theorem 1) *condition* (18) *also.*

P r o o f. Let $p \in p_{i_0} \ldots p_{i_k}$. Therefore, for every $j$ distinct from each of the numbers $i_0, \ldots, i_k$ we have $\lambda_j = 0$, i.e. $p \notin P_j$, whence $p \notin F_j$ by virtue of (20). By (12)

$$p \in (F_0 \cup \ldots \cup F_n) - F_j.$$

Since this formula holds for each $j$ such that $j \neq i_0, \ldots, j \neq i_k$, it follows that $p \in F_{i_0} \cup \ldots \cup F_{i_k}$.

Thus, inclusion (17) is proved.

THEOREM 3. $\dim \bar{S} = n$.

Proof. By the Corollary of § 3 of Chapter XIX we have $\dim \mathscr{I}^n \leqslant n$. Since $\bar{S}$ is homeomorphic to $\mathscr{I}^n$, it follows that $\dim \bar{S} \leqslant n$.

We must prove that $\dim \bar{S} > n-1$.

Let us note that the family of sets $P_i$ considered in Theorem 2 is an open cover of $\bar{S}$, i.e. $\bar{S} = P_0 \cup \ldots \cup P_n$.

Now suppose contrary to our assumption that $\dim \bar{S} \leqslant n-1$. Then there exists (according to the definition of dimension, see Chapter XIX, § 3) a system of closed sets $F_0, \ldots, F_n$ satisfying conditions (19), (20) and $(4_{n-1})$ of Chapter XIX, § 3. But then condition (9) is not fulfilled, contradicting Theorem 2.

## § 4. The fixed point theorem

Let $S$ be, as before, the simplex $p_0 \ldots p_n$.

BROUWER THEOREM. *For every continuous mapping $f\colon \bar{S} \to \bar{S}$ there exists a fixed point, i.e. a point $p$ such that*

(21) $$f(p) = p.$$

Proof. We shall use the following notation: for an arbitrary $p \in \bar{S}$ we write

(22) $$f(p) = \lambda_0^* p_0 + \ldots + \lambda_n^* p_n,$$

where (analogously to (2) and (5)):

(23) $$\lambda_0^* + \ldots + \lambda_n^* = 1,$$

(24) $$\lambda_i^* \geqslant 0.$$

We have to prove that there exists a point $p$ such that

(25) $$\lambda_i^* = \lambda_i \quad \text{for every } i.$$

Let us denote by $F_i$ the set of all points $p$ for which

(26) $$\lambda_i^* \leqslant \lambda_i.$$

By virtue of the continuity of the barycentric coordinates and of the function $f$, the sets $F_i$ are closed. We shall prove that condition (17) is satisfied.

Let $p \in p_{i_0} \ldots p_{i_k}$. This means that

(27) $$\lambda_{i_0} + \ldots + \lambda_{i_k} = 1.$$

But since by (23):

$$\lambda_{i_0}^* + \ldots + \lambda_{i_k}^* \leqslant 1, \tag{28}$$

hence from (27) and (28) it follows that

$$\lambda_{i_0}^* + \ldots + \lambda_{i_k}^* \leqslant \lambda_{i_0} + \ldots + \lambda_{i_k},$$

and therefore (cf. (24)) for some $j \leqslant k$ we have $\lambda_{i_j}^* \leqslant \lambda_{i_j}$. By (26) this means that $p \in F_{i_j}$. And therefore inclusion (17) is proved.

Owing to Theorem 1, § 3, inequality (18) is satisfied. Hence let $p \in F_0 \cap \ldots \cap F_n$. This means that

$$\begin{array}{c} \lambda_0^* \leqslant \lambda_0, \\ \ldots\ldots\ldots \\ \lambda_n^* \leqslant \lambda_n. \end{array} \tag{29}$$

Adding these inequalities we obtain

$$\lambda_0^* + \ldots + \lambda_n^* \leqslant \lambda_0 + \ldots + \lambda_n,$$

which yields, by (23) and (2),

$$\lambda_0^* + \ldots + \lambda_n^* = \lambda_0 + \ldots + \lambda_n.$$

Therefore in the system of inequalities (29) there cannot appear any strict inequality of the form $\lambda_i^* < \lambda_i$. In other words, formula (25) holds.

Remarks. 1. The Brouwer theorem for $n = 1$ states that for every continuous mapping $f$ of a closed interval into any one of its subsets there exists a fixed point. This is an immediate consequence of the Darboux property of the function $f(x) - x$.

2. The Brouwer theorem is obviously also applicable to the $n$-dimensional cube as well as to any set homeomorphic to $\bar{S}$. It is interesting to note that this theorem can be generalized also to the Hilbert cube $\mathscr{H}$ and to some function spaces.

This generalization has numerous applications in the theory of differential equations in proving the existence theorems.† For, a theorem on the existence of a solution of a differential equation can be formulated as a theorem on the existence of a fixed point of some mapping of the space of continuous functions into itself (under suitable hypotheses which we shall not give here).

---

† J. Schauder, Der Fixpunktsatz in Funktionalräumen, *Studia Mathematica* 2 (1930).

Let us illustrate this by an example (cf. Chapter XVI, Exercise 21).

To solve the differential equation

$$dy/dx = f(x, y) \tag{30}$$

with initial values $x_0, y_0$, means to find a function $g$ of the variable $x$ such that

$$dg(x)/dx = f(x, g(x)) \quad \text{and} \quad g(x_0) = y_0.$$

In other words, we must find a function $g$ such that

$$g(x) = y_0 + \int_{x_0}^{x} f(t, g(t))\, dt. \tag{31}$$

Let us denote by $h$ the mapping which assigns to each function $\varphi$ the function $h_\varphi$ of the variable $x$ defined by the condition

$$h_\varphi(x) = y_0 + \int_{x_0}^{x} f(t, \varphi(t))\, dt.$$

The fixed point of this mapping is a function $g$ such that

$$h_g = g, \quad \text{i.e.} \quad h_g(x) = g(x) \text{ for every } x,$$

which means that the function $g$ satisfies equality (31).

Thus the proof of the existence of a solution of equation (30) reduces to the proof of the existence of a fixed point for the mapping $h$ (which maps a certain function space into itself).

COROLLARY. *The surface $\mathscr{S}_n$ of the set $\mathscr{K}_{n+1} = \{z\colon |z| \leqslant 1\}$ is not a retract of it; i.e. there does not exist a continuous function $f\colon \mathscr{K}_{n+1} \to \mathscr{S}_n$ such that*

$$f(x) = x \quad \textit{for} \quad x \in \mathscr{S}_n. \tag{32}$$

P r o o f. If such a function $f$ existed, then the function

$$g(x) = -f(x) \tag{33}$$

would map $\mathscr{K}_{n+1}$ onto $g(\mathscr{K}_{n+1}) \subset \mathscr{K}_{n+1}$ without a fixed point, contrary to Brouwer's theorem (see Remark 2).

In fact, if $x \in \mathscr{K}_{n+1} - \mathscr{S}_n$ then $g(x) \neq x$, since $g(x) \in \mathscr{S}_n$. But if $x \in \mathscr{S}_n$ then $g(x) = -x$ by virtue of (33) and (32), and hence we also have $g(x) \neq x$.

This completes the proof of the corollary.

We shall now give another formulation of this corollary, using the concept of *homotopy*.

D e f i n i t i o n. Let there be given two continuous mappings of the space $X$ into the space $Y$, i.e. $f, g \in Y^X$. We say that these two functions are *homotopic* if there exists a continuous function $h$ of two variables $x$ and $t$, where $0 \leqslant t \leqslant 1$ such that

$$(34) \quad h(x,t) \in Y, \quad h(x,0) = f(x) \quad \text{and} \quad h(x,1) = g(x).$$

We can state this in more intuitive manner: there exists a continuous transition from the mapping $f$ to the mapping $g$ (we interpret the parameter $t$ to be time).

Let us note that if $Y = \mathscr{E}$ (or more generally, $Y = \mathscr{E}^n$), then the functions $f$ and $g$ are always homotopic.

For it suffices to set

$$h(x,t) = f(x) + t(g(x) - f(x)).$$

If, however, $Y = \mathscr{S}_n$, then this is no longer true. Namely, *the identity and a constant are not homotopic*. This means that if

$$X = Y = \mathscr{S}_n, \quad f(x) = x, \quad g(x) = c, \quad c \in \mathscr{S}_n,$$

then there does not exist a continuous function $h$ satisfying conditions (34).

For let us assume that such a function $h$ exists and set

$$f^*(tx) = h(x, 1-t) \quad \text{for} \quad x \in \mathscr{S}_n \text{ and } 0 \leqslant t \leqslant 1.$$

Since every point of $\mathscr{K}_{n+1}$ can be represented uniquely in the form $z = tx$ (with the exception of the point $z = 0$), therefore the function $f^*$ is continuous, i.e. $f^* \in (\mathscr{S}_n)^{\mathscr{K}_{n+1}}$. And at the same time we have

$$f^*(x) = h(x,0) = f(x) = x,$$

i.e. the function $f^*$ is a retract of $\mathscr{K}_{n+1}$ to its surface. But this is impossible by the last corollary.

**Exercises**

**1.** Let $S$ be an $n$-dimensional simplex lying in the space $\mathscr{E}^n$. Prove that the boundary of the simplex $S$ is the union of all its faces of dimension $< n$.

**2.** The continuum $C$ consists of the closure of the graph of the function $y = \sin(1/x)$ for $0 < |x| < 1/\pi$ and of an arc joining the points $(-1/\pi, 0)$,

$(1/\pi, 0)$ outside of the rest of the continuum $C$. Prove that under every continuous mapping of the set $C$ onto a subset there exists a fixed point.

**3.** Let $S = p_0 \dots p_n$ be a given simplex and let $X$ be a given metric space covered with open sets: $X = G_0 \cup \dots \cup G_n$.

Consider the mapping

$$\varkappa(x) = \lambda_0(x) \cdot p_0 + \dots + \lambda_n(x) \cdot p_n,$$

where

$$\lambda_i(x) = \varrho(x, X - G_i)/\{\varrho(x, X - G_0) + \dots + \varrho(x, X - G_n)\}$$

(this is the so-called *kappa mapping*).

Prove that

(a) $\varkappa(x) \in \bar{S}$ where $\lambda_i(x)$ is the $i$th barycentric coordinate of the point $\varkappa(x)$ (i.e. conditions (2) and (5) are satisfied);

(b) $\varkappa^{-1}(P_i) = G_i$, where $P_i$ has the same meaning as in Theorem 2 of § 3;

(c) $\varkappa^{-1}(p_{i_0} \dots p_{i_k}) = G_{i_0} \cap \dots \cap G_{i_k} - \bigcup_i G_i$, where the union is over all indices $i$ different from $i_0, \dots, i_k$;

(d) $\varkappa(X - G_i) \cap P_i = \emptyset$;

(e) if every intersection of $m+2$ of the sets $G_0, \dots, G_n$ is void, then $\dim \varkappa(X) \leqslant m$.

**4.** Let $S = p_0 \dots p_n$ be a given simplex and let $f$ be a continuous mapping of $\bar{S}$ into itself. We assume that, if $p \in \mathrm{Fr}(S)$, then $f(p) \in \mathrm{Fr}(S)$ and that $f(p) \neq p$. Prove that $f(\bar{S}) = \bar{S}$.

*Hint*: Assuming that $f(\bar{S}) \neq \bar{S}$ we denote by $r$ the point belonging to $S - f(S)$ and by $g(p)$ the projection of the point $f(p)$ from $r$ into $\mathrm{Fr}(S)$. We then arrive at a contradiction of Brouwer theorem.

**5.** Let $S = p_0 \dots p_n$ and let the sets $G_0, \dots, G_n$, open in $\bar{S}$, satisfy the conditions $\bar{S} = G_0 \cup \dots \cup G_n$ and $G_i \subset P_i$ for $i = 0, \dots, n$. Then $G_0 \cap \dots \cap G_n \neq \emptyset$.

*Hint*: Make use of Theorem 2 of § 3, and of Exercise 9 of Chapter XIII.

**6.** Let $T_i$ denote the closure of the face lying opposite the vertex $p_i$. Prove that if the closed sets $F_0, \dots, F_n$ satisfy the conditions $\bar{S} = F_0 \cup \dots \cup F_n$ and $T_i \subset F_i$, then $F_0 \cap \dots \cap F_n \neq \emptyset$.

*Hint*: Use Exercise 5.

**7.** Let $S = p_0 \dots p_n$ and let $f$ be a continuous mapping of $\bar{S}$ into itself such that $f(T_i) \subset T_i$ for $i = 0, \dots, n$. Then $f(\bar{S}) = \bar{S}$.

*Hint*: Argue as in the solution of Exercise 4 and set $F_i = g^{-1}(T_i)$. Then apply Exercise 5.

**8.** Show that the relation "$f$ is homotopic to $g$" is an equivalence relation.

**9.** If the homotopic mappings $f_0$ and $f_1$ have values in the space $Y$ and if $g_0$ and $g_1$ are homotopic and defined on $Y$, then the composed mappings $g_0 \circ f_0$ and $g_1 \circ f_1$ are homotopic.

**10.** The spaces $X$ and $Y$ are said to have the same *homotopy type* if there are continuous mappings $f\colon X \to Y$ and $g\colon Y \to X$ such that $g \circ f$ and $f \circ g$ are homotopic with the identity mappings. Show that the relation "$X$ and $Y$ have the same homotopy type" is an equivalence relation.

Show that $\mathscr{E}^n$ and $\mathscr{E}^m$ have the same homotopy type, while $\mathscr{I}^n$ and $\mathscr{S}_{n-1}$ have not.

**11.** Let $X$ and $Y$ be metric. Let $X$ be compact. Then the mappings $f, g \in Y^X$ are homotopic iff they can be joined in $Y^X$ by a continuum which is a continuous image of $\mathscr{I}$.

*Hint*: Use Exercise 8 of Chapter XVI.

Remark. Using the Theorem 2 of Chapter XVIII, § 5, one can replace the above condition by the existence of an arc joining $f$ to $g$ in $Y^X$.

**12.** Show that the Corollary of § 4 follows from the theorem (shown above) stating that on $\mathscr{S}_n$ the identity is not homotopic to a constant.

Show also that this Corollary implies the fixed point theorem.

*Hint*. Supposing that $f\colon \mathscr{K}_{n+1} \to \mathscr{K}_{n+1}$ has no fixed point, consider for each $x$ the intersection of $\mathscr{S}_n$ with the ray containing the points $f(x)$ and $x$, the first being its vertex.

## CHAPTER XXI

# CUTTINGS OF THE PLANE

### § 1. Auxiliary properties of polygonal arcs

As usual, we shall denote the plane of complex numbers by $\mathscr{E}^2$. By $\mathscr{S}_2$ we denote the plane $\mathscr{E}^2$ extended by the point at infinity (called the *Gauss plane*); topologically $\mathscr{S}_2$ does not differ from the surface of a three-dimensional ball.

THEOREM 1. *Any two points of a connected open set $R$ (i.e. of a region) situated on $\mathscr{S}_2$ can be joined by a polygonal arc.*

The proof is completely analogous to the proof of Theorem 4 of Chapter XVIII, § 4: by $F$ we denote the set of all points of the region $R$ which can be joined by a polygonal arc with the fixed point $p \in R$ and then we prove that this set is nonvoid and open and that the set $R-F$ is open; the connectedness of the set $R$ implies that $F = R$.

THEOREM 2. *If $L$ is a polygonal arc $\subset \mathscr{S}_2$, then the set $\mathscr{S}_2 - L$ is homeomorphic to the plane $\mathscr{E}^2$.*

The proof is by induction on the number $n$ of links in the polygonal arc.

For $n = 1$ we have to prove that the Gauss plane minus a segment is homeomorphic with the Gauss plane minus a point.

To this end, we describe about the centre of the segment $L$ a sequence of concentric circles $K_1, K_2, \ldots$ with radii tending to 0. Let $E_1, E_2, \ldots$ be a sequence of ellipses (together with their interiors) whose intersection is the segment $L$; we may assume here that $E_1 = K_1$ (Fig. 17).

We define the required homeomorphism $h$ as follows: on the exterior of the circle $E_1$ we set $h(z) = z$. Next we map the annulus $\overline{E_1 - E_2}$ homeomorphically onto the annulus $\overline{K_1 - K_2}$ without affecting the values of $h$ on $\mathrm{Fr}(E_1)$; in general, we map the annulus

$\overline{E_m - E_{m+1}}$ onto $\overline{K_m - K_{m+1}}$ without affecting the values of $h$ on $\mathrm{Fr}(E_m)$.

The theorem is thus proved for $n = 1$.

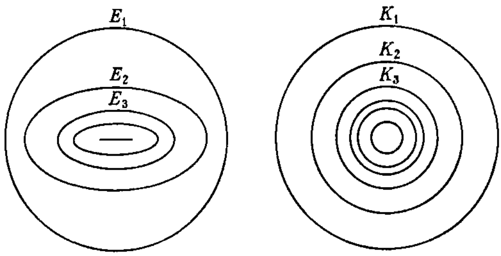



FIG. 17

For $n = 2$ the polygonal arc $L$ consists of two segments $A_1$ and $A_2$. We carry out a homeomorphic mapping of $\mathscr{S}_2$ onto $\mathscr{S}_2$, which leaves the segment $A_1$ invariant, but maps $A_2$ into a rectilinear extension of the segment $A_1$. The proof thus reduces to the case $n = 1$.

A similar method allows us, in the case where $L$ consists of $n+1$ segments, to "straighten out" the last segment (perhaps contracting it) in order to obtain a polygonal arc consisting of $n$ sides.

Remarks. Theorems 1 and 2 are valid in the space $\mathscr{E}^n$ for arbitrary $n$. For $n = 2$ Theorem 2 can be sharpened by replacing the polygonal arc $L$ by an arbitrary arc; namely, the complement of an arc contained in $\mathscr{E}^2$ is homeomorphic to the complement of a point. On the other hand, for $n = 3$ the theorem thus sharpened is not valid: there exists in $\mathscr{E}^3$ an arc, the so-called *Antoine's arc*, whose complement is not homeomorphic to the complement of a point.

## • § 2. Cuttings

We say that the (closed or open) set $A$ is a *cutting* of the space $\mathscr{S}_2$ (or: that it separates or cuts this space) if the set $\mathscr{S}_2 - A$ is not connected.

*A separates $\mathscr{S}_2$ between the points p and q* if these points belong to distinct components of $\mathscr{S}_2 - A$.

THEOREM. *If the closed set A cuts $\mathscr{S}_2$ between p and q, then there exist two closed sets R and Q such that*

$$\mathscr{S}_2 = R \cup Q, \quad p \in R, \quad q \in Q \quad \textit{and} \quad R \cap Q = A.$$

Proof. Let $M$ be a component of the set $\mathscr{S}_2 - A$ which contains the point $p$, and let $N$ be the union of all the remaining components of this set. Since the components of $\mathscr{S}_2 - A$ are open (see Chapter XVIII, § 2, Theorem 2) and the sets $M$, $N$ and $A$ are disjoint, then the sets $R = M \cup A$ and $Q = N \cup A$ are closed and, as can easily be seen, they satisfy the desired conditions.

## § 3. Complex functions which vanish nowhere. Existence of the logarithm

We shall denote by the letter $\mathscr{P}$ the plane minus the point 0, i.e.

$$\mathscr{P} = \mathscr{E}^2 - \{0\}.$$

We say that the function $f \in \mathscr{P}^A$ (i.e. continuous, defined on $A$, complex valued and everywhere different from 0) has a *single-valued continuous branch of the logarithm* if it is of the form

$$f(z) = e^{u(z)}, \quad \text{where} \quad u \in (\mathscr{E}^2)^A \tag{1}$$

(the function $u$ is this branch). We then write

$$f \sim 1.$$

More generally: if $f \in \mathscr{P}^A$ and $B \subset A$, then we write

$$f \sim 1 \quad \text{on} \quad B,$$

if there exists a function $u \in (\mathscr{E}^2)^B$ such that

$$f(z) = e^{u(z)} \quad \text{for} \quad z \in B. \tag{2}$$

A fundamental theorem for the topology of the plane, which is our nearest goal, is the following theorem:

EILENBERG THEOREM. *Let A be a compact or open subset of the space $\mathscr{P}$. A necessary and sufficient condition that the set A does not separate $\mathscr{S}_2$ between the points* 0 *and* $\infty$ *is, that the identity* *has, on the set $A$, a single-valued continuous branch of the logarithm, i.e. that there exists a function $u \in (\mathscr{E}^2)^A$ such that*

$$z = e^{u(z)} \quad \textit{for} \quad z \in A.$$

## § 4. Auxiliary theorems

THEOREM 1. *Let $R$ denote a ray lying in the plane and emanating from the point 0. Then $z \sim 1$ on the set $\mathscr{E}^2 - R$.*

P r o o f. Let $\varphi$ be the angle between $R$ and the positive direction of the $x$-axis; we assume that $0 \leqslant \varphi < 2\pi$.

Since every point $z$ of the plane is of the form $z = |z|e^{i\alpha}$, we can assume that $\varphi - 2\pi < \alpha < \varphi$ for points $z$ not belonging to $R$. The function

$$(3) \qquad u(z) = \log z = \log|z| + i\alpha$$

is continuous on $\mathscr{E}^2 - R$ and satisfies the identity

$$z = e^{u(z)} \quad \text{for} \quad z \in \mathscr{E}^2 - R.$$

From this we obtain the following theorem.

THEOREM 2. *If $f \in (\mathscr{E}^2 - R)$ then $f \sim 1$ (whatever $X$ is).*

For, the function $u(f(x))$ is continuous on the set $X$ and

$$f(x) = e^{u(f(x))} \quad \text{for} \quad x \in X$$

(where $u$ is the function defined by the formula (3)).

THEOREM 3. *Let $f \in \mathscr{P}^X$. To every point $x \in X$ there corresponds a certain neighbourhood $G$ such that*

$$(4) \qquad f \sim 1 \quad \textit{on} \quad G.$$

P r o o f. Let $R$ be a ray emanating from the point 0 and not containing the point $f(x)$ (such a ray exists because $f(x) \neq 0$). Owing to the continuity of $f$ we have

$$R \cap f(G) = \emptyset, \quad \text{i.e.} \quad f(G) \subset \mathscr{E}^2 - R$$

for some neighbourhood $G$ of the point $x$.

This means that the function $f$ restricted to $G$ satisfies the assumption of Theorem 2. Hence, we have formula (4).

THEOREM 4. *Let $f \in \mathcal{P}^X$, let $a \in X$ and let $c$ be one of the values of* $\log f(a)$. *If $f \sim 1$, we can choose the function $u$ satisfying formula* (1) *in such a way that it satisfies the "initial" condition*:

$$u(a) = c. \tag{5}$$

Proof. Since $f \sim 1$, the function $f$ is of the form

$$f(x) = e^{v(x)} \quad \text{where} \quad v \in (\mathcal{E}^2)^X.$$

Let us set

$$u(x) = v(x) - v(a) + c. \tag{6}$$

Hence, we have

$$e^{u(x)} = e^{v(x)} \cdot e^{-v(a)} \cdot e^c = f(x),$$

since

$$e^{-v(a)} = 1/f(a) \quad \text{and} \quad e^c = e^{\log f(a)} = f(a).$$

Hence the function $u$ satisfies condition (1). Moreover, formula (6) immediately implies formula (5).

Remark. The initial condition (5) in general does not determine the function $u$ uniquely. We have uniqueness, however, under the assumption that $X$ is connected. This follows from the following theorem:

THEOREM 5. *If $X$ is connected and*

$$f(x) = e^{u(x)} = e^{v(x)}, \tag{7}$$

*then* $v(x) = u(x) + \text{constant}$.

Proof. By virtue of (7), $e^{v(x)-u(x)} = 1$, and therefore for every $x$ there exists an integer $k(x)$ such that $v(x) - u(x) = 2k(x)\pi i$. Hence, the function $k(x)$ is continuous. Since $k(x)$ is defined on a connected set and has integral values, it is therefore constant (for the continuous image of a connected set is connected (cf. Chapter XVII, § 2, Theorem 1)).

THEOREM 6. *If $F$ is a closed subset of a metric space $X$ and the function $f \in \mathcal{P}^F$ satisfies the condition $f \sim 1$, then there exists a function $g \in \mathcal{P}^X$ which is an extension of the function $f$ and which satisfies the condition $g \sim 1$.*

Proof. By assumption, formula (1) is satisfied, and because of the Tietze Extension Theorem (Chapter XII, § 5, Corollary 1) the function $u$ can be extended to the entire space $X$. Let $v$ be this extension. Hence, we have

$$v \in (\mathscr{E}^2)^X \quad \text{and} \quad v(x) = u(x) \text{ for } x \in F.$$

The function $g(x) = e^{v(x)}$ is the desired function.

THEOREM 7. *Let $A$ and $B$ be two closed or two open sets with connected intersection. Let $f \in \mathscr{P}^{A \cup B}$. If $f \sim 1$ on $A$ and on $B$, then $f \sim 1$ on $A \cup B$.*

Proof. By assumption there exist two functions $u \in (\mathscr{E}^2)^A$ and $v \in (\mathscr{E}^2)^B$ such that

$$f(x) = \begin{cases} e^{u(x)} & \text{for} \quad x \in A, \\ e^{v(x)} & \text{for} \quad x \in B. \end{cases} \tag{8}$$

Let $A \cap B \neq \emptyset$, and $a \in A \cap B$. We can assume that $v$ has been so chosen that $v(a) = u(a)$ (cf. Theorem 4). Since the set $A \cap B$ is connected, it follows (by virtue of Theorem 5) that $v(x) = u(x)$ for every $x \in A \cap B$. Hence, if we assume that

$$w(x) = \begin{cases} u(x) & \text{for} \quad x \in A, \\ v(x) & \text{for} \quad x \in B, \end{cases} \tag{8'}$$

then—as can easily be verified (see Exercise 1 of Chapter XII)—the function $w$ is continuous, i.e. $w \in (\mathscr{E}^2)^{A \cup B}$. As $f(x) = e^{w(x)}$ for $x \in A \cup B$ (cf. (8)), hence $f \sim 1$.

We arrive at the same conclusion if $A \cap B = \emptyset$.

THEOREM 8. *Let $f \in \mathscr{P}^X$ and let $C_1, C_2, \ldots, C_n, \ldots$ be a sequence of connected sets such that*

$$X = C_1 \cup C_2 \cup \ldots \cup C_n \cup \ldots, \tag{9}$$

*and*

$$C_n \subset \operatorname{Int}(C_{n+1}) \quad \textit{for} \quad n = 1, 2, \ldots \tag{10}$$

*If $f \sim 1$ on $C_n$ for every $n$, then $f \sim 1$ (on $X$).*

Proof. Let $a \in C_1$. By assumption we have

$$f(x) = e^{u_n(x)} \quad \text{for} \quad x \in C_n \text{ and } u_n \in (\mathscr{E}^2)^{C_n}. \tag{11}$$

We can assume (see Theorem 4) that $u_n(a) = u_1(a)$. It follows, by virtue of the connectedness of the set $C_1$, that $u_n(x) = u_1(x)$ for $x \in C_1$, and since $u_{n+1}(a) = u_n(a)$, we have similarly

$$u_{n+1}(x) = u_n(x) \quad \text{for} \quad x \in C_n. \tag{12}$$

Let

(13) $$u(x) = u_n(x) \quad \text{for} \quad x \in C_n.$$

Because of (12) and (9), formula (13) defines the function $u$ uniquely for every $x \in G$. This is a continuous function. For, if $x_0 \in C_n$, then by virtue of (10) $x_0 \in \operatorname{Int}(C_{n+1})$; but since $u(x) = u_{n+1}(x)$ for $x \in C_n$, the continuity of the function $u_{n+1}$ at the point $x_0$ implies the continuity of the function $u$ at this point (cf. Chapter XII, Exercise 2). Finally, formulas (11) and (12) yield

$$f(x) = e^{u(x)} \text{ for } x \in X, \quad \text{i.e.} \quad f \sim 1.$$

THEOREM 9. *Let $G$ be an open subset of a metric separable (or, more generally, perfectly normal, see footnote on p. 242) and locally connected space. Let $f \in \mathscr{P}^G$. If*

(14) $$f \sim 1 \quad \text{on} \quad C$$

*for every closed connected subset of $G$, then $f \sim 1$ (on $G$).*

Proof. 1. Let us first assume that $G$ is connected. Then by Theorem 7 of Chapter XVIII, § 2, there exists a sequence $R_1, R_2, \ldots$ of open connected sets such that

(15) $$G = R_1 \cup R_2 \cup \ldots \quad \text{and} \quad \overline{R}_n \subset R_{n+1}.$$

Put $C_n = \overline{R}_n$. Obviously conditions (9) and (10) are fulfilled (replacing $X$ by $G$ and $C$ by $R$ in (9)), and therefore $f \sim 1$ on $G$.

2. If $G$ is not connected, we consider its decomposition into components

(16) $$G = G_1 \cup G_2 \cup \ldots \cup G_n \cup \ldots$$

Since $G_n$ is connected and open (by virtue of Theorem 2 of Chapter XVIII, § 2), it follows from the part of the theorem already proved that

$$f \sim 1 \quad \text{on} \quad G_n,$$

i.e. $f(x) = e^{v_n(x)}$ for $x \in G_n$, and $v_n \in (\mathscr{E}^2)^{G_n}$.

Let $v(x) = v_n(x)$ for $x \in G_n$. Since the sets $G_n$ are open, it follows (cf. Chapter XII, Exercise 2) that the function $v$ is continuous. Hence we have

(17) $$f(x) = e^{v(x)}, \text{ where } v \in (\mathscr{E}^2)^G, \quad \text{i.e.} \quad f \sim 1.$$

## § 5. Corollaries to the auxiliary theorems

COROLLARY 1. *If* $F = \bar{F} \subset \mathscr{I}$ *and* $f \in \mathscr{P}^F$, *then* $f \sim 1$.

P r o o f. By Theorem 3 of § 4, there is an open cover $\{G_t\}$, $t \in F$, of $F$ such that $t \in G_t$ and $f \sim 1$ on $F \cap G_t$. Since $F$ is compact, we may assume that this cover is finite. Accordingly there exists a system of points $0 = a_0 < a_1 < \ldots < a_n = 1$ such that $f \sim 1$ on the intersection $F \cap (a_{k-1}a_k)$ for $k = 1, 2, \ldots, n$.

The intersection $[F \cap (a_0a_1)] \cap [F \cap (a_1a_2)]$ being contained in $\{a_1\}$ is connected (perhaps void). Hence we have $f \sim 1$ on $F \cap (a_0a_1 \cup a_1a_2) = F \cap (a_0a_2)$ by virtue of Theorem 7 of § 4.

Similarly, $f \sim 1$ on $F \cap (a_0a_2 \cup a_2a_3) = F \cap (a_0a_3)$.

By induction we prove that $f \sim 1$ on $F \cap (a_0a_n) = F$.

COROLLARY 2. *Let* $K$ *be a square (with interior)* $\subset \mathscr{E}^2$. *Every function* $f \in \mathscr{P}^K$ *satisfies the formula* $f \sim 1$.

P r o o f. Let us decompose the square $K$ into a finite number of squares $A_1, A_2, \ldots, A_n$, enumerating them in such a way that the intersection

$$A_k \cap (A_1 \cup \ldots \cup A_{k-1}) \tag{18}$$

is connected for $k = 2, 3, \ldots, n$ (cf. Fig. 18). We assume that these squares are so small that $f \sim 1$ on each of them individually (according to Theorem 3 of § 4).

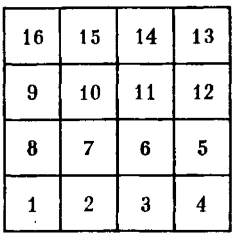



FIG. 18

Since $f \sim 1$ on $A_1$ and on $A_2$ and since the intersection $A_1 \cap A_2$ is connected, we have $f \sim 1$ on $A_1 \cup A_2$. Reasoning by induction and using the fact that the intersection (18) is connected, we deduce that $f \sim 1$ on $A_1 \cup \ldots \cup A_n$, i.e. on $K$.

COROLLARY 3. *Every function* $f \in \mathscr{P}^{(\mathscr{E}^2)}$ *satisfies the formula* $f \sim 1$.

P r o o f. Let $K_n$ be a square with side $n$ and with centre 0. Since

$$\mathscr{E}^2 = K_1 \cup K_2 \cup \ldots \cup K_n \cup \ldots$$

and since $f \sim 1$ on $K_n$ by virtue of the preceding theorem, we deduce from Theorem 8 of § 4, that $f \sim 1$ on $\mathscr{E}^2$.

R e m a r k. Corollaries 1–3 hold not only for the sets $\mathscr{I}$, $K$ and $\mathscr{E}^2$, but also for arbitrary sets which are homeomorphic to these sets; in particular, for arbitrary arcs, for a circular disk, for the complement (with respect to $\mathscr{S}_2$) of a polygonal arc.

Let us carry out the proof for the arc.

Let $h$ be a homeomorphic mapping of the segment $\mathscr{I}$ onto the arc $A$. Let $f \in \mathscr{P}^A$. Substituting $z = h(x)$ for $x \in \mathscr{I}$, we therefore have

$$f(z) = fh(x) = e^{u(x)} = e^{uh^{-1}(z)} = e^{v(z)},$$

where $v(z) = uh^{-1}(z)$, $z \in A$.

COROLLARY 4. *Let $C$ denote the circumference $|z| = r$. There does not exist a single-valued branch of the logarithm on $C$; that is, $z$ not $\sim 1$ on $C$.*

P r o o f. Let $z_0 = (r, 0)$, and $A = C - \{z_0\}$. For $z \in A$ we have

$$z = re^{i\alpha(z)}, \quad \text{where} \quad 0 < \alpha(z) < 2\pi. \tag{19}$$

Obviously the function $\alpha$ is continuous on $A$.

Let us assume that our theorem is false. Then

$$z = re^{i\beta(z)},$$

where $\beta$ is a real valued function continuous on $C$.

Since $A$ is connected we have (see Theorem 5 of § 4):

$$\alpha(z) = \beta(z) + \text{constant}. \tag{20}$$

It would then follow from this that the function $\alpha$ can be extended in a continuous manner onto $C$. But this is impossible. For, let $\lim\limits_{n\to\infty} z_n = z_0$. If the points $z_n$ lie above the $x$-axis, then $\lim\limits_{n\to\infty} \alpha(z_n) = 0$, and if they lie below the $x$-axis, then $\lim\limits_{n\to\infty} \alpha(z_n) = 2\pi$.

## § 6. Theorems on the cuttings of the plane

Proof of Eilenberg theorem (see § 3). Let $A \subset \mathscr{P}$. We shall consider separately the case where $A$ is a closed subset of $\mathscr{S}_2$ and the case where $A$ is an open set.

1. *A is closed.* Let us assume that $A$ does not separate $\mathscr{S}_2$ between the points $p = 0$ and $q = \infty$. We have to prove that

$$(21) \qquad z \sim 1 \quad \text{on} \quad A.$$

Since the points $p$ and $q$ lie in one of the components of the set $\mathscr{S}_2 - A$, there exists a polygonal arc $L$ (cf. Theorem 1 of § 1) such that

$$(22) \qquad L = pq \subset \mathscr{S}_2 - A.$$

By Theorem 2 of § 1, the set $\mathscr{S}_2 - L$ is homeomorphic to the plane $\mathscr{E}^2$, and hence by virtue of Corollary 3 of § 5 (cf. Remark) we have $z \sim 1$ on $\mathscr{S}_2 - L$, whence formula (21) follows, for $A \subset \mathscr{S}_2 - L$ by (22).

Let us assume next that $A$ separates $\mathscr{S}_2$ between the points $p = 0$ and $q = \infty$. Hence, there exist (see § 2) two closed sets $R$ and $Q$ such that

$$(23) \qquad \mathscr{S}_2 = R \cup Q, \quad p \in R, \quad q \in Q,$$

$$(24) \qquad R \cap Q = A.$$

We shall show that the assumption (21) leads to a contradiction. In fact, from (21) it follows that (cf. Theorem 6 of § 4)

$$(25) \qquad z = e^{u(z)} \text{ on } A, \quad \text{where} \quad u \in (\mathscr{E}^2)^{\mathscr{S}_2}.$$

Let us set

$$(26) \qquad f(z) = \begin{cases} e^{u(z)} & \text{if} \quad z \in Q, \\ z & \text{if} \quad z \in R \text{ and } z \neq 0. \end{cases}$$

By (24) and (25), the function $f$ is defined and continuous for every $z \neq 0$ (cf. Chapter XII, Exercise 1), i.e.

$$(27) \qquad f \in \mathscr{P}^{\mathscr{S}_2 - \{0\}}, \quad \text{whence} \quad f \sim 1$$

by virtue of Corollary 3 of § 5 (cf. Remark).

Since the point 0 does not belong to $Q$, there exists a disk with centre at the point 0 which is disjoint from $Q$ and hence contained in $R$. Let $C$ be the circumference of this disk. Hence we have (cf. (27)) $f \sim 1$ on $C$, i.e. by (26) $z \sim 1$ on $C$. But this contradicts Corollary 4 of § 5.

2. *A is open.* Let us assume that the set $A$ does not separate the plane $\mathscr{S}_2$ between the points $p = 0$ and $q = \infty$, i.e. that these points lie in the same component $T$ of the set $\mathscr{S}_2 - A$. Hence, if $F = \bar{F} \subset A$, then the points $p$ and $q$ lie in one component of the set $\mathscr{S}_2 - F$ (namely in the one that contains the set $T$). As already proved, we therefore have $z \sim 1$ on $F$. From this, by virtue of Theorem 9 of § 4, we obtain formula (21).

Let us assume next that the set $A$ separates the plane $\mathscr{S}_2$ between the points $p = 0$ and $q = \infty$, i.e. that these points belong to distinct components of the set $\mathscr{S}_2 - A$. Therefore, there exist two closed sets $M$ and $N$ (see Chapter XVII, § 6, Theorem 6) such that

$$\mathscr{S}_2 - A = M \cup N, \quad p \in M, \quad q \in N, \tag{28}$$

$$M \cap N = \emptyset. \tag{29}$$

Since the space $\mathscr{S}_2$ is normal (see Chapter XII, § 4, Theorem 6) and because of formula (29), there exist two open sets $R$ and $Q$ such that

$$M \subset R, \quad N \subset Q, \tag{30}$$

$$R \cap Q = \emptyset. \tag{31}$$

Let

$$F = \mathscr{S}_2 - (R \cup Q). \tag{32}$$

The set $F$ is therefore closed. Because of (30) and (28), we have

$$F = \mathscr{S}_2 - (R \cup Q) \subset \mathscr{S}_2 - (M \cup N) = A, \tag{33}$$

$p \in R$ and $q \in Q$.

Thus $\mathscr{S}_2 - F$ is the union of two open disjoint sets $R$ and $Q$ of which one contains $p$ and the other contains $q$ (cf. (32)). The set $F$ therefore separates $\mathscr{S}_2$ between these points. By virtue of the part of the theorem already proved, we have $z$ not $\sim 1$ on $F$.

But since $F \subset A$ (because of (33)) we have *a fortiori* $z$ not $\sim 1$ on $A$.

## § 7. Janiszewski theorems

THEOREM 1. *Let $A$ and $B$ be two closed or two open subsets of $\mathscr{S}_2$. If neither of these sets separates $\mathscr{S}_2$ between the points $p$ and $q$ and if the intersection $A \cap B$ is connected, then the union $A \cup B$ also does not separate $\mathscr{S}_2$ between these points.*

Proof. By means of the homographic transformation

$$h(z) = (z-p)/(z-q) \tag{34}$$

we reduce the proof to the case where

$$p = 0, \quad q = \infty. \tag{35}$$

Hence let us assume that the equalities (35) hold.

Since neither $A$ nor $B$ separates the plane $\mathscr{S}_2$ between the points $p$ and $q$, the relations

$$z \sim 1 \text{ on } A \quad \text{and} \quad z \sim 1 \text{ on } B$$

hold, by the Eilenberg theorem.

It follows from this, by virtue of Theorem 7 of § 4, that $z \sim 1$ on $A \cup B$. And therefore, by the Eilenberg theorem, $A \cup B$ does not separate $\mathscr{S}_2$ between $p$ and $q$.

THEOREM 2. *As before, let $A$ and $B$ be two open or two closed subsets of $\mathscr{S}_2$. If the sets $A$ and $B$ are connected but the intersection $A \cap B$ is not connected, then the union $A \cup B$ separates $\mathscr{S}_2$ between some pair of points.*

Proof. We use the usual notation

$$A^c = \mathscr{S}_2 - A, \quad B^c = \mathscr{S}_2 - B.$$

Let us assume—contrary to our theorem—that the set $A \cup B$ does not separate $\mathscr{S}_2$, i.e. that the set $\mathscr{S}_2-(A \cup B) = A^c \cap B^c$ is connected. We shall prove that then the assumptions of Theorem 1 are satisfied by the sets $A^c$ and $B^c$ where $p$, $q$ is an arbitrary pair of points belonging to $A \cap B$.

In fact, both $A^c$ and $B^c$ are open or both are closed, and their intersection $A^c \cap B^c$ is connected. It remains to prove that neither the set $A^c$ nor the set $B^c$ separates $\mathscr{S}_2$ between the points $p$ and $q$, i.e. that these points belong to some component of the complement of the set $A^c$, i.e. to some component of the set $A$, and, similarly, to some component of the set $B$. But this follows immediately from the assumption that the sets $A$ and $B$ are connected and contain the points $p$ and $q$.

Applying the first Janiszewski theorem to the sets $A^c$ and $B^c$, we deduce that the union $A^c \cup B^c$ does not separate $\mathscr{S}_2$ between $p$ and $q$, i.e. that $p$ and $q$ belong to the same component of the set $(A^c \cup B^c)^c = A \cap B$. But since the points $p$ and $q$ are arbitrary points belonging to $A \cap B$, it follows from this that the set $A \cap B$ is connected, contrary to assumption.

## § 8. Jordan theorem

*Every simple closed curve $C \subset \mathscr{S}_2$ (i.e. every set homeomorphic to $\mathscr{S}_1$) decomposes $\mathscr{S}_2$ into two regions and is their common boundary.*

We precede the proof with the following lemma.

LEMMA. *No arc or closed subset of an arc separates $\mathscr{S}_2$.*

Proof. Let us assume, on the contrary, that some closed subset $F$ of the arc $L$ separates $\mathscr{S}_2$ between the points $p$ and $q$. Applying the homographic transformation (34) we can assume that these points are $p = 0$ and $q = \infty$. By the Eilenberg theorem we have $z$ not $\sim 1$ on $F$. But this contradicts Corollary 1 of § 5 (see § 5, Remark).

Proof of the Jordan theorem. Since the curve $C$ can be represented as the union of two arcs whose intersection is not connected (namely consisting of two points) we deduce from the second Janiszewski theorem that $C$ separates $\mathscr{S}_2$.

Let

$$R_1, R_2, \ldots \tag{36}$$

denote the sequence of components of the set $\mathscr{S}_2 - C$. We have proved that this sequence contains at least two terms. It remains to prove that it does not contain more than two terms and that

$$\mathrm{Fr}(R_1) = C = \mathrm{Fr}(R_2). \tag{37}$$

We shall begin with the proof of formula (37). By virtue of Theorem 5 of Chapter XVIII, § 2, we have

$$\mathrm{Fr}(R_1) \subset C. \tag{38}$$

If formula (37) did not hold, then the set $\mathrm{Fr}(R_1)$ would be a closed subset of some arc (contained in $C$) and therefore, by the lemma, it would not separate $\mathcal{S}_2$. But this is impossible because $\mathrm{Fr}(R_1)$ obviously separates $\mathcal{S}_2$ between every point of $R_1$ and every point of $R_2$.

Hence the first of the identities (37) is proved and the second is obtained by symmetry.

It remains to prove that the sequence (36) consists of two terms.

Let us assume the contrary, i.e. that there exist at least three regions $R_1, R_2, R_3$. Let

(39) $$p_j \in R_j \quad \text{for} \quad j = 1, 2, 3.$$

Let us assume that the region $R_3$ is bounded. Let $Z$ be a straight line passing through the point $p_3$. This straight line therefore contains a segment $L = ap_3b$ lying in $R_3$ with the exception of the endpoints which belong to $C$:

(40) $$L \subset R_3 \cup \{a\} \cup \{b\}.$$

Let $aq_1b$ and $aq_2b$ be the arcs of the curve $C$ determined by the points $a$ and $b$ (see Fig. 19).

Hence we have

(41) $$aq_1b \cup aq_2b = C,$$

and

(42) $$aq_1b \cap aq_2b = \{a, b\}.$$

Let

(43) $$A_1 = aq_1b \cup L, \quad A_2 = aq_2b \cup L.$$

It follows from formulas (42) and (43) that

(44) $$A_1 \cap A_2 = L.$$

Since $q_1, q_2 \in C$, we therefore deduce from (37) that the sets $R_1 \cup \{q_1\} \cup R_2$ and $R_1 \cup \{q_2\} \cup R_2$ are connected, and from (39) that they contain the points $p_1$ and $p_2$. Since these sets are disjoint from $A_2$ and $A_1$ respectively (cf. (40) and (43)), the sets $A_1$ and $A_2$ do not separate $\mathcal{S}_2$ between $p_1$ and $p_2$. From formula (44) we deduce by virtue of the first Janiszewski theorem that $A_1 \cup A_2$ does not separate $\mathcal{S}_2$ between $p_1$ and $p_2$ either.

But this is impossible because (cf. (41) and (43)) $A_1 \cup A_2 = C \cup L$, and $C$ separates $\mathscr{S}_2$ between $p_1$ and $p_2$.

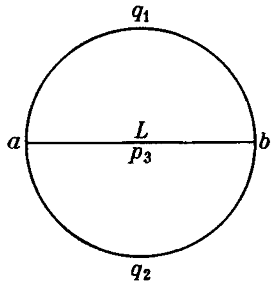



FIG. 19

*R e m a r k 1. We can sharpen the Jordan theorem by introducing the interesting concept of accessible point. Namely, we say that a point $p$ lying on the boundary of the region $R$ is *accessible* from this region if there exists an arc containing the point $p$ and lying entirely—with the exception of the point $p$—in the region $R$.

An example of a point which is not accessible is the following. Let $C$ be the closure of the curve $y = \sin(1/x)$, $0 < |x| \leqslant 1$, and let $R$ be the complement of the continuum $C$; the point $\langle 0, 0\rangle$ is not accessible from the region $R$.

One can prove that *every point of a simple closed curve is accessible from both regions into which the curve separates the plane.*

In the general case of an arbitrary region $R \subset \mathscr{E}^n$, the points which are accessible from $R$ form a dense set on its boundary.

For let $p \in \mathrm{Fr}(R)$. For $\varepsilon > 0$ there exists a point $q \in R$ at a distance $< \varepsilon$ from $p$. On the segment $qp$ let $r$ be the first point (starting from $q$) of the set $\mathrm{Fr}(R)$. Therefore the segment $qr$ lies—with the exception of the point $r$—entirely in the region $R$. Hence, the point $r$ is accessible from $R$. At the same time $|r-p| \leqslant |q-p| < \varepsilon$.

*R e m a r k 2. Another generalization of the Jordan theorem is the following theorem (of Schönflies):

*Let $C$ be a simple closed curve contained in $\mathscr{S}_2$. Every homeomorphism $h$ of $\mathscr{S}_1$ onto $C$ can be extended to a homeomorphism $h^*$ of the entire plane $\mathscr{S}_2$ onto itself; i.e. $h^*(\mathscr{S}_2) = \mathscr{S}_2$ and $h^*(p) = h(p)$ for $p \in \mathscr{S}_1$.*

On the basis of this theorem one can prove that every topological property of $\mathscr{S}_1$ with respect to the plane $\mathscr{S}_2$ (such as the number of components in $\mathscr{S}_2 - \mathscr{S}_1$ and the accessibility of points on the circumference) also holds for any simple closed curve.

An analogous theorem concerns arcs lying in $\mathscr{S}_2$: *every homeomorphism defined on the segment $\mathscr{I}$ can be extended to a homeomorphism of $\mathscr{S}_2$ onto $\mathscr{S}_2$.*

However, this theorem is not valid for arcs (nor for simple closed curves) lying in $\mathscr{E}^3$. The Antoine arc referred to in the remark at the end of § 1 is a counter-example.

*R e m a r k 3. Jordan theorem is a special case of the following theorem on the invariance of the number of components of the complement of a closed set lying on the sphere $\mathscr{S}_n$ (i.e. on the surface of the unit ball of Euclidean space $\mathscr{E}^{n+1}$): *if $F = \overline{F} \subset \mathscr{S}_n$ and if the set $\mathscr{S}_n - F$ has $k$ components, then for every homeomorphism $h: F \to \mathscr{S}_n$ the set $\mathscr{S}_n - h(F)$ also has $k$ components.*

The proof of this theorem can be carried out making use of the concept of homology extended to arbitrary compact sets.†

As for polyhedra, one proves that the Betti numbers are topological invariants and that the $(n-1)$th Betti number of the closed set $F$ lying in $\mathscr{S}_n$ equals the number of components of the set $\mathscr{S}_n - F$ minus 1.

For sets lying in $\mathscr{S}_2$ the proof of the above theorem can be carried out considering the function space $\mathscr{P}^F$ to be a group. Namely, the group operation is defined as follows.

Let $f_1, f_2$ and $f_3$ be three elements of the space $\mathscr{P}^F$. We assume that $f_3 = f_1 \cdot f_2$ when $f_3(z) = f_1(z) \cdot f_2(z)$ for every $z \in F$.

The functions $f$ satisfying the condition $f \sim 1$ form a subgroup of the group $\mathscr{P}^F$, as can easily be verified. Let us denote it by $\boldsymbol{G}$ and let us consider the quotient group $\boldsymbol{B}(F) = \mathscr{P}^F / \boldsymbol{G}$.

The rank of this group (the maximal number of linearly independent elements) equals the number of components of the set $\mathscr{S}_2 - F$ less one.

Let us note finally that the proof of the invariance of the property

---

† Another proof was given by K. Borsuk. This proof requires an apparatus which goes significantly beyond the scope of this book. See *Fundamenta Mathematicae* **37** (1950), pp. 217–241, and my *Topology*, vol. II, p. 495.

of a closed subset $F$ of $\mathscr{S}_n$ of separating $\mathscr{S}_n$ can be carried out without the use of homology. For, the connectedness of $\mathscr{S}_n - F$ and of $\mathscr{S}_{n-1}^F$ are equivalent.†

**Exercises**

**1.** Prove that $z^n$ is not $\sim 1$ for $n \neq 0$ on the circumference of a circle with centre 0.

**2.** Prove that if $f \in \mathscr{P}^{\mathscr{S}_2}$ then $f \sim 1$.

*Hint*: Decompose $\mathscr{S}_2$ by the equator and apply Corollary 2 of § 5.

**3.** Prove that the star-shaped curve consisting of $n$ arcs having one end in common, and having no other points in common, does not decompose the plane.

**4.** Prove that a curve consisting of three arcs having common endpoints, and having no other points in common (see Fig. 19), decomposes the plane into three regions.

**5.** A connected space is said to be *unicoherent* if $A \cap B$ is connected for every decomposition of the space into two closed connected sets $A$ and $B$. Prove that the circular disk and the space $\mathscr{S}_2$ are unicoherent.

**6.** Prove that if $C$ is a subcontinuum of the plane $\mathscr{S}_2$ (or more generally, of a connected unicoherent space), and $R$ is a component of the complement of $C$, then $\mathrm{Fr}(R)$ is a continuum.

*Hint*: Use Theorem 4 of Chapter XVII, § 3.

**7.** Let the space $X$ be a locally connected unicoherent continuum. If the closed set $F$ separates this space between the points $a$ and $b$, then it contains a subcontinuum which also separates the space between these points.

*Hint*: Consider the component $R$ of the set $X-F$ which contains the point $a$ and the component $P$ of the set $X-R$ which contains the point $b$, and apply Exercise 4, above, and Exercise 9 of Chapter XVIII.

**8.** Under the preceding assumptions on the space $X$, let $A$ and $B$ be two disjoint closed sets neither one of which separates $X$ between $p$ and $q$. Prove that $A \cup B$ does not separate the space between $p$ and $q$ either.

**9.** Show by an example that without the unicoherence assumption the theorems of Exercises 6–8 do not hold.

**10.** Let the function $f \in \mathscr{S}_1^{\mathscr{S}_1}$ satisfy the condition $f(-z) = -f(z)$ for every $z \in \mathscr{S}_1$. Then the condition $f \sim 1$ is not satisfied.

**11.** *The Borsuk–Ulam theorem on antipodes.* For every function $f \in (\mathscr{E}^2)^{\mathscr{S}_2}$ there exists a point $z_0$ such that $f(z_0) = f(-z_0)$.

*Hint*: For every point $p$ belonging to the disk $\mathscr{K}_2$ with radius 1 and centre 0, let us denote by $p^+$ the point belonging to the "upper half" of $\mathscr{S}_2$, whose projection is $p$. Let $h(p) = f(p^+) - f(-p^+)$. Let us assume, contrary to the assertion of the theorem, that $h(p) \neq 0$ for every $p$. Show (making use of

† Borsuk theorem, see *Monatshefte für Mathematik und Physik* **38** (1931), p. 218, and *Mathematische Annalen* **106** (1932), p. 239. Cf. also P. Aleksandrov, Dimensionstheorie, § 5, *Mathematische Annalen* **106** (1932), p. 218 or my *Topology*, vol. II, p. 470.

Corollary 2, § 5, and of the remark immediately following it) that this assumption leads to a contradiction to the theorem of Exercise 10.

**12.** A region $R \subset \mathscr{S}_2$ is said to be *simply connected* if the set $\mathscr{S}_2 - R$ is connected.

Prove that if a simply connected region $R \subset \mathscr{S}_2$ contains a simple closed curve $C$, then it also contains one of the two components of its complement. In particular, if $R$ does not contain the point at infinity, then it contains a bounded component of the set $\mathscr{S}_2 - C$.

*Hint*: Note that the set $\mathscr{S}_2 - R$ is contained in one of the components of the set $\mathscr{S}_2 - C$.

Remark. The property of simply connected regions formulated in the above theorem is also a sufficient condition for simple connectedness, as can be proved.

**13.** Let $R$ be a simply connected region contained in $\mathscr{S}_2$, and let $L$ be an arc which, except for its endpoints, lies in $R$. Prove that the arc $L$ separates the region $R$ (i.e. that $R - L$ is not connected).

**14.** Prove the following more general theorem: let $R$ be an arbitrary region contained in $\mathscr{S}_2$, and let $L$ be an arc which, except for its endpoints, lies in $R$; a necessary and sufficient condition for this arc to separate the region $R$, is that both its endpoints belong to the same component of the set $\mathscr{S}_2 - R$.

*Hint*: In the proof of the necessity of the above condition make use of Theorem 6 of Chapter XVII, § 6, and of the first Janiszewski theorem. Make use of the second Janiszewski theorem in the proof of its sufficiency. Compare my *Topology*, vol. II, pp. 438 and 562.

**15.** If $C$ is a continuum contained in $\mathscr{S}_2$, then each of the components of the set $\mathscr{S}_2 - C$ is a simply connected region.

*Hint*: Cf. Theorem 4 of Chapter XVII, § 3.

SUPPLEMENT

# ELEMENTS OF ALGEBRAIC TOPOLOGY

## Introduction

This Supplement contains an introduction to homology theory, which constitutes one of two basic branches of algebraic topology (the other being homotopy theory). We continue the considerations of Chapter XX, where the concepts of simplex, simplicial subdivision, etc., were defined; unlike the other parts of this book, however, the Supplement makes use of algebraic concepts, notably those of group theory. This explains the name "algebraic" as opposed to "point set" topology.

We shall restrict the presentation to the definition and fundamental properties of homology groups of polyhedra; this, however, will suffice to introduce a number of important concepts and illustrate the role of algebraic tools. The interplay of various branches of mathematics which we encounter here is worth noting: topology is a powerful tool in classical analysis, which, in turn, is connected with technology and natural sciences via its applications, while topology itself uses methods of algebra and set theory.

When writing this Supplement we utilized Chapter XXI (complexes, chains, homologies) of the previous edition of this book; the text was enlarged and supplemented.

The reader who would like to extend his knowledge of algebraic topology is referred to the following list of most popular books on the subject. Some of these books are primarily of historical interest, but those by Spanier and by Hilton and Wylie merit particular attention. The author of this Supplement is particularly indebted to the latter.

P. S. Alexandrov, *Combinatorial Topology*, Graylock, Rochester, 1956 and 1957.

D. G. Bourgin, *Modern Algebraic Topology*, MacMillan, New York, 1963.

S. Eilenberg, and N. Steenrod, *Foundations of Algebraic Topology*, Princeton University Press, Princeton, 1952.

P. J. Hilton, and S. Wylie, *Homology Theory*, Cambridge University Press, Cambridge, 1960.

S. T. Hu, *Homotopy Theory*, Academic Press, New York, 1959.

—— *Homology Theory: A First Course in Algebraic Topology*, Holden-Day, San Francisco, 1966.

L. S. Pontryagin, *Foundations of Combinatorial Topology*, Graylock, Rochester, 1952.

H. Seifert and W. Threlfall, *Lehrbuch der Topologie*, Teubner, Leipzig, 1934, and Chelsea, New York, 1947.

A. H. Spanier, *Algebraic Topology*, McGraw-Hill, New York, 1966.

A. H. Wallace, *An Introduction to Algebraic Topology*, Pergamon Press, Oxford, 1957.

## § 1. Complexes. Polyhedra. Simplicial approximation

A *simplicial complex* (or shortly: a *complex*) will be defined as a finite family of simplexes (lying in a Euclidean space of a fixed dimension) containing all the faces of its simplexes, and such that the intersection of the closures of each two of its elements is either empty or equal to the closure of their common face. The *dimension* of a complex will be defined as the least upper bound of the dimensions of its simplexes.

The family of all the faces of an arbitrary simplex is an example of a simplicial complex; the family of all the triangles in Fig. 16—their sides and vertices constitute another example—is a two-dimensional complex.

THEOREM 1. *A family of simplexes lying in a Euclidean space of a fixed dimension, containing with each simplex all its faces, forms a simplicial complex if, and only if, its elements are disjoint.*

Proof. If simplexes $S$ and $S'$ belong to a complex $K$ and $S \cap S' \neq \emptyset$, then $S$ and $S'$ have a common face $S''$ such that $\bar{S} \cap \bar{S}' = \bar{S}''$. If $S''$ were a proper face of one of these simplexes, we would have either $\bar{S}'' \cap S = \emptyset$ or $\bar{S}'' \cap S' = \emptyset$ (see formula (4), Chapter XX, § 1), and $\bar{S}'' \cap (S \cap S') = \emptyset$, contrary to the condition $\bar{S} \cap \bar{S}' = \bar{S}''$ and $S \cap S' \neq \emptyset$. We have, therefore, $S = S'' = S'$.

In view of the quoted formula, the fact that the elements of the family in question are disjoint implies that the intersection of the closures of any pair of its elements is either empty or equal to the closure of their common face.

The union of all simplexes of a complex $\boldsymbol{K}$ will be called the *underlying polyhedron* of $\boldsymbol{K}$ and will be denoted by $|\boldsymbol{K}|$. It follows from the definition of a complex that if $\bar{S} \in \boldsymbol{K}$, then $\bar{S} \subset |\boldsymbol{K}|$, and

$$|\boldsymbol{K}| = \bigcup_{S \in K} \bar{S}; \tag{1}$$

consequently, every polyhedron is compact.

The representation of a polyhedron in the form of $|\boldsymbol{K}|$ is not unique. Each complex $\boldsymbol{K}$ for which $|\boldsymbol{K}|$ is a given polyhedron will be called a *triangulation* (or *simplicial subdivision*) of this polyhedron.

A subset of a complex $\boldsymbol{K}$ which itself is a complex (i.e. it contains all the faces of its simplexes) will be called a *subcomplex* of $\boldsymbol{K}$. An example of a subcomplex is provided by the set of all simplexes of a given complex which are of dimension $\leqslant n$; this subcomplex will be called the *n-dimensional skeleton* of the complex in question. The elements of the 0-dimensional skeleton are called *vertices* of the complex. We easily verify that the union and the intersection of the subcomplexes $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$ are subcomplexes and that

$$|\boldsymbol{K}_1 \cup \boldsymbol{K}_2| = |\boldsymbol{K}_1| \cup |\boldsymbol{K}_2| \quad \text{and} \quad |\boldsymbol{K}_1 \cap \boldsymbol{K}_2| = |\boldsymbol{K}_1| \cap |\boldsymbol{K}_2|.$$

Any subset $K_0$ of a polyhedron $K$ such that for a certain triangulation $\boldsymbol{K}$ of $K$ there exists a subcomplex $\boldsymbol{K}_0$ whose underlying polyhedron is $K_0$ will be called a *subpolyhedron*. Subpolyhedra are closed sets.

Let $\boldsymbol{K}$ be a given simplicial complex. For each $S \in \boldsymbol{K}$ the barycentric subdivision $\boldsymbol{K}(S)$ of the set $S$ is a simplicial complex. We shall show that the family

$$\boldsymbol{K}' = \bigcup_{S \in K} \boldsymbol{K}(S) \tag{2}$$

is also a simplicial complex.

Note first that all simplexes of the family $\boldsymbol{K}'$ lie in a Euclidean space, the same as that in which lie the elements of the complex $\boldsymbol{K}$. Next, if a simplex belongs to $\boldsymbol{K}'$, the same holds for all its faces, since each of the complexes $K(S)$ has this property. Let us now consider a pair $S_1'$, $S_2'$ of elements of $K'$ such that $S_1' \cap S_2' \neq \emptyset$. Let $S_i' \in K(S_i)$, where $S_i \in K$, $i = 1, 2$, and let $S_3$ be the common face of $S_1$ and $S_2$ such that $\overline{S}_1 \cap \overline{S}_2 = \overline{S}_3$. We have, therefore, $\emptyset \neq S_1' \cap S_2' \subset \overline{S}_3$ and $S_1' \cap \overline{S}_3 \neq \emptyset \neq S_2' \cap \overline{S}_3$. Since the barycentric subdivision of $\overline{S}$ determines the barycentric subdivision of the closure of every face of $S$, we have $S_1', S_2' \in K(S_3)$ and $S_1' = S_2'$. In view of Theorem 1 it follows that $K'$ is a simplicial complex.

The complex $K^{(1)} = K'$ will be called the (*first*) *barycentric subdivision* of the complex $K$. Further barycentric subdivisions are defined by the relation

$$K^{(n)} = (K^{(n-1)})' \quad \text{for} \quad n = 2, 3, \dots$$

It can easily be seen that if $L$ is a subcomplex of $K$, then $L^{(n)}$ is a subcomplex of $K^{(n)}$.

From (1) and (2) it follows that

$$|K'| = \bigcup_{S' \in K'} \overline{S}' = \bigcup_{S \in K} \bigcup_{S' \in K(S)} \overline{S}' = \bigcup_{S \in K} \overline{S} = |K|,$$

hence the underlying polyhedra of a complex $K$ and all its barycentric subdivisions coincide. Theorem 2, Chapter XX, § 2, implies:

THEOREM 2. *For any simplicial complex $K$ and any $\varepsilon > 0$ there exists a natural number $n$ such that the simplexes of the barycentric subdivision $K^{(n)}$ have diameters $< \varepsilon$.*

COROLLARY. *Every polyhedron has, for any $\varepsilon > 0$, a triangulation with simplexes of diameters $< \varepsilon$.*

A mapping $\varphi$ of the set of vertices of a complex $K$ into the set of vertices of a complex $L$ will be called a *simplicial map*, if for every simplex $p_0 \dots p_n \in K$ the points $\varphi(p_0), \dots, \varphi(p_n)$ are vertices of a certain simplex in $L$; this simplex can have dimension smaller than $n$, since different vertices of $K$ can be mapped into the same vertex of $L$. By assigning to each simplex $p_0 \dots p_n \in K$ the simplex in $L$ whose set of vertices is $\{\varphi(p_0), \dots, \varphi(p_n)\}$ we define a mapping of the complex $K$ into the complex $L$; this mapping is an extension of $\varphi$ and will be denoted by the same letter $\varphi$. The term "simplicial map" will be used both for the mapping of the set of vertices, as well as for the corresponding mapping of the complex. We easily verify that the composition of simplicial maps is again a simplicial map.

Simplicial maps $\varphi, \psi: \boldsymbol{K} \to \boldsymbol{L}$ will be called *contiguous* if for every $S \in \boldsymbol{K}$ there exists a simplex $T(S) \in \boldsymbol{L}$ such that $\varphi(S)$ and $\psi(S)$ are faces of $T(S)$.

Let $\boldsymbol{K}$ be an arbitrary simplicial complex. Let us assign to every vertex $b(S)$ of the complex $\boldsymbol{K}'$ one of the vertices of the simplex $S \in \boldsymbol{K}$. If $b(S_0) \dots b(S_n)$ is a simplex in $\boldsymbol{K}'$, we may assume that $S_0 > \dots > S_n$, which implies that all the vertices of the simplexes $S_0, \dots, S_n$ are vertices of $S_0$, and, consequently, our mapping is a simplicial map. Mappings obtained in this manner will be called *standard translations* of $\boldsymbol{K}'$ in $\boldsymbol{K}$. Every two standard translations are contiguous. Iterations of standard translations will be called *standard maps*.

The set

$$\mathrm{St}_K(p) = \bigcup_{p<S\in K} S \subset |\boldsymbol{K}|$$

will be called the *star* of the vertex $p \in \boldsymbol{K}$. Since the family $\{S \in \boldsymbol{K}\colon (p < S)'\}$ is a subcomplex of $\boldsymbol{K}$, the complement of a star is a subpolyhedron of $|\boldsymbol{K}|$, and stars are open sets in $|\boldsymbol{K}|$.

Note that

(3) *if* $p < S \in \boldsymbol{K}$, *then* $\mathrm{St}_{K'}(b(S)) \subset \mathrm{St}_K(p)$.

Indeed, if a simplex $b(S_0) \dots b(S_n) \in \boldsymbol{K}'$, where $S_0 > \dots > S_n$ contains $b(S)$ among its vertices, then $p < S_0$ and $b(S_0) \dots b(S_n) \subset S_0 \subset \mathrm{St}_K(p)$.

THEOREM 3. *The intersection of stars* $\mathrm{St}_K(p_0) \cap \dots \cap \mathrm{St}_K(p_n)$, *where* $p_i \neq p_j$ *for* $i \neq j$ *is non-empty if, and only if,* $p_0 \dots p_n \in \boldsymbol{K}$.

Proof. If $\mathrm{St}_K(p_0) \cap \dots \cap \mathrm{St}_K(p_n) \neq \emptyset$, then there exist simplexes $S_0, \dots, S_n \in \boldsymbol{K}$ such that $p_i < S_i$ for $i = 0, 1, \dots, n$ and a point $p \in S_0 \cap \dots \cap S_n$. Since the simplexes of the complex $\boldsymbol{K}$ are pairwise disjoint, we have $S_0 = S_1 = \dots = S_n = S$ and $p_i < S$ for $i = 0, 1, \dots, n$. Thus, $p_0 \dots p_n$ is a face of the simplex $S$ and, consequently, is an element of $\boldsymbol{K}$.

If $p_0 \dots p_n \in \boldsymbol{K}$, then $p_0 \dots p_n \subset \mathrm{St}_K(p_i)$ for $i = 0, \dots, n$, and $\mathrm{St}_K(p_0) \cap \dots \cap \mathrm{St}_K(p_n) \neq \emptyset$.

Let $f$ be a continuous mapping of a polyhedron $K$ into a polyhedron $L$. We shall define a *simplicial approximation* of $f$ as a triple $(\boldsymbol{K}, \boldsymbol{L}, \varphi)$ where $\boldsymbol{K}$ and $\boldsymbol{L}$ are triangulations of $K$ and $L$ and $\varphi$ is a function mapping the set of vertices of $\boldsymbol{K}$ into the set of vertices of $\boldsymbol{L}$, such that for every $p \in \boldsymbol{K}$

$$\mathrm{St}_{\boldsymbol{K}}(p) \subset f^{-1}\big(\mathrm{St}_{\boldsymbol{L}}(\varphi(p))\big).$$

Applying Theorem 3 twice we show that for any simplicial approximation $(\boldsymbol{K}, \boldsymbol{L}, \varphi)$ of a continuous mapping $f$, the function $\varphi$ is a simplicial map; we will call it also a simplicial approximation of $f$.

THEOREM 4. *If* $(\boldsymbol{K}, \boldsymbol{L}, \varphi)$ *and* $(\boldsymbol{K}, \boldsymbol{L}, \psi)$ *are simplicial approximations of a mapping* $f\colon K \to L$, *then* $\varphi$ *and* $\psi$ *are contiguous.*

P r o o f. For every simplex $S = p_0 \dots p_n \in \boldsymbol{K}$ we have for $i = 0, \dots, n$:

$$\begin{aligned}\mathrm{St}_{\boldsymbol{K}}(p_i) &\subset f^{-1}\big(\mathrm{St}_{\boldsymbol{L}}(\varphi(p_i))\big) \cap f^{-1}\big(\mathrm{St}_{\boldsymbol{L}}(\psi(p_i))\big)\\ &= f^{-1}\big(\mathrm{St}_{\boldsymbol{L}}(\varphi(p_i)) \cap \mathrm{St}_{\boldsymbol{L}}(\psi(p_i))\big).\end{aligned}$$

Consequently,

$$\emptyset \neq \bigcap_{i=0}^{n} \mathrm{St}_{\boldsymbol{K}}(p_i) \subset f^{-1}\big(\bigcap_{i=0}^{n} \mathrm{St}_{\boldsymbol{L}}(\varphi(p_i)) \cap \mathrm{St}_{\boldsymbol{L}}(\psi(p_i))\big)$$

and

$$\mathrm{St}_{\boldsymbol{L}}(\varphi(p_0)) \cap \mathrm{St}_{\boldsymbol{L}}(\psi(p_0)) \cap \dots \cap \mathrm{St}_{\boldsymbol{L}}(\varphi(p_n)) \cap \mathrm{St}_{\boldsymbol{L}}(\psi(p_n)) \neq \emptyset.$$

As $T(S) \in \boldsymbol{L}$ we may take the simplex with the set of vertices

$$\{\varphi(p_0), \psi(p_0), \dots, \varphi(p_n), \psi(p_n)\}.$$

THEOREM 5. *If* $(\boldsymbol{K}, \boldsymbol{L}, \varphi)$ *is a simplicial approximation of a mapping* $f$, *and* $(\boldsymbol{L}, \boldsymbol{M}, \psi)$ *is a simplicial approximation of a mapping* $g$, *then* $(\boldsymbol{K}, \boldsymbol{M}, \psi\varphi)$ *is a simplicial approximation of the composition* $gf$.

P r o o f. For an arbitrary vertex $p \in \boldsymbol{K}$ we have

$$\mathrm{St}_{\boldsymbol{K}}(p) \subset f^{-1}\big(\mathrm{St}_{\boldsymbol{L}}(\varphi(p))\big)$$

and

$$\mathrm{St}_{\boldsymbol{L}}(\varphi(p)) \subset g^{-1}\big(\mathrm{St}_{\boldsymbol{M}}(\psi\varphi(p))\big).$$

Therefore

$$\mathrm{St}_K(p) \subset f^{-1}g^{-1}\big(\mathrm{St}_M(\psi\varphi(p))\big) = (gf)^{-1}\big(\mathrm{St}_M(\psi\varphi(p))\big).$$

It follows from (3) that standard translations are simplicial approximations of the identity. This fact, together with Theorem 5, implies the following two theorems:

THEOREM 6. *If* $(\boldsymbol{K}, \boldsymbol{L}', \varphi)$ *is a simplicial approximation of a mapping* $f$, *and* $\sigma$ *is a standard translation of* $\boldsymbol{L}'$ *in* $\boldsymbol{L}$, *then* $(\boldsymbol{K}, \boldsymbol{L}, \sigma\varphi)$ *is also a simplicial approximation of* $f$.

THEOREM 7. *If* $(\boldsymbol{K}, \boldsymbol{L}, \varphi)$ *is a simplicial approximation of a mapping* $f$ *and* $\sigma$ *is a standard translation of* $\boldsymbol{K}'$ *in* $\boldsymbol{K}$, *then* $(\boldsymbol{K}', \boldsymbol{L}, \varphi\sigma)$ *is also a simplicial approximation of* $f$.

Theorems 4 and 7 imply:

THEOREM 8. *If* $(\boldsymbol{K}, \boldsymbol{L}, \varphi)$ *and* $(\boldsymbol{K}^{(m)}, \boldsymbol{L}, \psi)$ *are simplicial approximations of the same mapping, and* $\sigma\colon \boldsymbol{K}^{(m)} \to \boldsymbol{K}$ *is a standard map, then* $\varphi\sigma$ *and* $\psi$ *are contiguous.*

THEOREM 9. *For every polyhedron* $L = |\boldsymbol{L}|$ *there exists an* $\varepsilon > 0$ *such that if* $f$ *and* $g$ *are continuous mappings of an arbitrary polyhedron* $K = |\boldsymbol{K}|$ *into* $L$, *and* $|f-g| < \varepsilon$, *then* $f$ *and* $g$ *have a common simplicial approximation* $(\boldsymbol{K}^{(m)}, \boldsymbol{L}, \varphi)$ *for a certain* $m \geq 1$.

Proof. The family of sets $\{\mathrm{St}_L(q)\}$, where $q$ is a vertex of $\boldsymbol{L}$, forms an open cover of the space $L$; let $\varepsilon$ be the Lebesgue coefficient of this cover (see Theorem 7, Chapter XVI, § 5). If $|f-g| < \varepsilon$, then the family of sets of the form $f^{-1}(\mathrm{St}_L(q)) \cap g^{-1}(\mathrm{St}_L(q))$, where $q$ is a vertex of $\boldsymbol{L}$, forms an open cover of $K$. Indeed, for every $x \in K$ there exists a vertex $q \in \boldsymbol{L}$ such that $\{f(x), g(x)\} \subset \mathrm{St}_L(q)$, i.e. $x \in f^{-1}(\mathrm{St}_L(q)) \cap g^{-1}(\mathrm{St}_L(q))$. By Theorem 2 and the theorem on the Lebesgue coefficient there exists a natural number $m$ such that for every vertex $p$ of the complex $\boldsymbol{K}^{(m)}$ one can find a vertex $\varphi(p) \in \boldsymbol{L}$ such that

$$\mathrm{St}_{K^{(m)}}(p) \subset f^{-1}\big(\mathrm{St}_L(\varphi(p))\big) \cap g^{-1}\big(\mathrm{St}_L(\varphi(p))\big).$$

The mapping $\varphi$ is a simplicial approximation of $f$ and $g$.

COROLLARY. *Every continuous mapping of a polyhedron into a polyhedron has a simplicial approximation.*

## § 2. Abelian groups

In this section we shall present those notions and theorems of the theory of groups which will be used in the construction of homology groups of simplicial complexes.

Definition 1. Let $G$ be a given set, and suppose that a binary operation (to be called *addition*) is defined, which to every pair $a, b$ of elements of $G$ assigns an element $a+b$ of $G$ (to be called the *sum* of $a$ and $b$). We say that $G$ is a *commutative* (or *Abelian*) *group* if the following conditions (axioms of group theory) are satisfied:

(i) $(a+b)+c = a+(b+c)$,

(ii) $a+b = b+a$,

(iii) *there exists a unique element of $G$ (denoted by $0$) such that $a+0 = a$ for every $a$ in $G$,*

(iv) *for every $a$ in $G$ there exists a unique inverse element (denoted by $-a$) such that $a+(-a) = 0$.*

Given $a \in G$ and a natural number $m$ we denote by $ma$ the element of $G$ obtained as the result of $m$-fold addition of $a$ (in view of (i) this element is well defined). Moreover, we put $0a = 0$ and $(-m)a = m(-a)$. We easily verify that

$$ma+na = (m+n)a$$

for any integers $m$ and $n$.

EXAMPLE 1. The set of all integers forms a group with respect to the operation of addition, but does not form a group with respect to the operation of multiplication, since axiom (iv) does not hold in the latter case.

EXAMPLE 2. The set of complex numbers $z$ such that $|z| = 1$ (i.e. numbers of the form $e^{ix}$) forms a commutative group with respect to multiplication.

EXAMPLE 3. The set of continuous non-vanishing complex-valued functions $f$ defined on a space $X$ forms a commutative group if the group operation is defined as

$$(f_3 = f_1 \cdot f_2) \equiv \bigwedge_x [f_3(x) = f_1(x) \cdot f_2(x)].$$

Definition 2. If a subset $G_0$ of $G$ forms a group with respect to the operation in $G$, i.e. if $0 \in G_0$ and $a, b \in G_0$ implies $(a+b) \in G_0$, $(-a) \in G_0$, we say that $G_0$ is a *subgroup* of $G$.

The equivalence relation mod $G_0$ for elements of the group $G$ is defined as

$$(4) \qquad [a \sim b \ (\mathrm{mod}\ G_0)] \equiv [(a-b) \in G_0].$$

THEOREM 1. *Relation* (4) *is an equivalence relation, i.e. it is reflexive, symmetric and transitive.*

Proof. $a \sim a \ (\mathrm{mod}\ G_0)$, i.e. $a-a = 0 \in G_0$ (since $G_0$ is a subgroup of $G$). Next,

$$[a \sim b \ (\mathrm{mod}\ G_0)] \equiv [(a-b) \in G_0] \equiv [(b-a) \in G_0] \equiv [b \sim a \ (\mathrm{mod}\ G_0)].$$

Let $a \sim b \ (\mathrm{mod}\ G_0)$ and $b \sim c \ (\mathrm{mod}\ G_0)$, i.e. $(a-b) \in G_0$ and $(b-c) \in G_0$. Then $(a-b)+(b-c) \in G_0$, whence $a \sim c \ (\mathrm{mod}\ G_0)$.

Relation (4) induces a partition of the elements of $G$ into disjoint sets of equivalent elements. These sets are called *cosets* mod $G_0$; thus, cosets mod $G_0$ are equivalence sets of the relation $a \sim b \ (\mathrm{mod}\ G_0)$ (see Exercise 19, Chapter IV).

Let $p(a)$ denote the set of elements equivalent to $a$ mod $G_0$, i.e. the coset mod $G_0$ containing $a$. We have

$$[p(a) = p(b)] \equiv [a \sim b \ (\mathrm{mod}\ G_0)].$$

In the class of cosets mod $G_0$ we define addition by the relation

$$(5) \qquad p(a)+p(b) = p(a+b).$$

We easily prove that addition of cosets as defined by (5) does not depend on the choice of elements of these cosets, and that the class of cosets with addition as defined by (5) forms a commutative group.

Definition 3. The group of cosets mod $G_0$ will be called the *quotient group* of $G$ with respect to $G_0$, and will be denoted by $G/G_0$. The zero of the group $G/G_0$ is $G_0$.

Definition 4. Let $G$ and $H$ be two commutative groups. A transformation $f\colon G \to H$ will be called a *homomorphism* if

$$(6) \qquad f(a+b) = f(a)+f(b).$$

If, in addition, $f$ is one-to-one, we call $f$ a *monomorphism*. If $f(G) = H$, we say that $f$ is an *epimorphism*. Transformations which are at the same time monomorphisms and epimorphisms are called *isomorphisms*.

Compositions of homomorphisms (monomorphisms, epimorphisms, isomorphisms) are homomorphisms (monomorphisms, epimorphisms, isomorphisms). A homomorphism $f$ is an isomorphism if, and only if, it has an inverse; then the inverse is also an isomorphism.

EXAMPLE 4. The transformation mapping a group $G$ into the zero of a group $H$ is obviously a homomorphism; we call it the *zero-homomorphism* of $G$ into $H$ and denote it by 0. The transformation of the group of integers into the group of complex numbers of the form $e^{ix}$ defined by assigning the complex number $e^{im}$ to the integer $m$ is a monomorphism. For any group $G$ the *identity*, i.e. the transformation defined by the relation $f(a) = a$ is an isomorphism of $G$ onto itself; this isomorphism will be denoted by $1_G$.

If there exists an isomorphism of $G$ onto $H$, we say that these groups are *isomorphic*, and write

$$G \approx H.$$

Clearly, the relation of isomorphism is reflexive, symmetric and transitive, i.e. it is an equivalence relation.

In much the same way as topology deals with invariants of homeomorphisms, the theory of groups deals with invariants of isomorphisms. From the point of view of the theory of groups, two isomorphic groups have the same properties. The role of continuous functions in topology is played in the theory of groups by homomorphisms.

THEOREM 2. *If $f$ is a homomorphism of a group $G$ into a group $H$, then*

1°. $f(0) = 0$,

2°. $f(-a) = -f(a)$,

3°. *the image* $\mathrm{Im} f = f(G)$ *is a subgroup of* $H$

*(on the left-hand side of* 1° *the symbol* 0 *denotes the zero element of the group $G$ and on the right-hand side that of the group $H$).*

Indeed, $f(0) = f(0+0) = f(0)+f(0)$, whence $f(0) = 0$. Thus, $0 = f[a+(-a)] = f(a)+f(-a)$, and $f(-a) = -f(a)$. Relation 3° follows from (6) and 2°.

THEOREM 3. *For any subgroup $G_0$ of a group $G$, the function mapping $a \in G$ into $p(a) \in G/G_0$ is an epimorphism.*

We shall call it the *natural epimorphism* $p\colon G \to G/G_0$.

Definition 5. The *kernel of a homeomorphism* $f\colon G \to H$ is defined as the set

$$\operatorname{Ker} f = \{x\colon f(x) = 0\}.$$

THEOREM 4. *If $f$ is a homeomorphism of a group $G$ into a group $H$, then*

1°. $\operatorname{Ker} f$ *is a subgroup of* $G$,

2°. $f$ *is a monomorphism if, and only if,* $\operatorname{Ker} f = 0$,

3°. *the groups* $G/\operatorname{Ker} f$ *and* $\operatorname{Im} f$ *are isomorphic.*

Proof. If $f(a) = 0$ and $f(b) = 0$, then $f(a+b) = 0$; if $f(a) = 0$, then $f(-a) = -f(a) = 0$, which proves that the kernel is a subgroup.

If $f$ is a monomorphism, then $\operatorname{Ker} f = 0$. If $\operatorname{Ker} f = 0$, and $f(x) = f(y)$, then $f(x-y) = f(x)-f(y) = 0$ and $x = y$.

The isomorphism between the groups $G/\operatorname{Ker} f$ and $\operatorname{Im} f$ can be obtained by assigning to the element $p(a) \in G/\operatorname{Ker} f$ the element $f(a) \in H$ (note that if $p(a) = p(a')$, then $f(a) = f(a')$).

Suppose now that we are given a homeomorphism $f\colon G \to H$ and subgroups $G_0 \subset G$ and $H_0 \subset H$ such that $f(G_0) \subset H_0$. If $a \sim b \pmod{G_0}$, then $f(a) \sim f(b) \pmod{H_0}$, since $f(a)-f(b) = f(a-b) \in f(G_0) \subset H_0$. The image under $f$ of every equivalence set of the relation $a \sim b \pmod{G_0}$ is thus contained in a certain equivalence set of the relation of equivalence mod $H_0$. If we assign to cosets mod $G_0$ the cosets mod $H_0$ containing their images under $f$, we define in fact the transformation $f'\colon G/G_0 \to H/H_0$, such that $pf = f'p$, i.e. that the diagram

$$\begin{array}{ccc} G & \xrightarrow{\;f\;} & H \\ {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle p} \\ G/G_0 & \xrightarrow{\;f'\;} & H/H_0 \end{array}$$

is commutative.

Since

$$f'(p(a)+p(b)) = f'p(a+b) = pf(a+b) = pf(a)+pf(b)$$
$$= f'(p(a))+f'(p(b)),$$

the transformation $f'$ is a homomorphism, called the *homomorphism induced* by $f$. We note that if $f$ is an isomorphism, and $f(G_0) = H_0$, then $f'$ is an isomorphism of the group $G/G_0$ onto $H/H_0$.

If $f, g\colon G \to H$ are homomorphisms, the functions $f+g\colon G \to H$ and $f-g\colon G \to H$ defined as

$$(7) \qquad (f+g)(a) = f(a)+g(a), \quad (f-g)(a) = f(a)-g(a)$$

are also homomorphisms, to be called the *sum* and the *difference* of the homomorphisms $f$ and $g$ respectively.

Definition 6. The Cartesian product $G_1 \times \ldots \times G_n$ of groups $G_1, \ldots, G_n$ with the addition defined as

$$(a_1, \ldots, a_n)+(b_1, \ldots, b_n) = (a_1+b_1, \ldots, a_n+b_n)$$

will be called the *direct sum* of the groups $G_1, \ldots, G_n$.

We can easily see that the operation defined above satisfies conditions (i)–(iv); the zero element is $(0, \ldots, 0)$, the inverse of $(a_1, \ldots, a_n)$ is $(-a_1, \ldots, -a_n)$.

Definition 7. We shall define the *free group* generated by a non-empty set $A$ as the set of all integer-valued functions defined on $A$ each of which assume non-zero values only for a finite number of elements of $A$, with the addition of functions (satisfying (i)–(iv)) defined by the first of formulas (7). Elements of $A$ will be called *generators* of the free group. In addition, we assume that the empty set generates the zero group, i.e. the group containing only the element 0.

If we introduce for every $a \in A$ a function $\hat{a}$ defined as

$$\hat{a}(x) = \begin{cases} 1 & \text{for} \quad x = a, \\ 0 & \text{for} \quad x \neq a \end{cases}$$

we can represent elements of the free group generated by $A$ in a unique way as

$$m_1\hat{a}_1 + \ldots + m_n\hat{a}_n, \quad \text{where} \quad a_1, \ldots, a_n \in A,$$

addition of elements being performed by adding coefficients at the corresponding functions.

In the sequel we shall indentify the function $\hat{a}$ with the generator $a$, and the elements of the free group generated by $A$ will be identified with formal linear combinations of the form

$$m_1 a_1 + \ldots + m_n a_n \quad \text{where} \quad a_1, \ldots, a_n \in A.$$

Addition of such linear combinations consists in adding coefficients at the correponding generators. Thus, the free group contains its generators. We can easily see that any free group with $n$ generators is isomorphic with the direct sum of $n$ copies of the group of integers.

THEOREM 5. *Every mapping of the set of generators of a free group into any group $H$ can be uniquely extended to a homomorphism of this group into the group $H$.*

Proof. Let $A$ be the set of generators of a free group $G$, and let $f$ be a mapping of $A$ into a group $H$. The required extension of $f$ is given by the homomorphism $\bar{f}\colon G \to H$ defined as

$$\bar{f}(m_1 a_1 + \ldots + m_n a_n) = m_1 f(a_1) + \ldots + m_n f(a_n).$$

## § 3. Categories and functors

In mathematics we often study a certain class of objects—which are generally sets satisfying certain additional conditions, such as the performability of certain operations in them under definite axioms—and a suitably chosen class of transformations of those objects, for instance transformations preserving the operations in question. We encounter such situations in the set theory (objects—all sets, transformations—all functions), in topology (objects—topological spaces, transformations—continuous functions) or in group theory (objects—groups, transformations—homomorphisms). The concept of category supplies an abstract description of such a situation.

Definition 1. A *category* will be defined as a class $\mathfrak{C}$ (see Chapter III, § 7)—whose elements will be called *objects*—given together with

(i) a set $M(A, B)$ assigned to each pair of objects $A, B \in \mathfrak{C}$; elements of $M(A, B)$ will be called *morphisms* of $A$ into $B$;

(ii) the operation of *composition* of morphisms, which assigns to each triple $A, B, C \in \mathfrak{C}$ a mapping of the Cartesian product

$M(A, B)\times M(B, C)$ into $M(A, C)$. The composition of morphisms $f\in M(A, B)$ and $g\in M(B, C)$ will be denoted by $g\circ f$ (we shall also write $gf$); the morphism $g\circ f$ belongs to $M(A, C)$.

We shall assume that

(iii) *if* $f\in M(A, B)$, $g\in M(B, C)$ *and* $h\in M(C, D)$, *then*

$$h\circ(g\circ f) = (h\circ g)\circ f;$$

(iv) *for each object* $A\in\mathfrak{C}$ *there exists a morphism* $e_A\in M(A, A)$, *called the identity morphism, such that*

$$f\circ e_A = f \text{ for } f\in M(A, B) \text{ and } e_A\circ f = f \text{ for } f\in M(B, A).$$

EXAMPLE 1. Taking as $\mathfrak{C}$ the class of all sets, and as $M(A, B)$ the class of all functions $f\colon A\to B$, we define the *category of all sets*. The composition of morphisms is here the usual composition of functions, and the identity morphism is $e_A(x) = x$.

EXAMPLE 2. The *category of topological spaces* is defined by taking as $\mathfrak{C}$ the class of all topological spaces, and as $M(A, B)$ the class of continuous functions $f\colon A\to B$. This category will be denoted by **T**. Restricting the class $\mathfrak{C}$ to more special topological spaces we obtain the *category* **M** *of metric spaces*, the *category* **H** *of Hausdorff spaces*, the *category* **C** *of compact spaces*, etc.

The *category* **P** *of polyhedra* is defined by taking as $\mathfrak{C}$ the class of all polyhedra, and their continuous mappings as morphisms.

EXAMPLE 3. The objects of the *category* **SC** *of simplicial complexes* are simplicial complexes, and the morphisms of this category are simplicial maps of complexes.

EXAMPLE 4. We define the *category* **G** *of Abelian groups* taking as $\mathfrak{C}$ the class of all Abelian groups, and as $M(A, B)$ the set of all homomorphisms of $A$ into $B$. The composition of morphisms in **G** is defined as the usual composition of functions, and the identity morphism of a group $A$ is the isomorphism $1_A$.

The basic concept of category theory is that of functor.

Definition 2. A function which to each object $A$ of a category **K** assigns an object $F(A)$ of a category **L**, and to each morphism $f\in M(A, B)$, where $A, B$ are objects of the category **K**, assigns a morphism $F(f)\in M(F(A), F(B))$ will be called a *covariant functor* if

(v) $F(e_A) = e_{F(A)}$,

(vi) $F(g \circ f) = F(g) \circ F(f)$.

If $F(f) \in M(F(B), F(A))$ for $f \in M(A, B)$ and condition (vi) is replaced by

(vi') $F(g \circ f) = F(f) \circ F(g)$

we shall speak of a *contravariant functor.*

Functors can be composed: the composition of two covariant or two contravariant functors is a covariant functor, while the composition of a covariant functor with a contravariant one is a contravariant functor.

EXAMPLE 5. Let $T$ be a fixed set. Let us assign to each set $X$ the Cartesian product $X \times T$, and to a function $f\colon X \to Y$ the function $f'\colon X \times T \to Y \times T$ defined as $f'(x, t) = (f(x), t)$. We easily show (see Exercises 17 and 23, Chapter IV) that such assigning is a covariant functor from the category of all sets into itself. From Theorem 4, Chapter XIII, § 2, it follows that if $T$ is a topological space, then the mapping defined above is a functor from the category of topological spaces into itself.

EXAMPLE 6. Let $T$ be a fixed set. Let us assign to each set $X$ the set of all functions defined on $X$ with values in $T$, i.e. the set $T^X$. To a function $f\colon X \to Y$, let us assign the function $d_f\colon T^Y \to T^X$ defined as $d_f(g) = gf$, where $g \in T^Y$. The reader will verify (see Exercise 22, Chapter IV) that such assigning is a contravariant functor from the category of all sets into itself. Let us now assume that $T$ is a topological space. For each topological space $X$, consider the compact-open topology in the set $T^X$ (see Chapter XVI, § 7). We shall show that the above assigning is a functor from the category $\mathbf{T}$ of topological spaces into itself.

It suffices to verify that, for any continuous mapping $f\colon X \to Y$, the mapping $d_f\colon T^Y \to T^X$ is also continuous. This, however, follows from Corollary 3 to Theorem 2, Chapter XII, § 1, since for any set $\Gamma(C, H)$ from the subbase of the space $T^X$ ($C$ is a compact subset of $X$ and $H$ is an open set in $T$), we have

$$d_f^{-1}(\Gamma(C, H)) = \{g\colon gf(C) \subset H\} = \Gamma(f(C), H)$$

and $f(C) \subset Y$ is compact by Theorem 3, Chapter XVI, § 2.

EXAMPLE 7. The Čech–Stone compactification is an important example of a covariant functor from the category of all completely regular $\mathscr{T}_1$-spaces to the category of compact spaces. This functor assigns to each space $X$ its compactification $\beta X$, and to each mapping $f\colon X \to Y \subset \beta Y$ its extension $g\colon \beta X \to \beta Y$; the latter exists in view of the theorem from Chapter XVI, § 4.

## § 4. Homology groups of simplicial complexes

We shall now describe a sequence of covariant functors (homology groups and induced homomorphisms) from the category **SC** of complexes and simplicial maps to the category **G** of Abelian groups.

The first step in the construction of homology groups will consist in the orientation of a simplicial complex. We start with the definition of an oriented simplex. Let $S = p_0 \dots p_n$ be an $n$-dimensional simplex, where $n \geqslant 1$. Each sequence $p_{j_0}, \dots, p_{j_n}$ consisting of $n+1$ of its vertices (without repetition) will be called a *n-dimensional oriented simplex*; we shall identify two oriented simplexes if one of them can be obtained from the other by an even permutation (we say in this case that these simplexes have the *same* orientation); for instance

$$(p_0, p_1, p_2) = (p_1, p_2, p_0) = (p_2, p_0, p_1),$$

$$(p_0, p_1, p_2) \neq (p_1, p_0, p_2).$$

Thus, each simplex $S$ of dimension $\geqslant 1$ determines two oriented simplexes. The simplex with an orientation *opposite* to that of the simplex $(p_0, \dots, p_n)$ will be denoted by $-(p_0, \dots, p_n)$.

We shall say that the oriented $(n-1)$-dimensional face $(p_1, \dots, p_n)$ of the oriented simplex $(p_0, p_1, \dots, p_n)$ is *coherently* oriented with this simplex. For instance, $(p_0, p_1)$ has coherent orientation with the simplex $(p_2, p_0, p_1) = (p_0, p_1, p_2)$, while $(p_0, p_2)$ is not coherently oriented with $(p_0, p_1, p_2)$ (but coherently oriented with $(p_1, p_0, p_2)$).

Thus, given an oriented $n$-dimensional simplex and its oriented $(n-1)$-dimensional face $(n \geqslant 2)$, we can decide whether or not it is coherently oriented with this simplex. We easily verify that the orientation of the face resulting from omitting the vertex $p_{j_i}$, coherent with the orientation of simplex $(p_{j_0}, \ldots, p_{j_n})$ is

$$(-1)^i(p_{j_0}, \ldots, p_{j_{i-1}}, p_{j_{i+1}}, \ldots, p_{j_n}).$$

Moreover, that definition of coherent orientation of a face does not depend on the choice of the sequence of vertices which gives the considered oriented $n$-dimensional simplex in question.

Remark. The notions of oriented simplex and coherent orientation have a geometrical interpretation. The $n$-dimensional simplex $S = p_0 \ldots p_n$ situated in the space $\mathcal{E}^m$ determines in this space the $n$-dimensional hyperplane $C(S)$, called the carrier of $S$, defined as the set of all points of the form

$$p = \lambda_0 p_0 + \ldots + \lambda_n p_n,$$

where $\lambda_0 + \ldots + \lambda_n = 1$. Formally, an oriented simplex is a pair consisting of $S$ and one of two possible orientations of its carrier, i.e. one of the two classes of bases of the vector space $C(S)$. Each sequence $(p_{j_0}, \ldots, p_{j_n})$ of vertices of the simplex $S$ determines a certain base of the space $C(S)$, namely the base consisting of vectors $\overrightarrow{p_{j_0}, p_{j_n}}, \overrightarrow{p_{j_1}, p_{j_n}}, \ldots, \overrightarrow{p_{j_{n-1}}, p_{j_n}}$. One can verify that sequences of vertices which give the same orientation of the simplex $S$ lead to the same orientation of the space $C(S)$, while sequences giving different orientations lead to different orientations of $C(S)$. Thus, instead of considering pairs consisting of a simplex and a basis of its carrier, one can consider sequences of vertices of the simplex, identifying those sequences which lead to the same orientation of the carrier.

If we are given an $n$-dimensional simplex $S$ ($n \geqslant 2$) and its $(n-1)$-dimensional face $S'$, then from each base of $C(S')$ one can obtain a base of $C(S)$ by adding one vector from $C(S)$ which is not parallel to $C(S')$. Our definition of the coherent orientations is equivalent to the agreement that if we start from an arbitrary base which determines a given orientation of $C(S')$ and add to it at the first place an arbitrary vector which pierces the face $S'$ from inside $S$, we obtain a base of the space $C(S)$ which gives the desired orientation of this space. This is illustrated in Fig. 20.

An *orientation* of a simplicial complex $K$ will be defined as a function $\alpha$ which to every simplex $S \in K$ of dimension $n \geqslant 1$

assigns one of the oriented simplexes determined by it. An *oriented* (*simplicial*) complex will be defined as a pair $(\boldsymbol{K}, \alpha)$ where $\boldsymbol{K}$ is the simplicial complex and $\alpha$ is its orientation. For the sake of

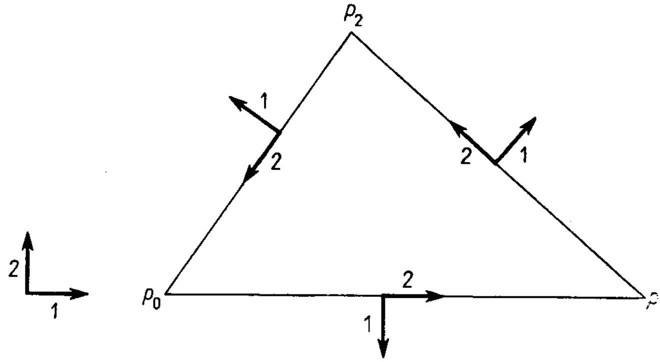



FIG. 20

unified notation we shall assume that $\alpha(p) = (p) = p$ for every vertex $p$.

For an arbitrary oriented complex $(\boldsymbol{K}, \alpha)$ and an integer $n \geqslant 0$ we shall denote by $C_n(\boldsymbol{K}, \alpha)$ the free group generated by the set of all oriented simplexes $\alpha(S)$, where $S$ is an $n$-dimensional simplex of $\boldsymbol{K}$. This group will be called the *group of n-dimensional chains* of the complex $(\boldsymbol{K}, \alpha)$. According to our convention concerning the representation of elements of the free group as linear combinations of generators, elements of $C_n(\boldsymbol{K}, \alpha)$ i.e. $n$-dimensional chains, have the form

$$k_1\alpha(S_1) + \ldots + k_m\alpha(S_m),$$

where $S_1, \ldots, S_m$ are $n$-dimensional simplexes of the complex $\boldsymbol{K}$. In particular, zero-dimensional chains are linear combinations

$$k_1 p_1 + \ldots + k_m p_m$$

of vertices of the complex $\boldsymbol{K}$, while $C_n(\boldsymbol{K}, \alpha) = 0$ for $n$ exceeding the dimension of the complex $\boldsymbol{K}$.

For each $n$-dimensional simplex $S = p_0 \ldots p_n \in \boldsymbol{K}$, where $n \geqslant 1$, the oriented simplex $\alpha(S)$ is an element of the group $C_n(\boldsymbol{K}, \alpha)$. Thus, the group $C_n(\boldsymbol{K}, \alpha)$ contains some sequences of the form $(p_{j_0}, \ldots, p_{j_n})$, namely those which determine the orientation $\alpha(S)$; each of these sequences is the same element of the group. For reasons of computational convenience, we assume that sequences giving the orientation opposite to $\alpha(S)$ also belong to $C_n(K, \alpha)$; each of these sequences denotes the element $-\alpha(S)$. This convention explains why we have denoted the simplex oriented opposite to $(p_0, \dots, p_n)$ by $-(p_0, \dots, p_n)$.

We now define the homomorphisms $\partial_n$: $C_n(K, \alpha) \to C_{n-1}(K, \alpha)$ for $n \geqslant 2$ by assigning to each generator $\alpha(S)$ of the group $C_n(K, \alpha)$ the sum of its $(n-1)$-dimensional faces oriented coherently with $\alpha(S)$ (see Theorem 5, § 2). We thus have

$$(8) \quad \partial_n(\alpha(S)) = \partial_n(p_0, \dots, p_n) = \sum_{i=0}^{n} (-1)^i (p_0, \dots, \hat{p}_i, \dots, p_n),$$

where the symbol ^ over a vertex signifies that this vertex is omitted, and $(p_0, \dots, p_n)$ is an arbitrary representation of the generator $\alpha(S)$. Since each generator of the group $C_1(K, \alpha)$ has exactly one representation in the form $(p_0, p_1)$, the formula

$$(9) \qquad \partial_1(p_0, p_1) = p_1 - p_0$$

defines a homomorphism $\partial_1$: $C_1(K, \alpha) \to C_0(K, \alpha)$. Note that formula (8) for $n = 1$ reduces to formula (9). The homomorphisms $\partial_n$, $n = 1, 2, \dots$ are called *boundary operators*.

THEOREM 1. *For every $n \geqslant 2$ we have*

$$(10) \qquad \partial_{n-1}\partial_n = 0.$$

Proof. It suffices to verify that $\partial_{n-1}\partial_n(\alpha(S)) = 0$, since the extension of the mapping $\partial_{n-1}\partial_n$ from the set of generators to the homomorphism of the whole group $C_n(K, \alpha)$ is uniquely determined. Let $\alpha(S) = (p_0, \dots, p_n)$. We then have

$$\begin{aligned}\partial_{n-1}\partial_n(\alpha(S)) &= \partial_{n-1}\Big(\sum_{i=0}^{n} (-1)^i (p_0, \dots, \hat{p}_i, \dots, p_n)\Big) \\ &= \sum_{i=0}^{n} (-1)^i \Big[\sum_{j=0}^{i-1} (-1)^j (p_0, \dots, \hat{p}_j, \dots, \hat{p}_i, \dots, p_n) \\ &\qquad + \sum_{j=i+1}^{n} (-1)^{j-1} (p_0, \dots, \hat{p}_i, \dots, \hat{p}_j, \dots, p_n)\Big].\end{aligned}$$

The simplex $(p_0, \dots, \hat{p}_k, \dots, \hat{p}_l, \dots, p_n)$ appears in the above linear combination twice: first with the coefficient $(-1)^l(-1)^k$

and then with the coefficient $(-1)^k(-1)^{l-1}$. Since the sum of these coefficients is zero, we have $\partial_{n-1}\partial_n(\alpha(S)) = 0$.

The group $Z_n(K, \alpha) = \operatorname{Ker} \partial_n \subset C_n(K, \alpha)$, where $n \geqslant 1$, will be called the *group of n-dimensional cycles* of the complex $(K, \alpha)$; we assume in addition that $Z_0(K, \alpha) = C_0(K, \alpha)$. The group $B_n(K, \alpha) = \operatorname{Im} \partial_{n+1} \subset C_n(K, \alpha)$, where $n \geqslant 0$ will be called the *group of n-dimensional boundaries* of the complex $(K, \alpha)$. It follows from (10) that $B_n(K, \alpha) \subset Z_n(K, \alpha)$; the quotient group $H_n(K, \alpha) = Z_n(K, \alpha)/B_n(K, \alpha)$ is called the *nth homology group* of the oriented complex $(K, \alpha)$. The index $n$ runs through the numbers $0, 1, 2, \ldots$, but all groups $H_n(K, \alpha)$ with indices exceeding the dimensions of $K$ are zero groups.

In the terminology of § 2, two $n$-dimensional cycles $z_1$ and $z_2$ are equivalent mod $B_n(K, \alpha)$ if $(z_1 - z_2) \in B_n(K, \alpha)$. In this case we say that $z_1$ and $z_2$ are *homologous to each other* in $(K, \alpha)$; instead of writing $z_1 \sim z_2 (\operatorname{mod} B_n(K, \alpha))$ we write

$$z_1 \approx z_2 \quad \text{in } (K, \alpha).$$

In particular, the condition $z = \partial_{n+1}(l)$ means that the cycle $z$ is homologous to zero in $(K, \alpha)$.

We shall show that the groups $H_n(K, \alpha)$ depend only on the complex $K$, and not on the choice of the orientation $\alpha$. For every pair of orientations $\alpha$, $\beta$ of $K$ define the homomorphisms $\theta_n^{\alpha\beta}\colon C_n(K, \alpha) \to C_n(K, \beta)$, putting for $n \geqslant 0$

$$\theta_n^{\alpha\beta}(\alpha(S)) = \begin{cases} \beta(S) & \text{if} \quad \alpha(S) = \beta(S), \\ -\beta(S) & \text{if} \quad \alpha(S) = -\beta(S). \end{cases}$$

We can easily see that $\theta_n^{\beta\alpha}\theta_n^{\alpha\beta}$ and $\theta_n^{\alpha\beta}\theta_n^{\beta\alpha}$ are identity isomorphisms of the groups $C_n(K, \alpha)$ and $C_n(K, \beta)$ respectively, whence $\theta_n^{\alpha\beta}$ is an isomorphism of the group $C_n(K, \alpha)$ onto $C_n(K, \beta)$. Moreover, transformations $\theta_n^{\alpha\beta}$ leave invariant every sequence $(p_0, \ldots, p_n)$; thus, the diagram

$$\begin{array}{ccc} C_n(K, \alpha) & \xrightarrow{\theta_n^{\alpha\beta}} & C_n(K, \beta) \\ \downarrow \partial_n & & \downarrow \partial_n \\ C_{n-1}(K, \alpha) & \xrightarrow{\theta_{n-1}^{\alpha\beta}} & C_{n-1}(K, \beta) \end{array}$$

is commutative, i.e. $\partial_n \theta_n^{\alpha\beta} = \theta_{n-1}^{\alpha\beta} \partial_n$ for $n = 1, 2, \ldots$ Consequently, under the isomorphism $\theta_n^{\alpha\beta}$ cycles are mapped into cycles and boundaries into boundaries, which implies that the quotient groups $H_n(K, \alpha)$ and $H_n(K, \beta)$ are isomorphic. The *nth homology group* of the (unoriented) complex $K$ will be defined as the group $H_n(K, \alpha)$ where $\alpha$ is an arbitrary orientation of the complex $K$; this group will be denoted by $H_n(K)$. Since in determining the homology, the choice of orientation is irrelevant, when considering groups of chains, cycles and boundaries, we shall, occasionally write $C_n(K)$, $Z_n(K)$ and $B_n(K)$, even though we shall always consider a fixed orientation of $K$. Moreover, we shall agree that in considering complexes with numbered vertices, the generators

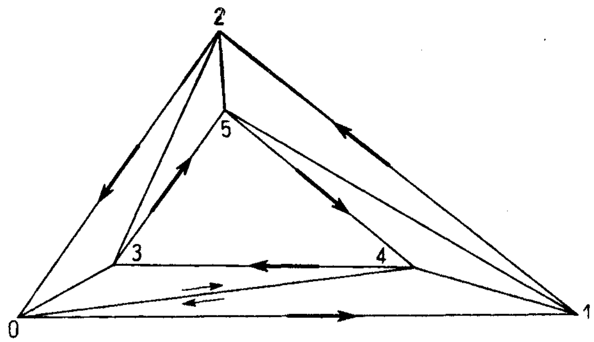



FIG. 21

of groups of chains will be taken as sequences of vertices with increasing indices.

EXAMPLE 1. Let $K$ be the complex consisting of all simplexes (0, 1 and 2-dimensional) given in Fig. 21, except the simplexes 012 and 345 (for simplicity we write $k$ instead of $p_k$). The chains

$$z_1 = (0, 1)+(1, 2)+(2, 0) \quad \text{and} \quad z_2 = (3, 4)+(4, 5)+(5, 3)$$

are cycles not homologous to zero in $K$, but homologous to one another.

Indeed, let

$$l = (0, 4, 3)+(0, 1, 4)+(1, 5, 4)+(1, 2, 5)+(2, 3, 5)+(0, 3, 2);$$

then $z_1 - z_2 = \partial_2 l$.

EXAMPLE 2. A complex $K$ will be called *connected* if it cannot be decomposed into two complexes without common vertices; in other words, if for every pair $p, q$ of its vertices, there exists a finite system of vertices $p = p_0, p_1, \ldots, p_r = q$ such that the simplexes $p_{i-1}p_i$ belong to $\boldsymbol{K}$ for $i = 1, 2, \ldots, r$. Clearly, the connectedness of $\boldsymbol{K}$ is equivalent to the connectedness of its underlying polyhedron $|\boldsymbol{K}|$.

We shall prove that

(11) *if a non-empty complex* $\boldsymbol{K}$ *is connected, then* $H_0(\boldsymbol{K}) \approx Z$,

where $Z$ is the group of integers.

Every element of the group $Z_0(\boldsymbol{K}) = C_0(\boldsymbol{K})$ is of the form

$$k_0 p_0 + \ldots + k_m p_m;$$

by assigning to it the sum $k_0 + \ldots + k_m$ we define an epimorphism of the group $Z_0(\boldsymbol{K})$ onto the group $Z$. To prove (11) it suffices to show that

$$(12) \qquad B_0(\boldsymbol{K}) = \{k_0 p_0 + \ldots + k_m p_m \colon k_0 + \ldots + k_m = 0\}.$$

Since $\partial_1(p_0, p_1) = p_1 - p_0$, the sum of the coefficients of every element in $B_0(\boldsymbol{K})$ is zero, i.e. the left-hand side of (12) is contained in the right-hand side. Let $q$ be a fixed vertex of $\boldsymbol{K}$. To prove the reverse inclusion it suffices to show that

$$k_0 p_0 + \ldots + k_m p_m \approx \Big(\sum_{i=0}^{m} k_i\Big) q \quad \text{in } \boldsymbol{K},$$

which in turn reduces to showing that for an arbitrary vertex $p \in \boldsymbol{K}$ and a number $k \in Z$ we have $kp \approx kq$ in $\boldsymbol{K}$.

Let us consider a system of vertices $p = p_0, p_1, \ldots, p_r = q$ such that $p_{i-1}p_i \in \boldsymbol{K}$ for $i = 1, 2, \ldots, r$ and the chain $l = k(p_1, p_0) + k(p_2, p_1) + \ldots + k(p_r, p_{r-1}) \in C_1(\boldsymbol{K})$. Since $\partial_1 l = kp_0 - kp_1 + kp_1 - kp_2 + \ldots + kp_{r-1} - kp_r = kp_0 - kp_r = kp - kq$, we get $kp \approx kq$ in $\boldsymbol{K}$, and the proof of (12) is complete.

We shall now introduce the concept of a cone on a chain, which will be used in the next example and in § 6.

Let $(\boldsymbol{K}, \alpha)$ be an arbitrary oriented complex, let $p$ be a vertex in $\boldsymbol{K}$, and let $\alpha(S) = (p_0, \ldots, p_n)$ be a generator of the group $C_n(\boldsymbol{K}, \alpha)$. If the simplex $pp_0 \ldots p_n$ belongs to $\boldsymbol{K}$ we say that $p$ can be *joined* with $\alpha(S)$ (or with $S$); it follows that $p \neq p_i$ for $i = 0, 1, \ldots, n$. If $p$ can be joined with $\alpha(S)$, then the oriented simplex $(p, p_0, \ldots, p_n)$ belongs to the group $C_{n+1}(\boldsymbol{K}, \alpha)$; this simplex will be denoted by $p\alpha(S)$. If $p$ can be joined with every simplex appearing in the chain $l \in C_n(K)$, with non-zero coefficient, we say that $p$ can be *joined with the chain l*. For an arbitrary vertex $p \in K$ which can be joined with a chain $l = k_1\alpha(S_1)+ \ldots$ $\ldots +k_m\alpha(S_m) \in C_n(K, \alpha)$, the *cone on l with vertex p* will be defined as the chain

$$pl = k_1 p\alpha(S_1)+ \ldots +k_m p\alpha(S_m) \in C_{n+1}(K, \alpha).$$

We can easily see that

$$\partial_{n+1}(pl) = l - p\partial_n l \quad \text{for} \quad n \geqslant 0, \tag{13}$$

where $\partial_0 l$ denotes the sum of the coefficients of the zero-dimensional chain $l$ and $pk = kp$.

EXAMPLE 3. If a complex $K$ has a vertex $p$ such that for any simplex $p_0 \ldots p_n \in K$ either $p$ is one of its vertices or $pp_0 \ldots p_n \in K$, we say that $K$ is a *cone*; $p$ is called the *vertex* of this cone, and the subcomplex $K_0 \subset K$ consisting of all simplexes which do not contain $p$ as a vertex will be called the *base* of this cone. Thus for instance, the complex $K^n$ consisting of all faces (including the improper one) of an arbitrary $n$-dimensional simplex is a cone; any vertex can serve in this case as the vertex of the cone, while the base will be the subcomplex consisting of the $(n-1)$-dimensional face opposite to the vertex of the cone (i.e. the face which does not contain the selected vertex), and of all its faces.

Let $K$ be a cone with vertex $p$ and base $K_0$. Let us select an orientation $\alpha_0$ of the complex $K_0$ and define the orientation $\alpha$ of the complex $K$ taking as generators of the group $C_n(K, \alpha)$, where $n \geqslant 1$, the generators of the group $C_n(K_0, \alpha_0)$ and simplexes of the form $p\alpha_0(S)$, where $\alpha_0(S)$ is a generator of the group $C_{n-1}(K_0, \alpha_0)$.

Any $n$-dimensional chain of the complex $(K, \alpha)$ has, for $n \geqslant 1$, the form

$$l = pl_1 + l_2, \quad \text{where} \quad l_1 \in C_{n-1}(K_0, \alpha_0), \quad l_2 \in C_n(K_0, \alpha_0).$$

In view of (13) the boundary of the chain $l$ is the chain

$$\partial_n l = l_1 - p\partial_{n-1} l_1 + \partial_n l_2 = -p\partial_{n-1} l_1 + (l_1 + \partial_n l_2).$$

If $l \in Z_n(K, \alpha)$, then $\partial_{n-1} l_1 = 0$ and $\partial_n l_2 = -l_1$; for the chain

$l' = pl_2 \in C_{n+1}(K, \alpha)$ we have

$$\partial_{n+1} l' = \partial_{n+1} p l_2 = l_2 - p\partial_n l_2 = pl_1 + l_2 = l,$$

i.e.

(14) $$Z_n(K, \alpha) = B_n(K, \alpha) \quad \text{for} \quad n \geqslant 1.$$

Since every cone is a connected complex, (14) and (11) imply:

(15) *for an arbitrary cone $K$ we have*

$$H_n(K) = 0 \text{ for } n \geqslant 1 \text{ and } H_0(K) \approx Z.$$

In particular, for the complex $K^n$ consisting of all the faces of a certain $n$-dimensional simplex, we have

(16) $$H_m(K^n) = 0 \quad \text{for} \quad m \geqslant 1, \quad H_0(K^n) \approx Z.$$

EXAMPLE 4. Consider the complex $S^n$ consisting of all the faces of dimension $\leqslant n$ of an $(n+1)$-dimensional simplex, where $n \geqslant 1$. Since $C_m(S^n) = C_m(K^{n+1})$ and $Z_m(S^n) = Z_m(K^{n+1})$ for $m \leqslant n$ and $B_m(S^n) = B_m(K^{n+1})$ for $m \leqslant n-1$, we have $H_0(S^n) \approx Z$ and $H_m(S^n) = 0$ for $m \leqslant n-1$. Let us find the $n$th homology group of $S^n$; clearly, $H_n(S^n) = Z_n(S^n)$, since $S^n$ has no simplexes, and hence no non-trivial chains, of dimension $n+1$. Let $l$ be a generator of the group $C_{n+1}(K^{n+1})$; the chain $\partial_{n+1} l \in C_n(S^n) = C_n(K^{n+1})$ is a non-zero cycle in $S^n$. Moreover, every $n$-dimensional cycle in $S^n$ is of the form $k\partial_{n+1} l$, since it is also a cycle in $K^{n+1}$, and in view of (16) it is a boundary of a certain $(n+1)$-dimensional chain. But every $(n+1)$-dimensional chain is of the form $kl$. Thus, for the complex $S^n$ consisting of all the faces of dimension $\leqslant n$ of an $(n+1)$-dimensional simplex $(n \geqslant 1)$, we have

(17) $$H_n(S^n) \approx Z, \quad H_m(S^n) = 0 \text{ for } m > n \text{ and } n > m \geqslant 1,$$
$$\text{and } H_0(S^n) \approx Z.$$

Let us now consider complexes $K$ and $L$ and a simplicial map $\varphi\colon K \to L$. Let $\alpha$ and $\alpha'$ be orientations of $K$ and $L$ respectively. Assuming for any generator $\alpha(S) = (p_0, \ldots, p_n)$ of the group $C_n(K, \alpha)$

$$\varphi_n^{\circ}(\alpha(S)) = \begin{cases} (\varphi(p_0), \ldots, \varphi(p_n)) & \text{if} \quad \varphi(p_j) \neq \varphi(p_k) \text{ for } j \neq k, \\ 0 \quad \text{otherwise}, \end{cases}$$

we define a homomorphism $\varphi_n^\circ\colon C_n(K,\alpha)\to C_n(L,\alpha')$ for $n = 0, 1, \ldots$

THEOREM 2. *For every $n \geqslant 1$ the diagram*

$$\begin{array}{ccc} C_n(K,\alpha) & \xrightarrow{\varphi_n^\circ} & C_n(L,\alpha') \\ \downarrow{\scriptstyle\partial_n} & & \downarrow{\scriptstyle\partial_n} \\ C_{n-1}(K,\alpha) & \xrightarrow{\varphi_{n-1}^\circ} & C_{n-1}(L,\alpha') \end{array}$$

*is commutative, i.e.* $\varphi_{n-1}^\circ\partial_n = \partial_n\varphi_n^\circ$.

Proof. It suffices to show that for every generator $\alpha(S) = (p_0, \ldots, p_n)$ of the group $C_n(K,\alpha)$ we have

(18) $$\varphi_{n-1}^\circ\partial_n(p_0, \ldots, p_n) = \partial_n\varphi_n^\circ(p_0, \ldots, p_n).$$

Equality (18) is obvious if $\varphi(p_j) \neq \varphi(p_k)$ for $j \neq k$.

Suppose now that $\varphi(p_j) = \varphi(p_k)$ for $j > k$. We then have

$$\varphi_{n-1}^\circ\partial_n(p_0, \ldots, p_n) = \varphi_{n-1}^\circ\Big[\sum_{i=0}^{n}(-1)^i(p_0, \ldots, \hat{p}_i, \ldots, p_n)\Big]$$

$$= \varphi_{n-1}^\circ[(-1)^k(p_0, \ldots, \hat{p}_k, \ldots, p_n)+(-1)^j(p_0, \ldots, \hat{p}_j, \ldots, p_n)+l],$$

where $l \in \operatorname{Ker}\varphi_{n-1}^\circ$.

If $(p_0, \ldots, \hat{p}_k, \ldots, p_n) \in \operatorname{Ker}\varphi_{n-1}^\circ$, i.e. if a certain pair of vertices is mapped into the same vertex of the complex $L$, then also $(p_0, \ldots, \hat{p}_j, \ldots, p_n) \in \operatorname{Ker}\varphi_{n-1}^\circ$ and $\varphi_{n-1}^\circ\partial_n(p_0, \ldots, p_n) = 0$. On the other hand, if $(p_0, \ldots, \hat{p}_k, \ldots, p_n) \notin \operatorname{Ker}\varphi_{n-1}^\circ$, then

$$\varphi_{n-1}^\circ(p_0, \ldots, \hat{p}_k, \ldots, p_n) = (\varphi(p_0), \ldots, \widehat{\varphi(p_k)}, \ldots, \varphi(p_n))$$

and

$$\varphi_{n-1}^\circ(p_0, \ldots, \hat{p}_j, \ldots, p_n) = (\varphi(p_0), \ldots, \widehat{\varphi(p_j)}, \ldots, \varphi(p_n)).$$

In view of the equality $\varphi(p_k) = \varphi(p_j)$ it follows that

$$\varphi_{n-1}^\circ(p_0, \ldots, \hat{p}_j, \ldots, p_n) = (-1)^{j-k-1}\varphi_{n-1}^\circ(p_0, \ldots, \hat{p}_k, \ldots, p_n)$$

and

$$\begin{aligned}\varphi_{n-1}^\circ\partial_n(p_0, \ldots, p_n) &= \varphi_{n-1}^\circ[(-1)^k(p_0, \ldots, \hat{p}_k, \ldots, p_n)]+\\ &\quad+\varphi_{n-1}^\circ[(-1)^j(p_0, \ldots, \hat{p}_j, \ldots, p_n)]\\ &= (-1)^k\varphi_{n-1}^\circ(p_0, \ldots, \hat{p}_k, \ldots, p_n)+\\ &\quad+(-1)^{2j-k-1}\varphi_{n-1}^\circ(p_0, \ldots, \hat{p}_k, \ldots, p_n) = 0.\end{aligned}$$

Under the assumption that $\varphi(p_j) = \varphi(p_k)$ for $j \neq k$, the left-hand side of (18) is thus equal to zero; in this case, the vanishing of the right-hand side is obvious, which proves (18).

It follows from Theorem 2 that the homomorphism $\varphi_n^\circ$ maps cycles into cycles and boundaries into boundaries, hence it determines the homomorphism $\varphi_n^*\colon H_n(K, \alpha) \to H_n(L, \alpha')$; we call $\varphi_n^*$ the *homomorphism induced by the simplicial map* $\varphi$.

We easily observe that for other orientations $\beta$ and $\beta'$ of the complexes $K$ and $L$ respectively, we have

$$(19) \qquad \theta_n^{\alpha'\beta'}\varphi_n^\circ = \varphi_n^\circ\theta_n^{\alpha\beta},$$

i.e. the diagram

$$\begin{array}{ccc} C_n(K, \alpha) & \xrightarrow{\varphi_n^\circ} & C_n(L, \alpha') \\ \downarrow \theta^{\alpha\beta} & & \downarrow \theta^{\alpha'\beta'} \\ C_n(K, \beta) & \xrightarrow{\varphi_n^\circ} & C_n(L, \beta') \end{array}$$

is commutative. Consequently, the homomorphism $\varphi_n^*$ of the group $H_n(K)$ into the group $H_n(L)$ induced by $\varphi\colon K \to L$ does not depend on the choice of orientations $\alpha$ and $\alpha'$ which served in defining the homology groups of non-oriented complexes.

Simple verification yields

THEOREM 3. *By assigning to simplicial complexes their nth homology groups ($n = 0, 1, 2, \ldots$) and to simplicial maps the induced homomorphisms, we define a covariant functor from the category* **SC** *to the category* **G**, *i.e. for simplicial maps* $\varphi\colon K \to L$ and $\psi\colon L \to M$ *we have*

$$(\psi\varphi)_n^* = \psi_n^*\varphi_n^*$$

*and for the identity map* $\iota\colon K \to K$ *we have*

$$\iota_n^* = 1_{H_n(K)}.$$

## § 5. Chain complexes

The construction of homology groups described in the preceding section can be split into two stages; more precisely, the functor which assigns homology groups to simplicial complexes, and induced homomorphisms to simplicial maps, can be represented as a composition of two functors. Such a decomposition allows a more thorough analysis of this functor and leads to computational simplifications. Defining the pair of functors, whose composition yields the homology functor requires introducing a new category.

Definition 1. By a *chain complex* we shall mean an infinite sequence

$$C_0 \xleftarrow{\partial_1} C_1 \leftarrow \dots \leftarrow C_{n-1} \xleftarrow{\partial_n} C_n \leftarrow \tag{20}$$

of Abelian groups $C_n$ and homomorphisms $\partial_n\colon C_n \to C_{n-1}$, to be called *boundary operators,* such that $\partial_{n-1}\partial_n = 0$ for $n = 2, 3, \dots$ The chain complex (20) will be denoted by $C = \{C_n, \partial_n\}$.

Definition 2. A *chain homomorphism* of a chain complex $C = \{C_n, \partial_n\}$ into a chain complex $D = \{D_n, \partial_n\}$ will be defined as an infinite sequence of homomorphisms $f = \{f_n\}$, where $f_n\colon C_n \to D_n$ and the diagram

$$\begin{array}{ccc} C_n & \xrightarrow{f_n} & D_n \\ {\scriptstyle\partial_n}\downarrow & & \downarrow{\scriptstyle\partial_n} \\ C_{n-1} & \xrightarrow{f_{n-1}} & D_{n-1} \end{array}$$

is commutative, i.e.

$$f_{n-1}\partial_n = \partial_n f_n \quad \text{for} \quad n \geqslant 1. \tag{21}$$

The category whose objects are chain complexes and whose morphisms are chain homomorphisms will be called the *category of chain complexes*, and will be denoted by **CC**. The composition of morphisms $f = \{f_n\}\colon C \to D$ and $g = \{g_n\}\colon D \to E$ is defined as the chain homorphism $gf = h = \{h_n\}\colon C \to E$ where $h_n = g_n f_n$; the identity morphism of a complex $C = \{C_n, \partial_n\}$ is the sequence $\{1_{C_n}\}$, denoted by $1_C$.

A morphisms $f = \{f_n\}$ of the category **CC** will be called a *monomorphisms* (*epimorphisms, isomorphisms*) if every homomorphism $f_n$ is a monomorphism (epimorphism, isomorphism). The concept of chain subcomplexes can be carried from the category **G** to the category **CC**.

A complex $C = \{C_n, \partial'_n\}$ is a *chain subcomplex* of $D = \{D_n, \partial_n\}$ if $C_n$ is a subgroup of $D_n$ and $\partial'_n = \partial_n | C_n$ for each $n$.

An example of a chain complex is provided by the sequence of groups of chains and boundary operators $C(\boldsymbol{K}, \alpha) = \{C_n(\boldsymbol{K}, \alpha), \partial_n\}$ of an oriented simplicial complex $(\boldsymbol{K}, \alpha)$.

Taking another orientation $\beta$ of the complex $K$ we obtain a chain complex $C(K, \beta)$ isomorphic to $C(K, \alpha)$; this means not only that the corresponding groups of complexes $C(K, \alpha)$ and $C(K, \beta)$ are isomorphic, but also that the boundary operators act in the same manner in both complexes. Thus, we may treat chain complexes as assigned to non-oriented complexes; the chain complex assigned to a complex $K$ will be denoted by $C(K)$. It follows from Theorem 2 of the last section and from formula (19) that to every simplicial map $\varphi\colon K \to L$ corresponds the chain homomorphism $\varphi^\circ = \{\varphi_n^\circ\}\colon C(K) \to C(L)$. We easily note that if $\varphi\colon K \to L$ and $\psi\colon L \to M$ are simplicial maps, then $(\psi\varphi)^\circ = \psi^\circ\varphi^\circ$, and that to the identity map $\iota\colon K \to K$ corresponds the identity $\iota^\circ = 1_{C(K)} = \{1_{C_n(K)}\}$.

We thus obtain

THEOREM 1. *By assigning to every simplicial complex $K$ the chain complex $C(K)$ and to every simplicial map $\varphi\colon K \to L$ the chain homomorphism $\varphi^\circ\colon C(K) \to C(L)$, we define a covariant functor from the category* **SC** *to the category* **CC**.

As in the case of complexes $C(K)$, for any chain complex $C = \{C_n, \partial_n\}$ one can define the *group of n-dimensional cycles* $Z_n(C) = \operatorname{Ker}\partial_n$ for $n \geqslant 1$, $Z_0(C) = C_0$, and the group of *n-dimensional boundaries* $B_n(C) = \operatorname{Im}\partial_{n+1}$. The equality $\partial_{n-1}\partial_n = 0$ implies that $B_n(C) \subset Z_n(C)$ for $n = 0, 1, 2, \ldots$; the quotient group $H_n(C) = Z_n(C)/B_n(C)$ will be called the *nth homology group* of the chain complex $C$.

It follows from (21) that a chain homomorphism $f\colon C \to D$ maps cycles into cycles and boundaries into boundaries; hence for each $n$ it determines $f_n^*\colon H_n(C) \to H_n(D)$, the *induced homomorphism.* We easily verify that if $f\colon C \to D$ and $g\colon D \to E$ are chain homomorphisms, then $(gf)_n^* = g_n^* f_n^*$ and that the identity $1_C\colon C \to C$ induces the identity $(1_C)_n^* = 1_{H_n(C)}$. We thus obtain

THEOREM 2. *By assigning to chain complexes their nth homology groups ($n = 0, 1, 2, \ldots$) and to the chain homomorphisms the induced homomorphisms of the nth homology groups, we define a covariant functor from the category* **CC** *to the category* **G**.

It is not difficult to see that the homology functor constructed in the preceding section is identical with the composition of functors described in Theorems 1 and 2. The passage from topology to algebra is described by the first of these functors, while the second has a purely algebraic character.

Definition 3. Two chain homomorphisms $f, g\colon C \to D$ where $f = \{f_n\}$, $g = \{g_n\}$, $C = \{C_n, \partial_n\}$ and $D = \{D_n, \partial_n\}$ will be called *chain homotopic*, and denoted by $f \approx g$, if there exists a sequence of homomorphisms $\Delta = \{\Delta_n\}$, where $\Delta_n\colon C_n \to D_{n+1}$ or $n = 0, 1, \ldots$ such that

(22) $$f_n - g_n = \partial_{n+1}\Delta_n + \Delta_{n-1}\partial_n \quad \text{for} \quad n = 1, 2, \ldots$$

and

(23) $$f_0 - g_0 = \partial_1 \Delta_0 .$$

The sequence of homomorphisms $\Delta$ will be called the *chain homotopy* between $f$ and $g$; note that it is not a chain homomorphism of $C$ into $D$, since $\Delta_n$ maps $C_n$ into $D_{n+1}$.

We easily verify that the relation of chain homotopy is reflexive, symmetric and transitive, i.e. it is an equivalence relation.

THEOREM 3. *If chain homomorphisms $f, g\colon C \to D$ are chain homotopic, then $f_n^* = g_n^*$ for $n = 0, 1, \ldots$*

Proof. Let $k$ be an arbitrary element of the group $H_n(C)$ and $z \in k$; we then have $\partial_n z = 0$, and in view of (22) and (23)

$$f_n(z) - g_n(z) = \partial_{n+1}\Delta_n(z).$$

Thus, $f_n(z) \sim g_n(z) \pmod{B_n(D)}$ and $f_n^*(k) = g_n^*(k)$.

We shall now describe a certain method of obtaining chain homotopies. This method will play a fundamental role in the construction of the homology functor from the category **P** of polyhedra and continuous mappings to the category **G**.

A chain complex $C = \{C_n, \partial_n\}$ will be called *geometric* if all groups $C_n$ are free groups with a finite number of generators, and for each $n$ a certain system $c_1^n, \ldots, c_{m_n}^n$ of generators of $C_n$ is selected. If in the representation of the boundary $\partial_{n+1} c_j^{n+1}$ as a linear combination of generators $c_1^n, \ldots, c_{m_n}^n$, the generator $c_k^n$ appears with a coefficient different from zero, we write $c_k^n < c_j^{n+1}$. By assigning to the element $c = k_1 c_1^0 + \ldots + k_{m_0} c_{m_0}^0$ of $C_0$ the integer $\partial_0(c) = \sum_{i=1}^{m_0} k_i$, we define a homomorphism of the group $C_0$ into the group $Z$ of integers.

A chain homomorphism $f\colon C \to D$, where $C$ and $D$ are geometric will be called *geometric* if $\partial_0(f_0(c)) = \partial_0(c)$ for every $c \in C_0$, i.e. if the diagram

$$\begin{array}{ccc} C_0 & \xrightarrow{f_0} & D_0 \\ \downarrow{\scriptstyle \partial_0} & & \downarrow{\scriptstyle \partial_0} \\ Z & \xrightarrow{1_Z} & Z \end{array}$$

is commutative.

Chain complexes assigned to oriented simplicial complexes are geometric, and so are chain homomorphisms assigned to simplicial maps.

A geometric chain complex $C$ will be called *acyclic*, if $B_0(C) = \operatorname{Ker}\partial_0$ and $H_n(C) = 0$ for $n \geqslant 1$. From (12) and (15) it follows that

(24) *if $K$ is a cone, then $C(K, \alpha)$ is an acyclic complex.*

Let $f = \{f_n\}\colon C \to D$ be a chain homomorphism of geometric complexes; a *carrier* of $f$ will be defined as a function assigning to each generator $c_i^m$ of the group $C_m$ (for $m = 0, 1, \ldots$) a subcomplex $D(c_i^m) = \{D_n(c_i^m), \partial_n\}$ of the complex $D$, where the group $D_n(c_i^m)$ is generated by a certain subset of the distinguished set of generators of $D_n$ (which implies that the operator $\partial_0$ in the complex $D(c_i^m)$ is the restriction of the operator $\partial_0$ in $D$), such that

(25) $D(c_k^n)$ is a subcomplex of $D(c_j^{n+1})$ if $c_k^n < c_j^{n+1}$

and

(26) $$f_n(c_i^n) \in D_n(c_i^n).$$

All the complexes $D(c_i^m)$ are geometric chain complexes.

The carrier will be called *acyclic* if each of the complexes $D(c_i^m)$ is acyclic.

THEOREM 4. *If $C$ and $D$ are geometric chain complexes, and $f$ and $g$ are geometric chain homomorphisms of $C$ into $D$ with a common acyclic carrier, then $f$ and $g$ are chain homotopic.*

Proof. Let the function assigning to each generator $c_i^m$ of the group $C_m$ the subcomplex $D(c_i^m)$ of the complex $D$ be an acyclic carrier of $f$ and $g$. We shall define recursively the homomorphisms $\Delta_n$ satisfying (22) and (23). Since $f$ and $g$ are geometric, for each generator $c_j^0$ of $C_0$ we have

$$\partial_0(f_0(c_j^0)-g_0(c_j^0)) = 0$$

and since the complex $D(c_j^0)$ is acyclic, there exists an element $\Delta_0(c_j^0) \in D_1(c_j^0)$ such that

$$(27) \qquad f_0(c_j^0)-g_0(c_j^0) = \partial_1\Delta_0(c_j^0).$$

Since the group $C_0$ is free, the choice of the element $\Delta_0(c_j^0)$ for each of its generators determines the homomorphism $\Delta_0\colon C_0 \to D_0$ satisfying, in view of (27), condition (23).

Assume that a homomorphism $\Delta_n\colon C_n \to D_{n+1}$ is defined for $n \leqslant m$ and satisfies (22) or (23) if $n = 0$, and

$$(28) \qquad \Delta_n(c_j^n) \in D_{n+1}(c_j^n)$$

for every generator $c_j^n$ of the group $C_n$.

Let $c_j^{m+1}$ be an arbitrary generator of the group $C_{m+1}$. The boundary $\partial_{m+1}c_j^{m+1}$ is an element of $C_m$, and $\Delta_m(\partial_{m+1}c_j^{m+1})$ is an element of the group $D_{m+1}$. Moreover, for every $c_k^m < c_j^{m+1}$ we have $\Delta_m(c_k^m) \in D_{m+1}(c_k^m) \subset D_{m+1}(c_j^{m+1})$, whence

$$d_j = f_{m+1}(c_j^{m+1})-g_{m+1}(c_j^{m+1})-\Delta_m(\partial_{m+1}c_j^{m+1}) \in D_{m+1}(c_j^{m+1}).$$

If $m \geqslant 1$, then from (22) for $n = m$ we deduce that

$$\partial_{m+1}(d_j) = f_m(\partial_{m+1}c_j^{m+1})-g_m(\partial_{m+1}c_j^{m+1}) \\ -[f_m(\partial_{m+1}c_j^{m+1})-g_m(\partial_{m+1}c_j^{m+1})-\Delta_{m-1}\partial_m\partial_{m+1}c_j^{m+1}] = 0;$$

on the other hand, if $m = 0$, relation (23) yields

$$\partial_1(d_j) = f_0(\partial_1 c_j^1)-g_0(\partial_1 c_j^1)-[f_0(\partial_1 c_j^1)-g_0(\partial_1 c_j^1)] = 0.$$

Since the complex $D(c_j^{m+1})$ is acyclic, there exists an element

$$\Delta_{m+1}(c_j^{m+1}) \in D_{m+2}(c_j^{m+1})$$

such that

$$(29) \quad f_{m+1}(c_j^{m+1})-g_{m+1}(c_j^{m+1})-\Delta_m(\partial_{m+1}c_j^{m+1}) = \partial_{m+2}\Delta_{m+1}(c_j^{m+1}).$$

Since the group $C_{m+1}$ is free, the choice of the element $\Delta_{m+1}(c_j^{m+1})$ for each of its generators determines the homomorphism $\Delta_{m+1}\colon C_{m+1} \to D_{m+2}$ satisfying, in view of (29), relation (22) for $n = m+1$. The condition (28) is also satisfied for $n = m+1$, which completes the construction of the chain homotopy $\Delta = \{\Delta_n\}$ between $f$ and $g$.

Using Theorem 4 we shall prove

THEOREM 5. *If simplicial maps* $\varphi, \psi$: $K \to L$ *are contiguous, then for every n they induce the same homomorphism of the group* $H_n(K)$ *into the group* $H_n(L)$.

P r o o f. Let $\alpha$ and $\beta$ be arbitrary orientations of complexes $K$ and $L$. For every simplex $S \in K$ the simplexes $\varphi(S)$ and $\psi(S)$ are faces of a certain simplex $T(S) \in L$. Let us assign to a generator $\alpha(S)$ of the group $C_n(K, \alpha)$ the subcomplex $D(\alpha(S)) = C(L(S))$ of the complex $D = C(L, \beta)$, where $L(S)$ is the subcomplex consisting of all faces of that simplex in $L$, whose set of vertices equals the union of the sets of vertices of $\varphi(S)$ and $\psi(S)$. We note that in this way we define an acyclic carrier of mappings $\varphi^\circ$ and $\psi^\circ$; therefore $\varphi^\circ \approx \psi^\circ$ and $\varphi_n^* = \psi_n^*$ in virtue of Theorem 3.

## § 6. Homology groups of polyhedra

We shall show now that homology groups of a simplicial complex depend only on the underlying polyhedron of that complex, and we shall describe homomorphisms of those groups induced by continuous mappings of polyhedra. This will lead to defining homology groups of polyhedra.

We start by showing that homology groups of an arbitrary simplicial complex and of its barycentric subdivision are isomorphic.

Suppose that we are given an arbitrary simplicial complex $K$ and let $\alpha$ be its orientation. We assign an orientation to the barycentric subdivision $K'$ by taking as generators of the $n$th group of chains decreasing sequences of vertices, i.e. sequences $(b(S_0), \ldots, b(S_n))$ where $S_0 > \ldots > S_n$. Denote the orientation so defined by $\beta$. The simplex $S_0 \in K$ will be called the *carrier* of the simplex $(b(S_0), \ldots, b(S_n))$; we can easily see that $b(S_0) \ldots b(S_n) \subset S_0$. If a simplex $S_0 \in K$ is the carrier of each of the simplexes appearing with non-zero coefficient in a certain chain $l \in C_n(K', \beta)$, we say that $S_0$ is the *carrier* of that chain.

In the first section we defined a class of simplicial maps of $K'$ into $K$, called standard translations; let $\sigma$: $K' \to K$ be one of them. It induces the chain homomorphism $\sigma^\circ$: $C(K', \beta) \to C(K, \alpha)$. We shall now define a chain homomorphism $s = \{s_n\}: C(K, \alpha) \to C(K', \beta)$ such that $s_n^*$ is the inverse of $\sigma_n^*$. This will imply that $\sigma_n^*$ is an isomorphism.

We define $s = \{s_n\}$ by induction. For $n = 0, 1, 2, \ldots$ we shall construct a homomorphism $s_n$: $C_n(K, \alpha) \to C_n(K', \beta)$ such that

(30) $$s_{n-1}\partial_n = \partial_n s_n \quad \text{for} \quad n \geqslant 1$$

and

(31) $S$ is the carrier of $s_n(\alpha(S))$ for every

$n$-dimensional simplex $S \in K$.

The group of zero-dimensional chains of $K$ is a subgroup of the group of zero-dimensional chains of $K'$. For $s_0$ we take the embedding, i.e. the monomorphism assigning to an element of $C_0(K, \alpha)$ the same element in $C_0(K', \beta)$; evidently $s_0$ satisfies condition (31) for $n = 0$.

Suppose now that homomorphisms $s_n$: $C_n(K, \alpha) \to C_n(K', \beta)$ are defined for $n \leqslant m$ and satisfy conditions (30) and (31).

Let $\alpha(S)$ be an arbitrary generator of the group $C_{m+1}(K, \alpha)$. The chain $\partial_{m+1}\alpha(S)$ belongs to the group $C_m(K, \alpha)$. By (31) for $n = m$, the carrier of any simplex in $K'$ which appears with a non-vanishing coefficient in the chain $s_m\partial_{m+1}\alpha(S)$ coincides with a face of $S$. Thus, the vertex $b(S)$ can be joined with the chain $s_m\partial_{m+1}\alpha(S)$ and the cone $b(S)s_m\partial_{m+1}\alpha(S)$ is a well defined element of the group $C_{m+1}(K', \beta)$. Putting

(32) $$s_{m+1}(\alpha(S)) = b(S)s_m\partial_{m+1}\alpha(S)$$

for every generator $\alpha(S)$ of the group $C_{m+1}(K, \alpha)$, we define the homomorphism $s_{m+1}$: $C_{m+1}(K, \alpha) \to C_{m+1}(K', \beta)$ satisfying condition (31) for $n = m+1$. Using (13) we obtain

$$\partial_{m+1} s_{m+1}(\alpha(S)) = s_m\partial_{m+1}\alpha(S) - b(S)\partial_m s_m\partial_{m+1}\alpha(S),$$

which implies that

(33) $$\partial_{m+1} s_{m+1}(\alpha(S)) = s_m\partial_{m+1}\alpha(S)$$

in view of (30) for $n = m$ if $m \geqslant 1$, or in view of the definition of $s_0$ and the easily verifiable identity $\partial_0\partial_1 = 0$ if $m = 0$. It follows from (33) that condition (30) holds for $n = m+1$, which completes the construction of the sequence $\{s_n\}$.

We shall now make a remark which will be used in the proof of Theorem 1. Suppose we are given simplicial complexes $\boldsymbol{K}$ and $\boldsymbol{L}$ and a simplicial map $\varphi\colon \boldsymbol{K}\to\boldsymbol{L}$. If $p$ is a vertex of $\boldsymbol{K}$ joined with a simplex $S$, then $\varphi(p)$ either can be joined with $\varphi(S)$ or is one of its vertices. To simplify the formulation of the relation between chains in $\boldsymbol{K}$ and $\boldsymbol{L}$ we shall treat $p$ as joined with every oriented simplex $\alpha(S)$ which contains $p$, and define the cone $p\,\alpha(S)$ as the zero of the corresponding group of chains. We note easily that (under the above convention), if $p$ can be joined with $l\in C_m(\boldsymbol{K})$, then $\varphi(p)$ can be joined with $\varphi_m(l)$ and

$$\varphi^{\circ}_{m+1}(pl)=\varphi(p)\varphi^{\circ}_m(l). \tag{34}$$

THEOREM 1. *If* $\sigma\colon \boldsymbol{K}'\to\boldsymbol{K}$ *is a standard translation and* $s=\{s_n\}\colon C(\boldsymbol{K},\alpha)\to C(\boldsymbol{K}',\beta)$ *is the chain homomorphism constructed above, then for every* $n$

$$\sigma^{\circ}_n s_n = 1_{C_n(\boldsymbol{K},\alpha)},\qquad \textit{hence}\qquad \sigma^*_n s^*_n = 1_{H_n(\boldsymbol{K},\alpha)} \tag{35}$$

*and*

$$s\sigma^{\circ}\approx 1_{C(\boldsymbol{K}',\beta)},\qquad \textit{hence}\qquad s^*_n\sigma^*_n = 1_{H_n(\boldsymbol{K}',\beta)}. \tag{36}$$

P r o o f. We shall prove (35) by induction. The validity of $\sigma^{\circ}_0 s_0 = 1_{C_0(\boldsymbol{K},\alpha)}$ follows from the fact that $\sigma$ keeps still all the vertices of the complex $\boldsymbol{K}'$ which belong to $\boldsymbol{K}$.

Suppose that (35) holds for $n\leqslant m$. Let $\alpha(S)=(p_0,\dots,p_{m+1})$ be a generator of the group $C_{m+1}(\boldsymbol{K},\alpha)$. Using (32), (34) and (35) for $n=m$ we have

$$\begin{aligned}\sigma^{\circ}_{m+1}s_{m+1}\big(\alpha(S)\big) &= \sigma^{\circ}_{m+1}[b(S)s_m\partial_{m+1}\alpha(S)]\\ &= \sigma\big(b(S)\big)\sigma^{\circ}_m s_m\partial_{m+1}\alpha(S)=\sigma\big(b(S)\big)\partial_{m+1}\alpha(S).\end{aligned} \tag{37}$$

Since $\sigma(b(S))=p_j$ for a certain $j\leqslant m+1$, we obtain

$$\begin{aligned}\sigma\big(b(S)\big)\partial_{m+1}\alpha(S) &= p_j\sum_{i=0}^{m+1}(-1)^i(p_0,\dots,\hat{p}_i,\dots,p_{m+1})\\ &= (-1)^j(p_j,p_0,\dots,\hat{p}_j,\dots,p_{m+1})\\ &= (-1)^{2j}(p_0,\dots,p_{m+1})=\alpha(S).\end{aligned} \tag{38}$$

Equations (37) and (38) imply the validity of (35) for $n=m+1$, which completes the proof of the first part of the theorem.

To prove the second assertion, it suffices to define a common acyclic carrier of $s\sigma^{\circ}$ and $1_{C(K',\beta)}$. Let $c_i^m = \big(b(S_0), \ldots, b(S_m)\big)$ be an arbitrary generator of the group $C_m(K', \beta)$. Denote by $S_0$ the subcomplex of $K'$ consisting of all the faces of the simplex $S_0$ and take as $D(c_i^m)$ the subcomplex $C(S_0', \beta)$ of the complex $C(K', \beta)$. Using (31) we easily verify that the function so defined is a carrier of $s\sigma^{\circ}$; the fact that it is also a carrier of $1_{C(K',\beta)}$ is obvious. The acyclicity follows from the fact that the complex $S_0'$ is a cone with vertex $b(S_0)$.

When defining homomorphisms $s_n$ we used a certain orientation of the complex $K$; one can verify that homomorphisms $s_n'$ defined in the same manner for another orientation $\alpha'$ of the complex $K$ satisfy the condition $s_n' = s_n\theta_n^{\alpha'\alpha}$. It follows that $s_n^*$ can be treated as a homomorphism of the group $H_n(K)$ into the group $H_n(K')$ since it does not depend on the choice of the orientation, which served for determining the homology group of the non-oriented complex $K$. The same conclusion can be reached from Theorem 1 since the homomorphism $s_n^*$ is inverse to the homomorphism $\sigma_n^*$ induced by a simplicial map, and such homomorphisms do not depend on orientation.

One considers also compositions of homomorphisms $s_n^*$. If $K$ is an arbitrary simplicial complex, and $K^{(m)}$ is its $m$th barycentric subdivision, then composing the homomorphisms

$$s_n^*\colon H_n(K) \to H_n(K'), \quad s_n^*\colon H_n(K') \to H_n(K^{(2)}), \quad \ldots,$$
$$s_n^*\colon H_n(K^{(m-1)}) \to H_n(K^{(m)}),$$

we obtain a homomorphism

$$s_n^{(m)*}\colon H_n(K) \to H_n(K^{(m)}).$$

Theorem 1 and Theorem 3 of § 4 imply the following

COROLLARY. *If* $\sigma\colon K^{(m)} \to K$ *is a standard map, and* $s_n^{(m)*}\colon H_n(K) \to H_n(K^{(m)})$ *is the composition of the corresponding homomorphisms* $s_n^*$, *then for every* $n$

$$\sigma_n^* s_n^{(m)*} = 1_{H_n(K)} \quad \textit{and} \quad s_n^{(m)*}\sigma_n^* = 1_{H_n(K^{(m)})}.$$

Suppose now that we are given two simplicial complexes $K$ and $L$ and a continuous mapping $f\colon |K| \to |L|$. We shall assign to this mapping a homomorphism $f_n$: $H_n(K) \to H_n(L)$ for $n = 0, 1, \ldots$ The definition of $f_n$ is illustrated by the following diagram:

$$\begin{array}{ccccc} H_n(K) & \xrightarrow{s_n^{(m)*}} & H_n(K^{(m)}) & \underset{\sigma_n^*}{\overset{s_n^{(k)*}}{\rightleftarrows}} & H_n(K^{(l)}) \\ & \searrow^{\bar{\varphi}_n} & \downarrow{\scriptstyle \varphi_n^*} & \swarrow_{\psi_n^*} & \\ & & H_n(L) & & \end{array}$$

Let $\varphi$: $K^{(m)} \to L$ be a simplicial approximation of $f$, existing by virtue of the corollary to Theorem 9 of § 1, and let $\bar{\varphi}_n = \varphi_n^* s_n^{(m)*}$. We shall show that the homomorphism of $H_n(K)$ into $H_n(L)$ defined in the above manner does not depend on the choice of a simplicial approximation $\varphi$. Let us consider another simplicial approximation $\psi$: $K^{(l)} \to L$ of $f$. Without loss of generality we may assume $l \geqslant m$, i.e. $l = m+k$ with $k \geqslant 0$. Let $\sigma$: $K^{(l)} \to K^{(m)}$ be a standard map; from Theorem 8 of § 1 it follows that maps $\varphi\sigma$ and $\psi$ are contiguous, whence by Theorem 3 of § 4 and Theorem 5 of § 5 we have

$$\varphi_n^* \sigma_n^* = (\varphi\sigma)_n^* = \psi_n^*. \tag{39}$$

From the corollary to Theorem 1 we obtain

$$\sigma_n^* s_n^{(k)*} = 1_{H_n(K^{(m)})}. \tag{40}$$

Using (39) and (40) we get

$$\bar{\psi}_n = \psi_n^* s_n^{(l)*} = \psi_n^* s_n^{(k)*} s_n^{(m)*} = \varphi_n^* \sigma_n^* s_n^{(k)*} s_n^{(m)*} = \varphi_n^* s_n^{(m)*} = \bar{\varphi}_n,$$

which shows that the homomorphism of $H_n(K)$ into $H_n(L)$ defined above depends only on $f$; we shall denote this homomorphism by $f_n^*$.

THEOREM 2. *If $f$: $|K| \to |L|$ and $g$: $|L| \to |M|$, then*

$$(gf)_n^* = g_n^* f_n^* \quad \textit{for} \quad n \geqslant 0.$$

P r o o f. The following diagram illustrates the proof:

$$\begin{array}{ccccc} H_n(K) & & H_n(L) & & \\ \downarrow{\scriptstyle s_n^{(l)*}} & & {\scriptstyle \sigma_n^*}\uparrow\downarrow{\scriptstyle s_n^{(m)*}} & & \\ H_n(K^{(l)}) & \xrightarrow{\varphi_n^*} & H_n(L^{(m)}) & \xrightarrow{\psi_n^*} & H_n(M) \end{array}$$

Let $\psi$: $L^{(m)} \to M$ be a simplicial approximation of $g$, and let $\varphi$: $K^{(l)} \to L^{(m)}$ be a simplicial approximation of the mapping $f$: $|K| \to |L^{(m)}| = |L|$. From Theorem 6 of § 1 it follows that $\sigma\varphi$, where $\sigma$: $L^{(m)} \to L$ is a standard map, is a simplicial approximation of $f$, whence

$$f_n^* = (\sigma\varphi)_n^* s_n^{(l)*} = \sigma_n^* \varphi_n^* s_n^{(l)*}.$$

Since $g_n^* = \psi_n^* s_n^{(m)*}$, it follows from the Corollary to Theorem 1 that

$$g_n^* f_n^* = \psi_n^* s_n^{(m)*} \sigma_n^* \varphi_n^* s_n^{(l)*} = \psi_n^* \varphi_n^* s_n^{(l)*} = (\psi\varphi)_n^* s_n^{(l)*},$$

and we infer that $g_n^* f_n^* = (gf)_n^*$, since $\psi\varphi$ is a simplicial approximation of $gf$ (see Theorem 5 of § 1).

Let us also note that for the identity transformation $i$: $|K| \to |K|$ we have $i_n^* = 1_{H_n(K)}$ for $n = 0, 1, \ldots$ Indeed, as a simplicial approximation of $i$ we can take the identity of $K$, which is a simplicial map of $K$ into $K$.

THEOREM 3. *If* $f$: $|K| \to |L|$ *is a homeomorphism, then* $f_n^*$: $H_n(K) \to H_n(L)$ *is an isomorphism for* $n = 0, 1, \ldots$

P r o o f. Let $g$: $|L| \to |K|$ be the inverse of $f$. Thus, $gf$ and $fg$ are identity mappings of the polyhedra $|K|$ and $|L|$ respectively. Using Theorem 2 we get

$$g_n^* f_n^* = (gf)_n^* = 1_{H_n(K)} \quad \text{and} \quad f_n^* g_n^* = (fg)_n^* = 1_{H_n(L)},$$

which implies that the homomorphism $g_n^*$ is inverse to $f_n^*$, and $f_n^*$ is an isomorphism.

It follows from Theorem 3 that if the underlying polyhedra of two complexes are homeomorphic, then the complexes have the same homology groups. This allows us to define homology groups of polyhedra and, more generally, homology groups of spaces homeomorphic with polyhedra. A topological space homeomorphic to a polyhedron will be called a *curvilinear polyhedron* (or a *triangulable space*). Thus, $X$ is a curvilinear polyhedron if, and only if, there exists a simplicial complex $K$ and a homeomorphism $t$: $|K| \to X$. The pair $(K, t)$ will be called a *triangulation* of the triangulable space $X$. Triangulable spaces and their continuous maps form the category **TS**. The *nth homology group* $H_n(X)$ of a triangulable space $X$ will be defined as the group $H_n(K)$ where

$(K, t)$ is a certain triangulation of $X$. Theorem 3 implies that this definition is independent of the choice of triangulation, for if $(K_1, t_1)$ is another triangulation of $X$, then $(t_1^{-1}t)_n^*$ is an isomorphism of the group $H_n(K)$ onto the group $H_n(K_1)$.

Suppose now that we are given two curvilinear polyhedra $X$ and $Y$ and let $(K, t)$ and $(L, u)$ be their triangulations; moreover, let $f\colon X \to Y$ be an arbitrary continuous mapping. The map $u^{-1}ft\colon |K| \to |L|$ induces a homomorphism of the group $H_n(K)$ into the group $H_n(L)$ for $n = 0, 1, 2, \ldots$ This homomorphism, to be denoted by $f_n^*$, will be regarded as a homomorphism of $H_n(X)$ into $H_n(Y)$ and will be called the *homomorphism induced by f*. We easily note that for other triangulations $(K_1, t_1)$ and $(L_1, u_1)$ of the spaces $X$ and $Y$ the diagram

$$\begin{array}{ccc} H_n(K_1) & \xrightarrow{(u_1^{-1}ft_1)_n^*} & H_n(L_1) \\ {\scriptstyle (t_1^{-1}t)_n^*}\uparrow & & \uparrow{\scriptstyle (u_1^{-1}u)_n^*} \\ H_n(K) & \xrightarrow{(u^{-1}ft)_n^*} & H_n(L) \end{array}$$

is commutative, i.e. the homomorphism $f_n\colon H_n(X) \to H_n(Y)$ does not depend on the choice of the triangulations of the spaces $X$ and $Y$, which served for its definition.

From Theorem 2 and the above definition we obtain

THEOREM 4. *By assigning to triangulable spaces the nth homology group ($n = 0, 1, \ldots$) and to continuous mappings of triangulable spaces the induced homomorphisms, we define a covariant functor from the category* **TS** *to the category* **G**.

EXAMPLE 1. The unit $n$-ball $\mathscr{K}_n$ of the space $\mathscr{E}^n$ is homeomorphic to the closure of the $n$-dimensional simplex, whence its triangulation is given by the complex $K^n$ consisting of all the faces of that simplex (together with a certain homeomorphism). Therefore, by (16) we have

$$(41) \qquad H_m(\mathscr{K}_n) = 0 \quad \text{for} \quad m \geqslant 1, \qquad H_0(\mathscr{K}_n) \approx Z.$$

A triangulation of the $n$-sphere $\mathscr{S}_n$, i.e. the surface of the ball $\mathscr{K}_{n+1}$, is given by the complex described in Example 4 of § 4 (together with a certain homeomorphism). Thus, in view of (17) we have for $n \geqslant 1$

(42) $$H_n(\mathscr{S}_n) \approx H_0(\mathscr{S}_n) \approx Z \quad \text{and} \quad H_m(\mathscr{S}_n) = 0 \quad \text{for other } m.$$

THEOREM 5. *Homotopic continuous mappings of triangulable spaces induce the same homomorphism of homology groups.*

Proof. It suffices to prove that if $f, g$: $|K| \to |L|$ are homotopic, then $f_n^* = g_n^*$ for every $n$. Note that if two maps of $|K|$ into $|L|$ have a common simplicial approximation, then they induce the same homomorphism of $H_n(K)$ into $H_n(L)$, since this simplicial approximation can be used for defining homomorphisms induced by those maps.

Let us take a continuous function $h$: $|K| \times \mathscr{I} \to |L|$ satisfying the condition

$$h(x, 0) = f(x) \quad \text{and} \quad h(x, 1) = g(x) \quad \text{for} \quad x \in |K|.$$

By the compactness of the Cartesian product $|K| \times \mathscr{I}$ and by Theorem 4 of § 5, Chapter XVI, for every $\varepsilon > 0$ there exist numbers $0 = t_0 < \ldots < t_m = 1$ such that $|f_i - f_{i-1}| < \varepsilon$ for $i = 1, \ldots, m$ where $f_i(x) = h(x, t_i)$. Taking as $\varepsilon$ the number satisfying the assertion of Theorem 9 of § 1, we obtain

$$f_n^* = (f_0)_n^* = (f_1)_n^* = \ldots = (f_{m-1})_n^* = (f_m)_n^* = g_n^*$$

for $n = 0, 1, \ldots$

EXAMPLE 2. It follows from Theorem 5 that the identity $f$: $\mathscr{S}_n \to \mathscr{S}_n$ and a constant mapping $g$: $\mathscr{S}_n \to \mathscr{S}_n$ are not homotopic for $n \geqslant 1$ (for $n = 0$ this fact is evident). As we know (see p. 270 and Exercise 12, Chapter XX), this is equivalent to the fixed-point theorem. Indeed, $g = g''g'$, where $g'$: $\mathscr{S}_n \to \{c\}$ and $g''$: $\{c\} \to \mathscr{S}_n$, whence $g_n^* = g_n''^* g_n'^*$ is the zero homomorphism (since $g_n'^*$ and $g_n''^*$ are zero homomorphisms, in view of the equality $H_n(c) = 0$), while $f_n^* = 1_Z$.

## § 7. Homology groups with coefficients

We shall now define homology groups of triangulable spaces with coefficients from an arbitrary Abelian group, and we shall assign homomorphisms of these groups to continuous maps of triangulable spaces. Different homology groups will correspond to different groups of coefficients, thus enlarging the number of algebraic invariants corresponding to triangulable spaces. The construction of homology groups with coefficients will consist in applying the tensor product functor to a chain complex before computing homology groups of this complex.

We start from the definition of the tensor product functor from the category **G** into itself.

Let $G$ be a fixed Abelian group. For every group $A$ from the category **G**, denote by $W(A, G)$ the free group generated by all pairs of the form $(a, g)$ with $a \in A$, $g \in G$, and by $R(A, G)$ the smallest subgroup of $W(A, G)$ containing all elements of the form

$$(43) \qquad (a_1+a_2, g)-(a_1, g)-(a_2, g) \quad \text{and} \quad (a, g_1+g_2)-(a, g_1)-(a, g_2).$$

The existence of such a subgroup follows from the fact that the intersection of an arbitrary family of subgroups of a given group is a subgroup of this group.

By the *tensor product* $A \otimes G$ we shall understand the quotient group $W(A, G)/R(A, G)$; the image of the generator $(a, g)$ under the natural epimorphism $p$ will be denoted by $a \otimes g$. Since elements of the form (43) are mapped into the zero of the tensor product, we have

$$(44) \qquad (a_1+a_2) \otimes g = a_1 \otimes g+a_2 \otimes g \quad \text{and} \quad a \otimes (g_1+g_2) = a \otimes g_1+a \otimes g_2.$$

It follows from (44) that for any integer $n \geqslant 0$ we have

$$(45) \qquad na \otimes g = n(a \otimes g) \quad \text{and} \quad a \otimes ng = n(a \otimes g);$$

in particular $0 \otimes g = 0 = a \otimes 0$, which in turn implies the validity of (45) for negative $n$.

THEOREM 1. *For any group* $G$

$$Z \otimes G \approx G.$$

P r o o f. Consider the homomorphism $f$ of the group $W(Z, G)$ into $G$ defined by assigning the element $kg \in G$ to the generator $(k, g)$. The homomorphism $f$ vanishes on elements of the form (43), whence $R(Z, G) \subset \operatorname{Ker} f$, and the formula

$$f'p(x) = f(x)$$

defines a homomorphism $f'$ of $Z\otimes G$ into $G$ such that $f'(k\otimes g) = kg$. Since $f'(1\otimes g) = g$, we conclude that $f'$ is an epimorphism. Every element $x$ of the group $Z\otimes G$ is of the form

$$m_1(n_1 \otimes g_1)+ \dots +m_k(n_k \otimes g_k)$$

and in view of (45) and (44) we get $x = 1\otimes g$, where $g = m_1n_1g_1+ + \dots +m_kn_kg_k$. If $f'(x) = 0$, then $g = 0$ and $x = 0$, hence $f'$ is also a monomorphism.

Suppose we are given a homomorphism $f$: $A \to B$. By assigning the element $(f(a), g) \in W(B, G)$ to the generator $(a, g) \in W(A, G)$ we define a homomorphism, to be denoted by $\bar{f}$, of the free group $W(A, G)$ into the group $W(B, G)$. We note immediately that $\bar{f}$ maps elements of the form (43) into elements of the same form in $W(B, G)$; consequently, $\bar{f}(R(A, G)) \subset R(B, G)$ and an induced homomorphism $f'$: $A \otimes G \to B \otimes G$ is defined such that

$$f'p(x) = p\bar{f}(x);$$

in particular, $f'(a \otimes g) = f(a) \otimes g$.

By simple checking we obtain

THEOREM 2. *By assigning to every group $A$ the tensor product $A \otimes G$ and to every homomorphism $f$: $A \to B$ the homomorphism $f'$: $A\otimes G \to B\otimes G$ we define a covariant functor from the category* **G** *into itself. Moreover, if $f$, $h$: $A \to B$, then $(f\pm h)' = f'\pm h'$.*

The tensor product functor can be applied also to chain complexes. Suppose we are given a chain complex $C = \{C_n, \partial_n\}$. The sequence $\{C_n \otimes G, \partial'_n\}$ of groups $C_n \otimes G$ and homomorphisms $\partial'_n$: $C_n \otimes G \to C_{n-1} \otimes G$ is a chain complex, since, by Theorem 2, $\partial'_{n-1}\partial'_n = (\partial_{n-1}\partial_n)'$ is the zero homomorphism. This complex will be called the *tensor product of the complex $C$ by $G$*, and will be denoted by $C\otimes G$. We can easily see that if $s = \{s_n\}$ is a chain homomorphism of $C$ into $D$, then $s' = \{s'_n\}$ is a chain homomorphism of $C\otimes G$ into $D\otimes G$, and that in this manner we define a covariant functor from category **CC** into itself.

Composing the functor described in Theorem 1 of § 5 with the tensor product functor, and then with the functor described in Theorem 2 of § 5, we obtain a covariant functor from the category **SC** to the category **G**. This functor assigns the group $H_n(K; G) = H_n(C(K)\otimes G)$ to any complex $K$ and the homomorphism $\varphi_n^{\circ\prime *}$: $H_n(K; G) \to H_n(L; G)$ to the simplicial map $\varphi$: $K \to L$. The group $H_n(K; G)$ will be called the *nth homology group of K with coefficients from G*, while the homomorphism $\varphi_n^{\circ\prime *}$, usually denoted shortly by $\varphi_n^*$, is called the *induced homomorphism.*

EXAMPLE. Let us compute the homology groups with coefficients from $G$ for the simplicial complex $K$ consisting of a single vertex $p$. We easily note that $C_n(K) = 0$ for $n \geqslant 1$ and $C_0(K) \approx Z$, hence all homomorphisms $\partial_n$ are zero homomorphisms. After taking the tensor product with the group $G$ we have $C_n(K) \otimes G = 0$ for $n \geqslant 1$, while $C_0(K) \otimes G \approx G$ by Theorem 1; the homomorphisms $\partial_n'$ are of course zero homomorphisms. It follows that $H_n(K; G) = 0$ for $n \geqslant 1$ and $H_0(K; G) \approx G$.

The construction of homology groups with coefficients from a group $G$ for triangulable spaces and of homomorphisms of those groups induced by continuous mappings is analogous to the construction shown in § 6. We shall merely sketch the main stages of this construction without going into detailed proof, which would merely repeat the computations of the preceding section.

To begin with, we show that homology groups with coefficients from $G$ of a complex $K$ and of its barycentric subdivision are isomorphic. We consider chain complexes $C(K) \otimes G$ and $C(K') \otimes G$ and chain homomorphisms $\sigma^{\circ\prime}$: $C(K') \otimes G \to C(K) \otimes G$ and $s'$: $C(K) \otimes G \to C(K') \otimes G$. The fact that $\sigma^{\circ\prime} s'$ is the identity is a consequence of the fact that the tensor product is a functor; the fact that $s'\sigma^{\circ\prime}$ is chain homotopic to identity follows from the fact that, in view of the last assertion of Theorem 2, the tensor product functor maps chain homotopies into chain homotopies.

Theorem 1 of § 6 and its corollary remain, therefore, valid for homology groups with coefficients. This allows us to assign a homomorphism $f_n^*$: $H_n(K; G) \to H_n(L; G)$ to every continuous mapping $f$: $|K| \to |L|$. Next, we prove theorems corresponding to Theorems 2 and 3 of § 6 for homologies of simplicial complexes with coefficients. Using these theorems we define the functor which assigns the group $H_n(X; G)$ to the triangulable space $X$, and the homomorphism $f_n^*$: $H_n(X; G) \to H_n(Y; G)$ to the continuous mapping $f$: $X \to Y$. We can also prove a theorem corresponding to Theorem 5, which asserts that for a pair of homotopic mappings $f, g: X \to Y$ the homomorphisms $f_n^*$ and $g_n^*$ of the group $H_n(X; G)$ into $H_n(Y; G)$ coincide.

From the above example it follows that the 0th homology group of a point with coefficients from $G$ is the group $G$. Therefore the homology functors with coefficients differ for different groups of coefficients. Note also that there exists spaces for which some homology groups without coefficients are zero groups, but groups with suitably chosen coefficients are non-zero groups (see Exercise 15).

One can easily show that the group $G \otimes Z$ is isomorphic to $G$ and that under identification of groups $G$ and $G \otimes Z$, and $H$ and $H \otimes Z$ the homomorphisms $f: G \to H$ and $f': G \otimes Z \to H \otimes Z$ are identified (the proof is similar to that of Theorem 1). It follows that homology groups defined in §§ 4 and 6 are identical with homology groups with integer coefficients.

## § 8. Cohomology groups

We shall define one more functor from the category of triangulable spaces to the category of Abelian groups, namely cohomology groups with coefficients from a group $G$. The construction of this functor, as in the case of the construction of homology groups with coefficients, consists in applying a certain functor to chain complexes before computing homology groups. Unlike the functors described in the preceding sections, the functor introduced in this section will be contravariant.

We shall start from defining the contravariant functor Hom from the category **G** into itself.

Let $G$ be a fixed Abelian group. For every group $A$ from the category **G** denote by $\mathrm{Hom}(A, G)$ the set of all homomorphisms of $A$ into $G$. With the addition of homomorphisms defined by formula (7) the set $\mathrm{Hom}(A, G)$ is an Abelian group; the role of zero is played by the zero homomorphism of $A$ into $G$. The group $\mathrm{Hom}(A, G)$ will be called the *group of homomorphisms of $A$ into $G$*.

THEOREM 1. *For any group $G$*

$$\mathrm{Hom}(Z, G) \approx G.$$

P r o o f. Since $Z$ is the free group generated by the number 1, each homomorphism of $Z$ into $G$ is uniquely determined by its value at 1, and by assigning an arbitrary element of $G$ to 1 we define a certain homomorphism. We easily verify that the correspondence between $f \in \mathrm{Hom}(Z, G)$ and $f(1) \in G$ establishes the isomorphism between these groups.

Suppose we are given a homomorphism $f\colon A \to B$. By assigning to the element $g \in \mathrm{Hom}(B, G)$ the element $f'(g) \in \mathrm{Hom}(A, G)$, defined as

$$[f'(g)](a) = gf(a) \tag{46}$$

we define a homomorphism $f'\colon \mathrm{Hom}(B, G) \to \mathrm{Hom}(A, G)$. Indeed

$$\begin{aligned}[f'(g_1+g_2)]\,(a) &= (g_1+g_2)f(a) = g_1f(a)+g_2f(a)\\ &= [f'(g_1)]\,(a)+[f'(g_2)]\,(a),\end{aligned}$$

i.e.

$$f'(g_1+g_2) = f'(g_1)+f'(g_2).$$

THEOREM 2. *By assigning the group of homomorphisms* $\mathrm{Hom}(A, G)$ *to every group* $A$, *and the homomorphism* $f'\colon \mathrm{Hom}(B, G) \to \mathrm{Hom}(A, G)$ *to every homomorphism* $f\colon A \to B$ *we define a contravariant functor from the category* $\mathbf{G}$ *into itself. Moreover, if* $f, h\colon A \to B$, *then* $(f \pm h)' = f' \pm h'$.

P r o o f. From (46) it follows immediately that $1'_A = 1_{\mathrm{Hom}(A, G)}$. If $f\colon A \to B$ and $h\colon B \to C$ are homomorphisms, then for $g \in \mathrm{Hom}(C, G)$ we have

$$[(hf)'(g)]\,(a) = ghf(a) = [h'(g)]f(a) = [f'h'(g)]\,(a),$$

i.e.

$$(hf)' = f'h'.$$

Finally,

$$\begin{aligned}[(f\pm h)'(g)]\,(a) &= g(f\pm h)\,(a) = gf(a)\pm gh(a)\\ &= [f'(g)](a)\pm [h'(g)](a),\end{aligned}$$

i.e.

$$(f\pm h)'(g) = f'(g)\pm h'(g) \quad \text{and} \quad (f\pm h)' = f'\pm h'.$$

The functor Hom can also be applied to chain complexes; it leads from the category **CC** to a certain category which will be described below.

D e f i n i t i o n 1. By a *cochain complex* we shall mean an infinite sequence

$$C^0 \xrightarrow{\delta^1} C^1 \xrightarrow{\delta^2} \ldots \to C^{n-1} \xrightarrow{\delta^n} C^n \to \ldots \tag{47}$$

of Abelian groups $C^n$ and homomorphisms $\delta^n\colon C^{n-1} \to C^n$ called *coboundary operators,* such that $\delta^n\delta^{n-1} = 0$ for $n = 2, 3, \ldots$ The cochain complex (47) will be denoted by $C = \{C^n, \delta^n\}$.

D e f i n i t i o n 2. A *cochain homomorphism* of cochain complex $C = \{C^n, \delta^n\}$ into a cochain complex $D = \{D^n, \delta^n\}$ will be defined as an infinite sequence of homomorphisms $f = \{f^n\}$, where $f^n\colon C^n \to D^n$ and the diagram

$$\begin{array}{ccc} C^n & \xrightarrow{f^n} & D^n \\ \uparrow{\scriptstyle\delta^n} & & \uparrow{\scriptstyle\delta^n} \\ C^{n-1} & \xrightarrow{f^{n-1}} & D^{n-1} \end{array}$$

is commutative, i.e.

$$f^n\delta^n = \delta^n f^{n-1} \quad \text{for} \quad n \geqslant 1. \tag{48}$$

The category whose objects are cochain complexes, and whose morphisms are cochain homomorphisms will be called the *category of cochain complexes,* and will be denoted by **CCC**. The composition of morphisms, and the identity morphism are defined as in the category **CC**.

Given a cochain complex $C = \{C^n, \delta^n\}$, the *group of n-dimensional cocycles* will be defined as the *groups* $Z^n(C) = \operatorname{Ker}\delta^{n+1}$, and the group of *n-dimensional coboundaries* will be defined as $B^n(C) = \operatorname{Im}\delta^n$ for $n \geqslant 1$; the *group of* 0-*dimensional coboundaries* $B^0(C)$ will be defined as the zero group. The identity $\delta^n\delta^{n-1} = 0$ implies that $B^n(C) \subset Z^n(C)$ for $n = 0, 1, \ldots$; the quotient group $H^n(C) = Z^n(C)/B^n(C)$ will be called the *nth cohomology group* of the cochain complex $C$.

It follows from (48) that a cochain homomorphism $f\colon C \to D$ maps cocycles into cocycles and coboundaries into coboundaries;

thus, it determines for every $n$, the *induced homomorphism* $f_*^n$: $H^n(C) \to H^n(D)$. The following theorem holds:

THEOREM 3. *By assigning the nth cohomology groups to cochain complexes, and the induced homomorphisms of those groups to cochain homomorphisms, we define a covariant functor from the category* **CCC** *to the category* **G**.

Definition 3. Two cochain homomorphisms $f, g$: $C \to D$ where $f = \{f^n\}$, $g = \{g^n\}$, $C = \{C^n, \delta^n\}$ and $D = \{D^n, \delta^n\}$ will be called *cochain homotopic*, and denoted by $f \approx g$, if there exists a sequence of homomorphisms $\Delta = \{\Delta^n\}$ where $\Delta^n$: $C^{n+1} \to D^n$ for $n = 0, 1, 2, \ldots$ such that

(49) $$f^n - g^n = \delta^n \Delta^{n-1} + \Delta^n \delta^{n+1} \quad \text{for} \quad n = 1, 2, \ldots$$

and

(50) $$f^0 - g^0 = \Delta^0 \delta^1.$$

The sequence of homomorphisms $\Delta$ will be called a *cochain homotopy* between $f$ and $g$.

We easily prove

THEOREM 4. *If cochain homomorphisms* $f, g$: $C \to D$ *are cochain homotopic, then* $f_*^n = g_*^n$ *for* $n = 0, 1, \ldots$

The functor Hom applied to a chain complex yields a certain cochain complex. Suppose we are given a chain complex $C = \{C_n, \partial_n\}$. The sequence $\{C^n, \delta^n\}$ where $C^n = \mathrm{Hom}(C_n, G)$ and $\delta_n = \partial_n'$ is a cochain complex, since $\delta^n \delta^{n-1} = \partial_n' \partial_{n-1}' = (\partial_{n-1} \partial_n)'$ is the zero homomorphism. This complex will be denoted by $\mathrm{Hom}(C, G)$. We easily verify that if $s = \{s_n\}$ is a chain homomorphism of $C$ into $D$, then $s' = \{s_n'\}$ is a cochain homomorphism of $\mathrm{Hom}(D, G)$ into $\mathrm{Hom}(C, G)$ and that the mapping defined above is a contravariant functor from the category **CC** to the category **CCC**.

Composing the functor described in Theorem 1 of § 5 with functor Hom, and then with the functor described in Theorem 3, we obtain a contravariant functor from the category **SC** to the category **G**; the group $H^n(K; G) = H^n(\mathrm{Hom}(C(K), G))$ is assigned to any simplicial complex $K$, and the homomorphism $\varphi_*^{\circ\prime n}$: $H^n(L; G) \to H^n(K; G)$ is assigned to any simplicial map $\varphi$: $K \to L$. The group $H^n(K; G)$ will be called the *nth cohomology* *group of $K$ with coefficients from $G$*, and the homomorphism $\varphi_*^{0,n}$, denoted shortly by $\varphi_*^n$, will be called the *induced homomorphism.*

EXAMPLE. We shall determine the cohomology groups with coefficients from $G$ of the simplicial complex $K$ consisting of a single vertex $p$. We know that $C_n(K) = 0$ for $n \geqslant 1$, $C_0(K) \approx Z$ and the homomorphisms $\partial_n$ are zero homomorphisms. After applying the functor Hom we obtain $\mathrm{Hom}(C_n(K), G) = 0$ for $n \geqslant 1$, while $\mathrm{Hom}(C_0(K), G) \approx G$ in virtue of Theorem 1; clearly the homomorphisms $\delta^n = \partial_n^{\text{'}}$ are zero homomorphisms. It follows that $H^n(K; G) = 0$ for $n \geqslant 1$ and $H^n(K; G) \approx G$.

The construction of cohomology groups with coefficients from $G$ for triangulable spaces, and of homomorphisms of those groups induced by continuous mappings is analogous to the construction presented in § 7. We shall merely sketch the main steps of this construction.

To begin with, we show that cohomology groups with coefficients from $G$ of a complex $K$ and of its barycentric subdivision $K'$, are isomorphic. We consider cochain complexes $\mathrm{Hom}(C(K), G)$ and $\mathrm{Hom}(C(K'), G)$ and cochain homomorphisms $\sigma^{\circ\text{'}}$: $\mathrm{Hom}(C(K), G) \to \mathrm{Hom}(C(K'), G)$ and $s$': $\mathrm{Hom}(C(K'), G) \to \mathrm{Hom}(C(K), G)$. Since Hom is a functor, $s\text{'}\sigma^{\circ\text{'}}$ is the identity; next, $\sigma^{\circ\text{'}}s\text{'}$ is cochain homotopic with the identity, since (in view of the last assertion of Theorem 3) the functor Hom maps chain homotopies into chain homotopies.

We leave it to the reader to formulate a corollary, corresponding to the corollary to Theorem 1 of § 6, which permits the assigning of a homomorphism $f^n$: $H^n(L; G) \to H^n(K; G)$ to every continuous mapping $f$: $|K| \to |L|$. Next we prove theorems corresponding to Theorems 2 and 3 of § 6 for cohomologies of simplicial complexes; using those theorems we define a functor which assigns to the triangulable space $X$ the group $H^n(X; G)$ and to the mapping $f$: $X \to Y$ the homomorphism $f_*^n$: $H^n(Y; G) \to H^n(X, G)$. One can also prove the theorem corresponding to Theorem 5 of § 6, which asserts that for a pair of homotopic mappings $f, g$: $X \to Y$ the homomorphisms $f_*^n$ and $g_*^n$ of the group $H^n(Y, G)$ into $H^n(X, G)$ are identical.

**Exercises**

**1.** Show that every simplex determines its vertices, i.e. that if $p_0 \dots p_n = q_0 \dots q_n$, then the sets $\{p_0, \dots, p_n\}$ and $\{q_0, \dots, q_n\}$ coincide.

*Hint.* Every point of the closure of a simplex $S$ other than one of its vertices is an interior point of some segment contained in $\bar{S}$.

**2.** Show that the simplicial subdivision of the set $\bar{S}$ defined in Chapter XX forms a simplicial complex, i.e. show that it contains all faces of its simplexes.

**3.** Two homomorphisms $f_1$ and $f_2$ such that

$$G_1 \xrightarrow{f_1} G_2 \xrightarrow{f_2} G_3$$

are said to form an *exact sequence*, if the group $f_1(G_1)$ is the kernel of $f_2$.

Assuming that $f_1$ and $f_2$ form an exact sequence, show that:

1°. If $G_1$ reduces to the zero element, then $f_2$ is a monomorphism.

2°. If $G_3$ reduces to the zero element, then $f_1$ is an epimorphism.

**4.** More generally, we say that a given (finite or infinite) sequence of homomorphisms is *exact*, if every successive triple of its elements forms an exact sequence.

Show that the following sequence of homomorphisms is exact:

$$0 \to G_0 \xrightarrow{i} G \xrightarrow{p} G/G_0 \to 0,$$

where $i$ denotes the identity (defined on elements of a subgroup $G_0$ of $G$) and $p$ is the natural epimorphism of $G$ onto $G/G_0$.

**5.** The Cartesian product $\prod_t F_t$ of any family of groups $F_t, t \in T$ (see Chapter IV, § 8) becomes a group if we define the addition of elements $f$ and $g$ of this product (called the *direct product* of groups $F_t$) by the formula

$$(h = f+g) \equiv \bigwedge_t (h_t = f_t + g_t).$$

Prove that

1°. The zero element of the group $\prod F_t$ is the function $f$ equal to zero for every $t \in T$.

2°. $(-f)_t = -f_t$ for every $t \in T$.

3°. The projection $f \to f_t$ is a homomorphism mapping $\prod_t F_t$ onto $F_t$.

**6.** Suppose we are given an inverse system $\{T, F, \varphi\}$ (see Chapter VII, § 5), where for each $t \in T$, $F_t$ is a group, and $\varphi_{t_0}^{t_1}$ is a homomorphism:

$$\varphi_{t_0}^{t_1}\colon F_{t_1} \to F_{t_0} \quad \text{for} \quad t_0 \prec t_1.$$

Show that the limit of this system, to be denoted by $\underleftarrow{\mathrm{Lim}}\{T, F, \varphi\}$, is a subgroup of the group $\prod_t F_t$.

**7.** Show that if $T$ is the set of all positive integers, $F_{n+1} \subset F_n$ and $\varphi_m^n$ is the identity for $n > m$, then the elements of the limit $\underleftarrow{\mathrm{Lim}}$ are constant functions with values in the intersection of the sets $F_n$.

**8.** Let $T$ be a directed set, let $F_t$ be an Abelian group for every $t \in T$, and let $\varphi_{t_0}^{t_1}$ be a homomorphism

$$\varphi_{t_0}^{t_1}\colon F_{t_0} \to F_{t_1} \quad \text{for} \quad t_0 \prec t_1 .$$

The system $\{T, F, \varphi\}$ will be called *direct*, if

$$\varphi_t^t = 1_{F_t}, \quad \varphi_{t_1}^{t_2}\varphi_{t_0}^{t_1} = \varphi_{t_0}^{t_2} \quad \text{for} \quad t_0 \prec t_1 \prec t_2 .$$

The limit $\underrightarrow{\mathrm{Lim}}\{T, F, \varphi\}$ is defined as follows: the elements of this limit are sets obtained by including to the same set two elements $x \in F_t$ and $x^* \in F_{t*}$ if there exists $t_0$ such that $t \prec t_0, t^* \prec t_0$ and $\varphi_t^{t_0}(x) = \varphi_{t*}^{t_0}(x)$.

Addition of these sets is defined as follows: let $x_j \in X_j \in \underrightarrow{\mathrm{Lim}}$, where $x_j \in F_{t_j}$ for $j = 0, 1$, let $t_0 \prec t_2, t_1 \prec t_2$ and let $[\varphi_{t_0}^{t_2}(x_0) + \varphi_{t_1}^{t_2}(x_1)] \in X_2$. In this case we put $X_2 = X_0 + X_1$.

Prove that according to the above definition, $\underrightarrow{\mathrm{Lim}}$ is an Abelian group.

**9.** Show that if $T$ is the set of all positive integers, $F_n \subset F_{n+1}$ and $\varphi_m^n$ is the identity for $m < n$, then $\underrightarrow{\mathrm{Lim}}$ consists of one-point sets, with elements from the union of sets $F_n$.

**10.** The definition of connectedness of a complex leads in a natural way to the definition of a component of a complex. Show that the homology groups of a complex are equal to direct sums of the corresponding homology groups of its components.

Prove analogous theorems for homology groups with coefficients and for cohomology groups.

**11.** The complex $K$ consists of all segments $S_1 = p_0p_1$, $S_2 = p_1p_2, \ldots,$ $S_m = p_{m-1}p_m$ and vertices $p_0, p_1, \ldots, p_m$ of the polygonal line $L$ (Fig. 22).

Let the chain $z = k_1S_1 + k_2S_2 + \ldots + k_mS_m$ be a cycle. Show that $k_1 = k_2 = \ldots = k_m = 0$.

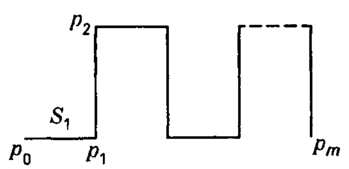



FIG. 22

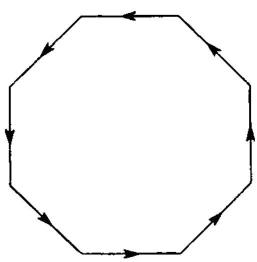


FIG. 23

**12.** The complex $K$ consists of all the segments and vertices of the polygonal line $L$ presented in Fig. 23.

We orient the segments of $L$ in the manner shown in Fig. 23. Denote the one-dimensional simplexes obtained in this way by $S_1, \ldots, S_8$.

Prove that

(a) Every chain of the form

(i) $$k \sum_{i=0}^{8} S_i$$

is a cycle.

(b) Every one-dimensional cycle of the complex $K$ is of the form (i).

(c) From (a) and (b) deduce that the first homology group of $K$ is isomorphic to the group of integers.

**13.** The complex $K$ consists of all the segments and vertices of two polygonal lines $L_1, L_2$ with one common vertex (Fig. 24). The segments of the complex

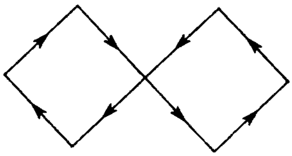


FIG. 24

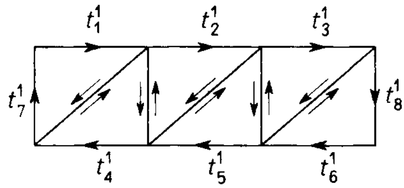



FIG. 25

$K$ are oriented as shown in Fig. 24; we denote them by $S_{1i}$ or $S_{2i}$, the indices 1 and 2 referring to lines $L_1$ and $L_2$. Put

$$z_1 = \sum_{i=1}^{4} S_{1i}, \quad z_2 = \sum_{i=1}^{4} S_{2i}.$$

Prove that

(a) Every one-dimensional cycle $z$ of the complex $K$ is of the form

(ii) $$z = k_1 z_1 + k_2 z_2.$$

(b) Every chain of the form (ii) is a cycle.

Deduce the structure of the homology groups of the complex $K$ from (a) and (b).

**14.** Fig. 25 with the sides $t_7^1$ and $t_8^1$ of the rectangle joined together represents a triangulation of the *Möbius band.* The orientation of the triangles of this triangulation, and also of the oriented segments $t_j^1$ is given in the figure. Putting

$$l = \sum_{i=1}^{6} t_i^2, \quad z = \sum_{i=1}^{6} t_i^1$$

show that

$$\partial l = z + 2t_7^1.$$

**15.** A triangulation of the *projective plane* can be obtained as follows: consider the triangulation of the rectangle presented in Fig. 26, consisting of 24 oriented triangles $t_1^2, t_2^2, \ldots, t_{24}^2$, of 12 segments $t_1^1, t_2^1, \ldots, t_{12}^1$, of interior segments, and of vertices; then join $t_1^1$ with $t_7^1$, $t_2^1$ with $t_8^1$, $t_3^1$ with $t_9^1$, $t_4^1$ with $t_{10}^1$, $t_5^1$ with $t_{11}^1$ and $t_6^1$ with $t_{12}^1$.

Instead of 12 segments $t_1^1, t_2^1, \ldots, t_{12}^1$ we obtain only 6, denoting them as before by $t_1^1, \ldots, t_6^1$.

In this manner we obtained a triangulation of the projective plane. Putting

$$l = \sum_{i=1}^{24} t_i^2, \quad z = \sum_{j=1}^{6} t_j^1$$

show that $\partial_2 l = 2z$, and that $z$ is not a boundary of any two-dimensional chain of $K$.

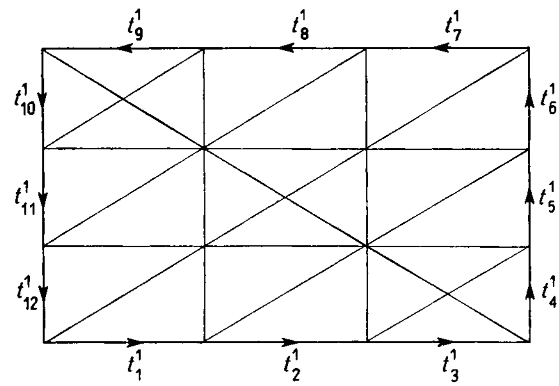



FIG. 26

*Hint.* If $\partial_2 l_1 = z$, then $l_1 = kl$.

Compute the homology groups of the projective plane and its homology groups with coefficients from the group $Z_2$ (i.e. the quotient group $Z/2Z$, where $2Z$ denotes the group of even numbers).

**16.** Show that if a group $H$ is the direct sum of $m$ copies of the group $Z$, then the group $H \otimes G$ is the direct sum of $m$ copies of the group $G$.

**17.** Prove that if $K$ is a non-empty connected complex, then $H_0(K; G) = G$.

**18.** Show that if the group $H$ is the direct sum of $m$ copies of the group $Z$, then the group $\mathrm{Hom}(H, G)$ is the direct sum of $m$ copies of the group $G$.

**19.** Show that if $K$ is a non-empty connected complex, then $H^0(K; G) = G$.

**20.** Let $\{G_0, \ldots, G_n\}$ be a system of open sets in an (arbitrary) space $X$. Let $p_0, \ldots, p_n$ be a system of points of a Euclidean space such that

$$(*) \qquad G_{i_0} \cap \cdots \cap G_{i_k} \neq 0$$

implies linear independence of points $p_{i_0}, \ldots, p_{i_k}$. The complex $N$ consisting of simplexes $p_{i_0} \ldots p_{i_k}$ such that (*) holds will be called the *nerve* of the system $\{G_0, \ldots, G_n\}$.

Prove that

1°. If $X = \overline{p_0 \ldots p_n}$ (where the vertices of the simplex in question are linearly independent), and if $G_i$ denotes the star of $p_i$, then the nerve of the system $\{G_0, \ldots, G_n\}$ is given by the complex consisting of all the faces of the simplex $p_0 \ldots p_n$.

2°. If $X$ is a compact $n$-dimensional space, then for every $\varepsilon > 0$ there exists a continuous mapping $f$ of this space into an $n$-dimensional polyhedron $N$ such that $\delta[f^{-1}(y)] < \varepsilon$ for every $y \in N$.

*Hint.* Consider the nerve of the cover given in the theorem from § 3, Chapter XIX, and the kappa mapping from Exercise 3, Chapter XX.

**21.** Let $\{G_0, \ldots, G_m\}$ and $\{H_0, \ldots, H_n\}$ be two covers of $X$ consisting of open non-empty sets. Moreover, suppose that the second cover is a refinement of the first one (i.e. for every $j \leqslant n$ there exists an $i \leqslant m$ such that $H_j \subset G_i$).

Suppose that a complex $K$ with vertices $p_0, \ldots, p_m$ is the nerve of the system $\{G_0, \ldots, G_m\}$, and a complex $L$ with vertices $q_0, \ldots, q_n$ is the nerve of the system $\{H_0, \ldots, H_n\}$. Let $f$ be the function assigning to every $j \leqslant n$ a number $f(j) \leqslant m$ such that

$$H_j \subset G_{f(j)}.$$

Show that the function $\pi$ defined by the condition

$$\pi(q_j) = p_{f(j)}$$

is a simplicial map of $L$ into $K$ (such a simplicial map is called a *projection*).

Remark. The concept of homology group for an arbitrary compact space $X$ is introduced as follows. Let $T$ be the family of all finite (open) covers of $X$. This family is directed by the relation $\prec$, where $A \prec B$ means that $B$ is a refinement of $A$. For each $A \in T$ denote by $K_A$ the nerve of $A$ (see Exercise 20), and let $\pi_A^B$ be the projection of $K_B$ into $K_A$ (this projection is a simplicial map). Finally, for a given $n$, denote by $p_A^B$ the homomorphism of the group $H_n(K_B)$ into $H_n(K_A)$ induced by $\pi_A^B$.

One can show that the system $\{T, H_n(K_A), p\}$ is an inverse system. The limit of this system is called the *nth homology group of the space X*.

In an analogous manner one defines the cohomology groups of compact spaces, using direct systems instead of inverse ones (see Exercise 8).

# LIST OF IMPORTANT SYMBOLS

# INDEX

# OTHER TITLES IN THE SERIES IN PURE AND APPLIED MATHEMATICS

Vol. 1. WALLACE — *An Introduction to Algebraic Topology*
Vol. 2. PEDOE — *Circles*
Vol. 3. SPAIN — *Analytical Conics*
Vol. 4. MIKHLIN — *Integral Equations*
Vol. 5. EGGLESTON — *Problems in Euclidean Space: Application of Convexity*
Vol. 6. WALLACE — *Homology Theory on Algebraic Varieties*
Vol. 7. NOBLE — *Methods Based on the Wiener-Hopf Technique for the Solution of Partial Differential Equations*
Vol. 8. MIKUSIŃSKI — *Operational Calculus*
Vol. 9. HEINE — *Group Theory in Quantum Mechanics*
Vol. 10. BLAND — *The Theory of Linear Viscoelasticity*
Vol. 11. KURTH — *Axiomatics of Classical Statistical Mechanics*
Vol. 12. FUCHS — *Abelian Groups*
Vol. 13. KURATOWSKI — *Introduction to Set Theory and Topology*
Vol. 14. SPAIN — *Analytical Quadrics*
Vol. 15. HARTMAN and MIKUSIŃSKI — *Theory of Lebesgue Measure and Integration*
Vol. 16. KULCZYCKI — *Non-Euclidean Geometry*
Vol. 17. KURATOWSKI — *Introduction to Calculus (2nd edition)*
Vol. 18. GERONIMUS — *Polynomials Orthogonal on a Circle and Interval*
Vol. 19. ELSGOLC — *Calculus of Variations*
Vol. 20. ALEXITS — *Convergence Problems of Orthogonal Series*
Vol. 21. FUCHS and LEVIN — *Functions of a Complex Variable, Volume II*
Vol. 22. GOODSTEIN — *Fundamental Concepts of Mathematics*
Vol. 23. KEENE — *Abstract Sets and Finite Ordinals*
Vol. 24. DITKIN and PRUDNIKOV — *Operational Calculus in Two Variables and its Applications*
Vol. 25. VEKUA — *Generalized Analytic Functions*
Vol. 26. AMIR-MOÉZ and FASS — *Elements of Linear Spaces*
Vol. 27. GRADSHTEIN — *Direct and Converse Theorems*
Vol. 28. FUCHS — *Partially Ordered Algebraic Systems*
Vol. 29. POSTNIKOV — *Foundations of Galois Theory*
Vol. 30. BERMANT — *A Course of Mathematical Analysis, Part II*
Vol. 31. ŁUKASIEWICZ — *Elements of Mathematical Logic*
Vol. 32. VULIKH — *Introduction to Functional Analysis for Scientists and Technologists*
Vol. 33. PEDOE — *An Introduction to Projective Geometry*
Vol. 34. TIMAN — *Theory of Approximation of Functions of a Real Variable*
Vol. 35. CSÁSZÁR — *Foundations of General Topology*
Vol. 36. BRONSHTEIN and SEMENDYAYEV — *A Guide-Book to Mathematics for Technologists and Engineers*
Vol. 37. MOSTOWSKI and STARK — *Introduction to Higher Algebra*